



Détection d'anomalies dans les données mesurées dans les bâtiments par association de connaissances expertes et de techniques d'apprentissage automatique

Enguerrand de Rautlin de la Roy

► To cite this version:

Enguerrand de Rautlin de la Roy. Détection d'anomalies dans les données mesurées dans les bâtiments par association de connaissances expertes et de techniques d'apprentissage automatique. Mécanique [physics]. Université de Bordeaux, 2023. Français. NNT : 2023BORD0344 . tel-04438194

HAL Id: tel-04438194

<https://theses.hal.science/tel-04438194>

Submitted on 5 Feb 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE

présentée pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE BORDEAUX

ÉCOLE DOCTORALE SCIENCES PHYSIQUES ET DE L'INGÉNIEUR (SPI)

Spécialité : MÉCANIQUE

Par **Enguerrand DE RAUTLIN DE LA ROY**

Détection d'anomalies dans les données mesurées dans les
bâtiments par association de connaissances expertes et de
techniques d'apprentissage automatique

Sous la direction de : **Laurent MORA**
Akka ZEMMARI

Encadrée par : **Thomas RECHT**

Soutenue le 27 novembre 2023

Membres du jury :

Mme. VALE Zita, Professeure des universités, Polytechnic of Porto, GECAD, Présidente
M. KUZNIK Frédéric, Professeur des universités, INSA de Lyon, Rapporteur
M. GHAMRI-DOUDANE Yacine, Professeur des universités, Université de la Rochelle, L3i, Rapporteur
M. WURTZ Frédéric, Directeur de recherche, CNRS, G2Elab, Examineur
Mme. PANNIER Marie-Lise, Maitresse de conférence, Université d'Angers, POLYTECH, Examinatrice
M. MORA Laurent, Professeur des universités, Université de Bordeaux, I2M, Directeur de thèse
M. ZEMMARI Akka, Professeur des universités, Université de Bordeaux, LaBRI, Directeur de thèse

Invités :

M. RECHT Thomas, Maître de conférence, Université de Bordeaux, I2M, Encadrant de thèse
M. BOURREAU Pierre, Docteur, NOBATEK/INEF4, Encadrant de thèse

Remerciements

Qu'est-ce qu'une thèse ? Quelques mots de Robert Jordan – un de mes auteurs favoris et père du magnifique cycle littéraire « La Roue du Temps » – me reviennent quand je pense à ma thèse : « *If a sword had a memory, it might be grateful to the forge fire, but never fond of it* ». À mon avis, il existe autant de définitions, de significations et d'interprétations qu'il y a de thèses, rendant chacune aussi unique que leurs auteurs. Bien entendu, les doctorants seuls n'ont pas l'apanage de la réponse à cette question et j'imagine que toi, qui lis ces quelques lignes, as également ton propre avis sur la question. De mon côté, tu n'auras rien de plus que cette seule et unique phrase. Cependant, lorsque le feu s'estompe, que les braises s'endorment et que le marteau s'éloigne, que reste-t-il à l'épée ? De la gratitude pour la main ferme mais chaleureuse qui l'a accompagnée, soutenue et encouragée pendant tout le façonnage !

Cette main a pris différents aspects au cours de mes quatre années de thèse.

Tout d'abord, je tiens à remercier Laurent Mora qui m'a fait confiance pour cette thèse et qui n'a pas été effrayé par ma propension, étonnamment régulière, à mettre des chats, chiens, lapins et personnages Disney en tout genre sur bon nombre de mes présentations pendant mon stage de fin d'études (et après). Merci également d'avoir été aussi patient que ferme quand il le fallait. J'ai beaucoup appris et évolué à ton contact et pas seulement sur le plan scientifique.

Merci également à Thomas Recht qui a été une source d'inspiration et de motivation pendant toutes ces années. Merci pour toutes les discussions, tant professionnelles que personnelles, que ce soit au téléphone, dans le bureau de l'ENSAM, au A11, sur le terrain de tennis ou même sur un bout de chemin du retour le soir ou encore, la veille de ma soutenance. Tu as toujours été de bons conseils et je pars avec un bon paquet d'entre eux dans mes bagages.

Enfin, merci à Akka Zemhari qui m'a accueilli à bras ouverts dans son domaine et qui a fait preuve d'une grande confiance alors même qu'au tout début, je ne savais qu'à peine coder sous Python. Merci pour ta gentillesse et ta patience tout au long de ces quatre années, j'ai découvert un univers que j'adore, et si j'en parle avec autant de passion aujourd'hui, c'est grâce à toi.

Je souhaite également remercier Pierre Bourreau pour son temps pendant nos échanges, sa gentillesse et sa présence aussi constante que rassurante. Je souhaite également remercier Muriel qui a été d'une aide conséquente, sans parler de sa patience phénoménale. Un grand merci aussi à l'ensemble des équipes GP2E et E3BUS, ce fut un plaisir de travailler comme de discuter avec vous. Merci aussi à tous mes camarades non permanents et à nos réunions mensuelohebdomadairobimensuelles du vendredi ! Merci, Huynh, Ryad, Quentin, Anass, Myriame, Zineb, Clémence, Phong, Asmae, Ana, Assia, Tosin, Horly, Férial, Jinxia, Valentin, Fatima, Maxime pour tous ces bons moments et supers échanges.

Un merci un peu plus spécial à Manon, la première personne que j'ai rencontrée sur Bordeaux, qui est également ma colocataire et sœur de cœur. À nos soirées sushi et série, à la force que tu m'as apportée ainsi qu'à Lilas, la petite furie de l'appartement. Merci à Lorena, qui prend toujours soin des autres comme le ferait une maman et Nicolas qui a été comme un grand frère. Vous étiez là quand j'ai commencé et je vous dois tellement. Merci Fanny pour ces supers moments passés à Bordeaux comme à Paris à surfer entre le « désagréable » et le « bichette ». Merci d'avoir été présente quand j'en avais besoin et merci d'être venue me voir le jour de ma soutenance. Merci Tessa pour tes encouragements et ta gentillesse, maintenant je sais qu'il y a tellement plus à y gagner à ne pas baisser les yeux en marchant dans les couloirs. Merci Benjamin d'avoir été là, pour toutes nos discussions qui étaient aussi intéressantes qu'exutoires, pour ta gentillesse, pour toute l'expérience que tu m'as transmise et que tu continues à me transmettre, et pour toutes ces super sorties dans les bars avec ou sans liqueur de melon. Merci Aurore pour ta prévenance et ta délicatesse de tout instant. Grâce à toi, j'ai pu être collègue avec un rayon de soleil l'espace d'un an. Ton arrivée sur Bordeaux a été une véritable bouffée d'air pur. Merci Nallya (et Monsieur) pour nos matinées « café FF » mais pas que, merci pour ton support sans faille, tes encouragements, toute ta gentillesse, nos appels à rallonge et ton amitié. C'est grâce à toi que je pouvais m'évader l'espace de quelques heures et recharger mes batteries. Enfin, merci à mes deux frères de cœur. Merci Enzo d'être qui tu es, haut en couleur comme en paroles, un mélange aussi formidable qu'attachant de Réunionnais-Marseillais. Tous ces samedis en centre-ville, les soirées chez vous, ou les débarquements impromptus chez moi, m'ont fait le plus grand bien. Un grand merci pour votre soutien de tout instant à toi et Cynthia. Merci Mathieu, on a commencé notre thèse ensemble et on a pu se voir grandir mutuellement sur ces quatre années. Sans ton amitié, ta gentillesse, ton soutien, nos sorties, nos soirées et nos cafés balade autour de l'ENSAM, je n'aurais pas tenu mais surtout, je ne serais pas le même aujourd'hui. Du fond du cœur je vous remercie toi et Léa, pour tout ce que vous m'avez apporté. Vous serez des parents formidables. J'ai hâte d'assister à vos soutenances !

Enfin, je souhaite remercier du fond du cœur ma famille qui n'a cessé de m'épauler depuis que j'ai commencé mon (chaotique ?) parcours d'études. Sans vous, je ne serais pas en train de rédiger ces quelques lignes. Merci Maman, Papa, mes sœurs, Constance, Aliénor, Astrid et Laurent, mon parrain, qui n'ont eu de cesse d'être à mes côtés et de m'encourager. Merci pour tout.

Merci Lily.

Résumé

L'essor croissant de systèmes et d'équipements connectés dans le bâtiment a conduit à une diffusion généralisée d'édifices appelés « bâtiments connectés » ou « bâtiments intelligents ». Ces derniers peuvent être considérés comme une association diversifiée de systèmes, logiciels, capteurs et actionneurs qui génèrent et collectent quantité de données de nature variée. Ces données peuvent être analysées et exploitées pour fournir des applications ou des services utiles à destination de l'occupant, d'un utilisateur tiers ou du bâtiment lui-même. De ce fait, leur accessibilité comme leur disponibilité sont devenu des préoccupations majeures. Cependant, en raison d'événements rares et anormaux susceptibles de se produire de leur production à leur exploitation, la quantité ou la qualité des données peut souvent devenir limitante, entravant l'utilisation efficace de tels services. Pour pallier ce problème, il devient essentiel d'identifier ces événements ou observations anormales pour espérer les corriger à terme, permettant ainsi de disposer des données en quantité et d'une qualité suffisante pour une utilisation optimale. Dans ce contexte, les techniques d'identification d'événements anormaux, également connues sous le nom de détection d'anomalies, constituent un domaine en constante évolution qui a suscité un vif intérêt et la contribution de diverses disciplines au sein de la littérature. Cette implication interdisciplinaire a conduit à un ensemble aussi vaste que varié de techniques dont le choix et l'application, souvent propres à des données ou anomalies spécifiques, complexifient la mise en œuvre.

Ces travaux de thèse se sont intéressés à l'identification d'événements ou observations anormales dans un ensemble de données et sur diverses anomalies issues de mesures dans le bâtiment, avec deux objectifs principaux. D'une part, la création d'une méthodologie de sélection de techniques de détection d'anomalies, à base d'apprentissage automatique (ou *machine learning*), adaptée aux anomalies pouvant être fréquemment rencontrées dans les mesures du bâtiment. Cette démarche vise à évaluer et classifier différents types de techniques de détection d'anomalies (p. ex. prédiction, reconstruction, classification) selon la nature des données et le type d'anomalies rencontrées. Ceci afin de mettre en lumière les avantages, inconvénients et spécificités de plusieurs méthodes relevées dans la littérature, et actuellement utilisées ou potentiellement applicables aux mesures du bâtiment. D'autre part, cette thèse a intégré une expertise métier tout au long du processus pour adapter et améliorer, lorsque cela était possible, l'entraînement, l'application des techniques ainsi que l'interprétation des résultats pouvant être apportés.

Le but était de proposer un ensemble de bonnes pratiques applicables et éprouvées dans le domaine. À titre d'exemple, un travail préparatoire a été effectué pour transformer les données d'entraînement, incluant l'ajout de métriques expertes dans le cadre d'un problème de classification.

Mots clés : Intelligence Artificielle (IA), bâtiments connectés (ou intelligents), détection d'anomalies, apprentissage automatique, connaissances expertes

Abstract

The growing prevalence of connected systems and equipment in buildings has led to widespread adoption of 'smart buildings' or 'intelligent buildings.' These can be seen as a diverse combination of systems, software, sensors, and actuators that collect and generate abundant data of various types. This data can be analysed and leveraged to provide applications or services beneficial to occupants, third-party users, or the building itself. However, this significant volume of data raises questions regarding both its accessibility and availability. Due to rare and abnormal events that may occur at various stages, from data generation to data utilization, the quantity or quality of data can often become limiting, hindering the efficient use of such services. To address this issue, it becomes essential to identify these abnormal events or observations with the hope of correcting them over time, thus enabling the availability of data in sufficient quantity and quality for optimal use. In this context, techniques for identifying abnormal events, also known as anomaly detection, constitute a constantly evolving field that has attracted significant interest and contributions from various disciplines in the literature. The involvement of multiple disciplines has resulted in a vast and varied set of techniques, but their selection and application, often tailored to specific data or anomalies, complicate implementation.

This thesis work has focused on the identification of abnormal events or observations in a dataset and various anomalies derived from measurements in buildings, with two main objectives. Firstly, the establishment of a methodology for selecting machine learning-based anomaly detection techniques tailored to the types of anomalies frequently encountered in building measurements. This approach aims to evaluate and classify different types of anomaly detection techniques (e.g., prediction, reconstruction, classification) based on the nature and type of data or anomalies encountered. This is done to highlight the advantages, disadvantages, and specificities of several methods found in the literature, which are currently in use or potentially applicable to building measurements. Secondly, this thesis has incorporated domain expertise throughout the process to adapt and enhance, when possible, the training, application of techniques, interpretation of results, and the quality of corrections that can be made. The goal was to propose a set of applicable and proven best practices in the field.

Keywords : Artificial Intelligence (AI), smart buildings, anomaly detection, machine learning, expert knowledge.

Nomenclature

Notations et abréviations

<i>Adam</i>	<i>Adaptive Moment Estimation</i>
ADEME	Agence de l'Environnement et de la Maîtrise de l'Énergie
ADL	Analyse Discriminante Linéaire
AE	<i>Autoencoder</i>
ARIMA	<i>Autoregressive Integrated Moving-Average</i>
AUC	<i>Area Under Curve</i>
AUC-ROC	<i>Area Under the Curve ROC</i>
BBC	Bâtiment Basse Consommation
BEPOS	Bâtiment à Énergie POSitive
BIPM	Bureau International des Poids et des Mesures
BIRCH	<i>Balanced Iterative Reducing and Clustering using Hierarchies</i>
CBLOF	<i>Clustering-Based Local Outlier Factor</i>
CCP	Communication en Champ Proche
CE	<i>Cross-Entropy</i>
CITEPA	Centre Interprofessionnel Technique d'Études de la Pollution Atmosphérique
COF	<i>Connectivity-based Outlier Factor</i>
CVC	Chauffage, Ventilation et Climatisation
DAE	<i>Denoising Autoencoder</i>
DBSCAN	<i>Density-Based Spatial Clustering of Applications with Noise</i>
DED	<i>Denoising Encoder-Decoder</i>
Dense-p	Réseau de neurones Dense appliqué à la tâche de Prédiction
DMA	Distance Moyenne par Assemblage
DMM	Différence MME-Mesure
DTW	<i>Dynamic Time Warping</i>
ED	<i>Encoder-Decoder</i>
EIF	<i>Extended Isolation Forest</i>
FN	Faux Négatif
FP	Faux Positif
GAN	<i>Generative Adversarial Network</i>
GES	Gaz à Effet de Serre
GIEC	Groupe d'experts Intergouvernemental sur l'Évolution du Climat
GRU	<i>Gated Recurrent Unit</i>
GTB	Gestion Technique de Bâtiment
IA	Intelligence Artificielle

IdO	Internet des Objets
IF	<i>Isolation Forest</i>
IoT	<i>Internet of Things</i>
KNN	<i>K-Nearest Neighbours</i>
LOF	<i>Local Outlier Factor</i>
LSTM	<i>Long Short-Term Memory</i>
LSTM-p	Réseau de neurones LSTM appliqué à la tâche de Prédiction
M2M	<i>Machine to Machine</i>
MAE	<i>Mean Absolute Error</i>
ML	<i>Machine Learning</i>
MME	Moyenne Mobile Exponentielle
MSE	<i>Root-Mean-Square-Error</i>
ONPE	Observatoire National de la Précarité Énergétique
PCA	<i>Principal Component Analysis</i>
RAE	<i>Recurrent Autoencoder</i>
RBF	<i>Radial Basis Function</i>
RE	Réglementation Environnementale
RED	<i>Recurrent Encoder-Decoder</i>
ReLu	<i>Rectified Linear Unit</i>
RL	Régression Logistique
RMSE	<i>Root Mean Squared Error</i>
RN	Réseaux de Neurones
RNR	Réseaux de Neurones Récurrents
ROC	<i>Receiver Operating Characteristic</i>
RT	Réglementation Thermique
SAE	<i>Sparse Autoencoder</i>
SARIMA	<i>Seasonal Auto-Regressive Integrated Moving Average</i>
SED	<i>Sparse Encoder-Decoder</i>
STL	<i>Seasonal-Trend decomposition using LOESS</i>
sub-IF	<i>Subsequence Isolation Forest</i>
sub-KNN	<i>Subsequence K-Nearest Neighbor</i>
sub-LOF	<i>Subsequence Outlier Factor</i>
SVM	Séparateurs à Vaste Marge
VAE	<i>Variational Autoencoder</i>
VARIMA	<i>Vector Auto-Regression Integrated Moving Average</i>
VN	Vrai Négatif
VP	Vrai Positif

Lettres latines

b	Valeur de biais
C	Taux de contamination d'une variable
C_{int}	Concentration de CO ₂ intérieure
d	Distance
H_{ext}	Humidité extérieure
H_{int}	Humidité intérieure
I	Intensité d'une anomalie
K_c	Nombre de clusters
K	Nombre de voisins proches
$\max_X$	Maximum d'une variable X
$\min_X$	Minimum d'une variable X
N_e	Taille des séquences d'observations en entrée d'un modèle
N_s	Taille des séquences d'observations en sortie d'un modèle
S_X	Séquence d'observations de la variable X
S_X^*	Séquence d'observations contaminée de la variable X
S_y	Séquence d'observations de la variable Y
t	Temps
T_{ext}	Température extérieure
T_{int}	Température intérieure
X	Variable
x	Observation d'une variable X
X^*	Variable contaminée
x^*	Observation contaminée d'une variable contaminée X^*
Y	<i>Variable</i>
y	Observation d'une variable Y

Lettres grecques

Δt	Intervalle de temps
α	Coefficient de lissage de la moyenne mobile exponentielle
σ_X	Écart-type de la variable X

Sommaire

Remerciements.....	ii
Résumé.....	v
<i>Abstract</i>	vii
Nomenclature	viii
Introduction générale.....	1
Chapitre 1 Données issues des bâtiments connectés.....	7
1.1. De la communication à « l'intelligence » des objets.....	9
1.1.1. L'objet connecté, un objet communicant	9
1.1.2. L'objet intelligent, un objet connecté et bien plus	9
1.2. Introduction aux bâtiments connectés et aux bâtiments intelligents.....	10
1.2.1. Définition des concepts de bâtiment connecté et de bâtiment intelligent.....	10
1.2.2. La donnée, composante centrale des bâtiments connectés comme intelligents.....	13
1.3. Bâtiments connectés : de la production des données... ..	16
1.3.1. La couche de perception : à l'origine des données.....	16
1.3.2. Mesures, capteurs et incertitudes	17
1.3.3. Les mesures issues du bâtiment : quelques notions sur les séries temporelles.....	18
1.4. ... à leur utilisation : enjeux, bénéfices	19
1.4.1. La couche d'application : des données aux services	19
1.4.2. Enjeux et bénéfices des bâtiments connectés ou intelligents	20
1.5. Les anomalies, un facteur limitant à considérer.....	22
1.5.1. De l'importance des données à leurs limites.....	22
1.5.2. Apprentissage automatique pour la détection d'anomalies, une solution ?.....	23
1.6. Conclusion	24
Chapitre 2 État de l'art sur la détection d'anomalies et les méthodes d'apprentissage automatique	27
2.1. Introduction.....	29
2.2. Qu'est-ce qu'une anomalie ?	29
2.3. Une introduction à la détection d'anomalies	31
2.3.1. Contexte et difficultés	31
2.3.2. Illustration par l'exemple.....	32
2.3.3. Domaines d'application	33
2.4. Familles et méthodes d'apprentissage automatique	35
2.4.1. Les méthodes d'apprentissage automatique.....	35
2.4.2. Estimation des performances : métriques d'évaluation.....	37

2.4.3. La famille des méthodes ou techniques de détection d'anomalies.....	40
2.5. La diversité, à la fois une force et une contrainte	49
2.6. Contexte et motivations de la thèse	52
2.7. Conclusion	53
Chapitre 3 Méthodologie de sélection de techniques de détection d'anomalies appliquée aux mesures du bâtiment	55
3.1. Contexte et démarche	57
3.2. La nature des données	57
3.2.1. La nature des données issues de mesures du bâtiment	58
3.2.2. Caractéristiques des données considérées dans la littérature	62
3.2.3. Caractéristiques considérées pour la thèse	63
3.3. Les types d'anomalies.....	64
3.3.1. Contexte	64
3.3.2. Les types d'anomalies rencontrées dans les mesures.....	64
3.3.3. La caractérisation des anomalies au travers les typologies	67
3.3.4. Mise en place d'une typologie d'anomalies adaptée	70
3.4. Méthodologie de sélection des techniques de détection	75
3.4.1. La taxonomie des problématiques rencontrées dans les mesures du bâtiment	75
3.4.2. Problématique du contexte dans les cas d'applications réels.....	77
3.5. Conclusion	79
Chapitre 4 Détection supervisée d'anomalies : application aux ouvertures de fenêtres	81
4.1. Contexte et démarche	83
4.2. État de l'art : la détection d'ouvertures de fenêtres.....	85
4.2.1. Détection ou prédiction du comportement des occupants	85
4.2.2. Modélisation des ouvertures de fenêtres.....	86
4.2.3. Apport de l'expertise métier pour l'optimisation des performances	88
4.3. Description du protocole expérimental	89
4.3.1. Processus d'évaluation des performances	89
4.3.2. Présentation des modèles.....	92
4.3.3. Configuration et entraînements des modèles	97
4.4. Description et prétraitement des données	100
4.4.1. Collecte et préparation des données.....	100
4.4.2. Analyse des données de test.....	102
4.4.3. Prétraitement des données par transformation et association	103
4.4.4. Combinaisons et associations des données	107
4.5. Résultats	107
4.5.1. Évaluation basée sur le F ₁ -score	108
4.5.2. Évaluation avec métriques additionnelles	114
4.5.3. Synthèse appliquée à la détection d'ouvertures des fenêtres	118
4.6. Conclusion	119

Chapitre 5 Détection semi-supervisée et non supervisée d'anomalies dans des mesures du bâtiment	123
5.1. Contexte et démarche.....	125
5.2. État de l'art pour le choix des techniques.....	126
5.3. Méthodologie de l'expérimentation	128
5.3.1. La génération d'anomalies.....	128
5.3.2. Modèles d'apprentissage automatique	133
5.3.3. Configuration et entraînement des modèles	142
5.3.4. Processus d'évaluation des modèles	145
5.4. Description, prétraitement des données et génération des ensembles d'évaluation	151
5.4.1. Collecte et préparation des données.....	151
5.4.2. Génération des ensembles d'évaluation anormaux.....	153
5.4.3. Transformation et association des données.....	155
5.5. Résultats	156
5.5.1. Contexte sur l'entraînement et l'exploitation des modèles	157
5.5.2. Le cas des mesures de température d'air intérieur.....	160
5.5.3. Le cas des mesures d'humidité relative d'air intérieur	165
5.5.4. Les cas des mesures de concentration de CO ₂ intérieure	166
5.5.5. Le cas des mesures de consommation électrique	172
5.5.6. Observations communes à l'ensemble des variables	175
5.5.7. Synthèse des résultats	183
5.6. Exemple d'application sur des anomalies issues de mesures	186
5.6.1. Contexte	186
5.6.2. Visualisation des résultats.....	187
5.7. Conclusion	191
Conclusion et perspectives.....	193
Publications.....	200
Bibliographie.....	201
Annexes	222
Liste des figures.....	232
Liste des tableaux.....	237

Introduction générale

Contexte et enjeux autour du bâtiment

Le secteur du bâtiment est l'un des plus importants consommateurs d'énergie en France, avec une part s'élevant à 44 % de la consommation énergétique finale (ADEME, 2023)¹ mais aussi en Europe avec une part de 40 % (Commission européenne, 2020)². Cette tendance ne semble pas sur le point de s'essouffler au regard des consommations passées, comme le présente la Figure 1 (à gauche) pour le secteur résidentiel et tertiaire, malgré une légère baisse des consommations totales à l'échelle du pays, imputable à la crise de la Covid-19. Ce secteur est par ailleurs responsable d'environ 18 % des émissions de Gaz à Effet de Serre (GES) (CITEPA, 2021)³. En outre, des analyses de cycle de vie de bâtiments révèlent que la phase d'exploitation est principalement responsable de 80 à 90 % de la consommation énergétique totale tout au long de la durée de vie d'un bâtiment, loin devant la phase de construction (10 à 20 %) et la phase de destruction (négligeable) (Ramesh et al., 2010). Face à ce constat et dans le contexte de réchauffement climatique (GIEC, 2022), de précarité énergétique (ONPE, 2023)⁴ et même de crise énergétique (liée au conflit entre l'Ukraine et la Russie), le bâtiment et son exploitation constituent de réels enjeux tant politiques, qu'économiques et scientifiques.

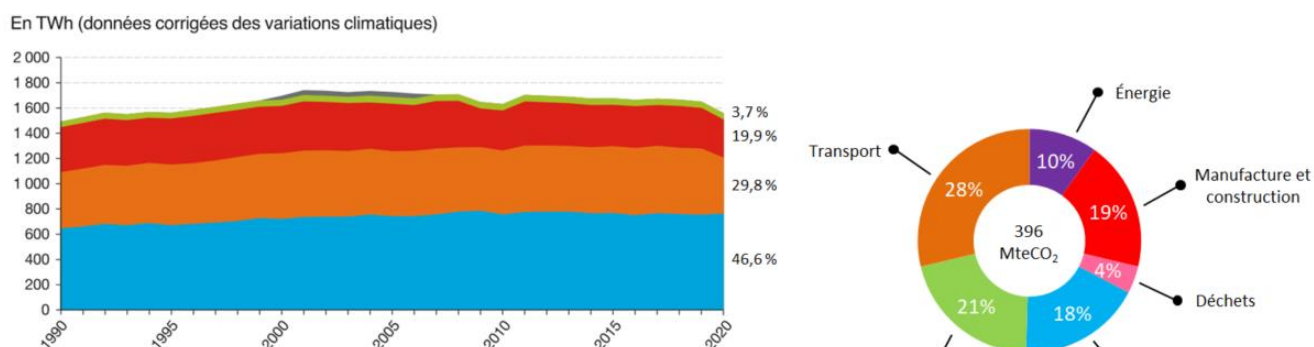


Figure 1 Gauche - Consommation finale énergétique par an et par secteur en France en TWh de 1990 à 2020 avec la répartition des consommations pour l'année 2020 (SDSE). Droite - Répartition des émissions de GES par secteur en France pour l'année 2020 (CITEPA).

¹ <https://www.ademe.fr/les-defis-de-la-transition/batiments/>

² <https://www.europarl.europa.eu/news/fr/press-room/20230206IPR72112/performance-energetique-des-batiments-neutralite-climatique-d-ici-2050>

³ <https://www.citepa.org/fr/secten/>

⁴ https://onpe.org/sites/default/files/chiffres_cles_2022_s2.pdf

Vers des bâtiments plus performants

D'un point de vue réglementaire, plusieurs directives internationales et nationales ont successivement encadré la construction des bâtiments neufs mais aussi la rénovation du parc existant avec pour objectif la neutralité carbone en Europe à l'horizon 2050 (European Climate Law, 2021). À titre d'exemple, pour les bâtiments neufs, on peut citer l'introduction en 2010 du concept de « bâtiment à consommation énergétique quasi-nulle » (Directive 2010/31/UE, 2010) sous l'appellation de bâtiment passif mais aussi de bâtiment à énergie positive (BÉPOS) avec la Réglementation Environnementale (RE) de 2020 faisant suite au bâtiment à basse consommation (BBC) de la Réglementation Thermique (RT) de 2012. Ces directives promeuvent une meilleure conception ainsi qu'une utilisation, plus intelligente et raisonnée des bâtiments afin d'atteindre de meilleures performances énergétiques tout en assurant le confort de l'occupant. Recht (2016) montre à travers une analyse comparative de plusieurs maisons à énergie positive existantes, tout en s'appuyant sur les travaux de Musall et al. (2010), que les hautes performances énergétiques atteintes par ces bâtiments se traduisent souvent par un investissement économique conséquent orienté vers l'enveloppe afin de réduire les besoins de chaleur en hiver et optimiser le confort des occupants été comme hiver (p. ex. placement des ouvrants raisonné, sur-isolation des parois, vitrages très isolants, pare soleils). Ce travail conséquent autour de l'enveloppe n'est pas le seul atout de ces bâtiments puisqu'ils sont régulièrement associés à des systèmes de distribution, de production et de stockage d'énergie automatiques et flexibles capables de gérer tant la demande que le surplus de production (Recht, 2016). Ainsi le bâtiment est pensé dès sa phase de conception afin d'être optimisé tant sur son enveloppe que les systèmes qui l'équipent. Cependant et bien que significatives et nécessaires à bien des égards, ces avancées n'apportent de réponses qu'à un seul aspect du problème, le bâtiment.

L'utilisateur entre la conception et l'exploitation du bâtiment

En effet, selon Janda (2011), ce sont les utilisateurs plutôt que les bâtiments eux-mêmes qui consomment de l'énergie. À cet égard, le rôle de l'occupant dans la phase d'exploitation du bâtiment, de ses habitudes à son confort en passant par ses activités et, par conséquent, son influence sur la consommation énergétique, est largement étudié et reconnu (Bonte et al., 2014; Delzendeh et al., 2017; Phan, 2022). Cette influence devient d'autant plus importante à mesure que les bâtiments deviennent énergétiquement performants et automatisés (Scanu et al., 2017), car toute action contraire aux usages recommandés en conception peut affecter négativement les performances du bâtiment (que ce soit par méconnaissance, erreur ou rejet de l'occupant) (Scanu, 2017). Cependant, certaines études montrent que l'implication de l'occupant dans la gestion de ses propres

consommations, par exemple grâce à un retour d'information en temps réel (Assegond, 2015; Vellei et al., 2016), peut générer des avantages significatifs. Cette évolution vers une implication active de l'occupant dans la gestion de son bâtiment marque un changement par rapport aux approches antérieures (Scanu, 2017) qui se concentraient principalement sur de la phase de conception (isolation, ponts thermiques, *etc.*), considérant souvent l'occupant comme un élément perturbateur à contrôler. Désormais, l'utilisateur est considéré comme un acteur central dans la gestion de son bâtiment, lui permettant de visualiser, contrôler et le cas échéant réduire sa consommation tout en optimisant son confort, et ainsi de s'approprier davantage son habitation.

Vers des bâtiments connectés et intelligents producteurs de données

C'est dans ce contexte que l'Internet des Objets (IdO), également connu sous le nom de *l'Internet of Things* (IoT), a pris un rôle central. Ce terme, défini en 1999 par Kévin Ashton, servait à décrire un système permettant de connecter à Internet des objets physiques (Rose et al., 2015). Système pouvant aussi se voir comme l'ajout d'une propriété de communication à ces mêmes objets. Aujourd'hui, ce concept a transformé nos objets du quotidien (téléphones, capteurs, ordinateurs, *etc.*) en un réseau d'objets interconnectés *via* internet, créant une infrastructure commune qui facilite les interactions interhumaines, inter-objets et humaine-objets (Tomat et al., 2020). Utilisés dans de nombreux domaines (industrie, médical, transport, *etc.*) et à de nombreuses fins (contrôle, économies, sécurité, bien-être, *etc.*), ces objets permettent d'améliorer et de simplifier le cadre de vie personnel ou professionnel de leurs utilisateurs (Rose et al., 2015). Tendance à laquelle le domaine du bâtiment n'échappe pas (Buckman et al., 2014). Que ce soit en raison d'obligations réglementaires, de choix écologique ou économiques, de préoccupations liées au confort, par commodité ou même par simple curiosité, l'utilisation croissante d'équipements et de systèmes connectés dans le secteur du bâtiment a conduit à la généralisation de bâtiments dits « connectés » ou « intelligents ». Ces derniers peuvent être considérés comme une combinaison complexe de systèmes, logiciels et capteurs (Aste et al., 2017) qui génèrent et collectent une quantité massive de données (température, occupation, concentration de CO₂, *etc.*). Données qui peuvent être analysées et utilisées afin de fournir des applications ou services utiles à destination de l'utilisateur, du gestionnaire ou du système de commande lui-même (Fan et al., 2015). Ces services, gourmands en données, sont variés et incluent des fonctions telles que l'automatisation et l'optimisation des systèmes de gestion du bâtiment, la visualisation des consommations ou encore la possibilité de contrôler à distance certains équipements. Ils tiennent ainsi un rôle central en assistant le bâtiment et ses usagers à relever les différents défis évoqués précédemment. Par exemple, Moudgil et al. (2023) ont montré que l'utilisation d'un système de Chauffage, Ventilation et Climatisation (CVC) optimisé

pouvait réduire la consommation énergétique de 20 à 30 %, alors que Vellei et al. (2016) ont montré l'impact positif de la visualisation en temps réel des températures intérieures sur les habitudes des occupants, leurs réponses aux changements de températures et leur confort ressenti.

Cependant, en raison d'événements rares et anormaux qui peuvent survenir à toutes les étapes, de la collecte des données à leur exploitation (tels que des défaillances de transmission, des incidents, du bruit de mesure, *etc.*), les données recueillies peuvent être d'une qualité ou en quantité insuffisante pour les applications auxquelles elles sont prévues (Erhan et al., 2021). Pour pallier ce problème, il devient essentiel d'identifier les événements ou observations anormales (Gupta et al., 2014) au sein d'un ensemble de données variées. Cette étape est essentielle pour espérer, *in fine*, les corriger et ainsi disposer de données de qualité et en quantité suffisante (Erhan et al., 2021). Par exemple, bien que peu d'études aient abordé les effets et impacts de ces événements sur les données, les travaux de Do et Cetin (2018) ont montré que l'identification et la suppression des données anormales dans un panel de maisons ont permis d'améliorer significativement, dans 20 % des cas, les performances et la fiabilité d'un modèle de prédiction de consommation énergétique.

Périmètres et objectifs des travaux de la thèse

Ces travaux de thèse s'inscrivent dans le domaine de la détection d'événements et d'observations anormales dans les mesures issues du bâtiment. Ils visent à atteindre deux objectifs majeurs. Tout d'abord, ils cherchent à développer une méthodologie de sélection de techniques de détection d'anomalies basées sur l'apprentissage automatique, spécifiquement adaptée aux données et aux anomalies provenant du domaine du bâtiment. Cette méthodologie a pour principal objectif de permettre la comparaison de plusieurs techniques de détection relevées dans la littérature, mettant en lumière leurs avantages, leurs limites, ainsi que leurs caractéristiques spécifiques dans le contexte des problématiques propres à ce domaine.

Le deuxième objectif principal de cette thèse consiste à intégrer une expertise métier de manière continue tout au long de ces travaux, afin d'optimiser l'application et les performances des techniques de détection.

En combinant ces deux objectifs, la finalité de ces travaux est de fournir des directives pratiques pour orienter le choix des techniques de détection les mieux adaptées en fonction des données disponibles ou utilisables, des types d'anomalies rencontrées ou ciblées, ainsi que des contraintes propres aux modèles envisagés.

Plan du manuscrit

Pour répondre à ces objectifs, le **premier chapitre** introduit le contexte de la thèse en présentant un état de l'art sur les « bâtiments connectés » et les « bâtiments intelligents ». Nous commençons par définir leurs concepts, rôles et fonctionnements. Ce chapitre explore le processus d'acquisition et d'utilisation des données au sein de ces bâtiments, mettant en avant la grande diversité des données recueillies et des services qui y sont liés. Nous examinons également le caractère limitant imposé par la présence d'évènements anormaux, ou anomalies, et mettons en avant le potentiel des techniques d'apprentissage automatique dans l'identification de ces événements.

Face à ce constat, le **deuxième chapitre** se consacre à un état de l'art orienté sur la détection d'anomalies. Nous introduisons plusieurs notions essentielles, allant de la définition même de ce qu'est une anomalie, aux limites et aux défis inhérents à sa détection, en passant par la diversité des domaines d'application et des familles de techniques disponibles. Ce chapitre nous permet d'aborder la complexité du choix des techniques de détection adaptées aux données issues des bâtiments, complexité due en partie à la diversité des données étudiées et à la richesse de techniques disponibles.

Le **troisième chapitre** se concentre sur l'élaboration d'une méthodologie de comparaison et de sélection de techniques de détection d'anomalies adaptée aux données et aux anomalies spécifiques du bâtiment. Nous conduisons un état de l'art pour identifier les caractéristiques inhérentes aux données et aux anomalies ayant une influence sur le choix des techniques. À partir de cette base, nous construisons une taxonomie modulable permettant d'évaluer plusieurs méthodes de détection d'anomalies sur différents cas et limites d'applications représentatifs du domaine du bâtiment, en vue de sélectionner celles qui conviennent le mieux.

Le **quatrième chapitre** vise à mettre en œuvre la méthodologie de comparaison et de sélection de techniques de détection. Nous présentons une preuve de concept employant différentes techniques d'apprentissage supervisé pour détecter un type spécifique d'anomalies étiquetées, en l'occurrence la détection d'ouvertures de fenêtres, un problème de classification binaire similaire à la détection d'anomalies. Ce chapitre souligne également l'importance de l'apport d'une expertise métier dans le processus de détection d'évènements anormaux.

Le **cinquième et dernier chapitre** illustre l'ensemble de la méthodologie à travers une application sur une étude de cas, en utilisant des techniques d'apprentissage semi-supervisé et non supervisé. Cette étude de cas inclut plusieurs types d'anomalies différents recensés dans la littérature, qui ont été artificiellement injectés dans différentes

mesures. De plus, nous présentons un cas portant sur des anomalies naturelles issues de mesures réelles.

Enfin, nous abordons les apports et les limites de ces travaux pour conclure ce mémoire, en formulant une conclusion générale et en évoquant des perspectives pour des recherches futures.

Chapitre 1

Données issues des bâtiments connectés

Ce chapitre a pour but de présenter le contexte des travaux de thèse. Dans un premier temps, quelques notions fondamentales seront introduites sur les objets et les bâtiments dits « connectés » ou « intelligents » (sections 1.1 et 1.2). Ces deux premières sections visent à définir leur concept, leur rôle et leur fonctionnement. Ensuite, la production des données au sein de ces bâtiments sera discutée, en détaillant le processus d'acquisition et en mettant en avant la diversité des données générées (section 1.3). Après leur production, la section 1.4 traitera de leur utilisation par le biais d'applications associées, et mettra en évidence le rôle essentiel de la donnée pour le bâtiment. Enfin, la section 1.5 abordera les limites inhérentes aux données récoltées, prenant la forme d'évènements anormaux et indésirables. Le chapitre se terminera par une synthèse dans la section 1.6.

« Une civilisation peut tout autant se noyer dans ce qu'elle sait que dans ce qu'elle ignore. »

Mappo à Icarium

Extrait traduit du Livre des Martyrs, les portes de la maison des morts, de Steven Erikson

Sommaire du chapitre 1

1.1. De la communication à « l'intelligence » des objets	9
1.1.1. L'objet connecté, un objet communicant	9
1.1.2. L'objet intelligent, un objet connecté et bien plus	9
1.2. Introduction aux bâtiments connectés et aux bâtiments intelligents	10
1.2.1. Définition des concepts de bâtiment connecté et de bâtiment intelligent	10
1.2.2. La donnée, composante centrale des bâtiments connectés comme intelligents	13
1.3. Bâtiments connectés : de la production des données... ..	16
1.3.1. La couche de perception : à l'origine des données	16
1.3.2. Mesures, capteurs et incertitudes	17
1.3.3. Les mesures issues du bâtiment : quelques notions sur les séries temporelles	18
1.4. ... à leur utilisation : enjeux, bénéfices.....	19
1.4.1. La couche d'application : des données aux services	19
1.4.2. Enjeux et bénéfices des bâtiments connectés ou intelligents	20
1.5. Les anomalies, un facteur limitant à considérer	22
1.5.1. De l'importance des données à leurs limites	22
1.5.2. Apprentissage automatique pour la détection d'anomalies, une solution ?.....	23
1.6. Conclusion.....	24

1.1. De la communication à « l'intelligence » des objets

Avant d'aborder les bâtiments connectés ou intelligents, concepts clés de cette thèse, il est intéressant d'examiner au préalable les objets dits connectés ou intelligents. Bien que le terme « objet » puisse désigner un appareil ou un équipement, il partage l'idée centrale de connectivité et d'intelligence avec les bâtiments. Cette section pose donc les bases visant à définir les bâtiments connectés (ou intelligents).

1.1.1. L'objet connecté, un objet communicant

La première utilisation d'un objet connecté remonterait au début des années 1980 sur le campus de l'université de Carnegie Mellon à Pittsburg, en Pennsylvanie, aux États-Unis. Des étudiants auraient connecté un distributeur de boissons au réseau de l'université afin de pouvoir visualiser en temps réel la disponibilité des boissons. Cette anecdote illustre l'objectif premier des objets connectés : la collecte d'informations à distance. Le terme « connecté » traduit alors l'ajout d'une fonction de **communication** à un objet. Les objets connectés peuvent donc se définir comme des objets collectant (capteurs) ou produisant (actionneurs) des données et disposant d'un système de transmission (Hoffman and Novak, 2015). Outre Internet, plusieurs types de réseaux existent et sont également utilisés comme système de transmission de données, tels que le *Machine to Machine* (M2M), la *Communication en Champ Proche* (CCP), le *Bluetooth*, etc. (Sethi and Sarangi, 2017). Actuellement, de nombreux objets tels que les smartphones, les capteurs, les caméras, les équipements électroménagers, et autres objets pouvant se connecter à Internet ou à un autre réseau pour transmettre des données sont autant d'exemples d'objets connectés potentiels. Appliqué au domaine du bâtiment, des thermostats (capteurs) qui affichent la température d'une pièce sur un smartphone ou des compteurs d'électricité de type *Linky* (capteurs et actionneurs) sont des exemples d'objets connectés.

1.1.2. L'objet intelligent, un objet connecté et bien plus

Malgré leur utilisation fréquente comme synonymes, les *objets intelligents* se distinguent des *objets connectés* de par la notion même d'intelligence qui leur est associée. Il est important de noter qu'il n'existe pas consensus clair quant à la signification exacte de l'adjectif « intelligent » dans ce contexte, puisqu'il s'agit d'un terme évolutif qui n'a cessé de gagner ou de perdre en attributs au cours du temps ou des auteurs, bien que certaines caractéristiques communes demeurent (Nihtianov and Luque, 2018). Ces caractéristiques correspondent à la définition courante du terme « intelligent » apportée par le dictionnaire Le Robert⁵ à savoir « *qui a la faculté de connaître et de comprendre* ». Selon Bourdon (2021), un objet est considéré comme intelligent s'il est capable d'analyser

⁵ <https://dictionnaire.lerobert.com/definition/intelligence>

les informations des données qu'il récolte afin de s'adapter ou de prendre des décisions en toute autonomie. Une définition similaire est apportée par Nihtianov et Luque (2018), proche de celle de McFarlane et al. (2003) qui décrivent l'intelligence d'un objet (ici un capteur) comme sa capacité à percevoir, transmettre, accumuler, analyser des informations et prendre des décisions. Ainsi, un objet intelligent présente des facultés **d'adaptation, d'autonomie** et de **communication**. Il est à noter que ces fonctions n'ont pas nécessairement besoin d'être intégrées à l'objet lui-même, mais peuvent être déployées en modules distincts ou prendre la forme de services à distance (Bourdon, 2021). En ce sens, un capteur d'intrusion est un objet intelligent, au même titre qu'un smartphone.

Les objets connectés et les objets intelligents connaissent une popularité croissante depuis plusieurs années (Tomat et al., 2020), en raison de divers facteurs, notamment la diminution des coûts de production des capteurs et des dispositifs électroniques, l'amélioration de la connectivité et des protocoles de communication ainsi que le développement de logiciels dédiés et optimisés (Bourdon, 2021; Hammi, 2019). Bien que leurs objectifs diffèrent, les objets connectés et intelligents reposent tous deux sur une caractéristique commune : la communication et par conséquent, la **production de données**. Aujourd'hui, avec des dizaines de milliards d'objets sur le marché et dans nos habitations, c'est tout autant de données potentiellement exploitables pour qui, ou quoi, sait en tirer profit (Bansal et al., 2021), avec des applications dans divers domaines tels que l'agriculture, la santé, l'automobile et le bâtiment (Bashir et al., 2022).

1.2. Introduction aux bâtiments connectés et aux bâtiments intelligents

L'introduction des concepts de connectivité et d'intelligence a posé les bases de ces deux domaines, bien que leurs applications puissent varier. De même, nous distinguons les bâtiments « connectés » des bâtiments « intelligents » en raison de leurs fonctions distinctes, que nous examinerons dans cette partie.

1.2.1. Définition des concepts de bâtiment connecté et de bâtiment intelligent

En France, les termes « bâtiments connectés » et « bâtiments intelligents » sont régulièrement utilisés comme synonymes alors qu'ils décrivent deux concepts sensiblement différents. Afin de clarifier leur utilisation ultérieure dans la thèse, cette partie a pour but de les introduire et de les définir.

1.2.1.1. Le bâtiment connecté, un producteur d'informations

Pour définir un bâtiment connecté (resp. une maison, un habitat, *etc.*), une première approche peut consister à le considérer comme l'ensemble de l'écosystème formé par les objets connectés en son sein (Nabil, 2018). Plus précisément, Josse (2017) définit un bâtiment connecté comme un bâtiment composé de capteurs pouvant mesurer et enregistrer des données, parfois associés à des actionneurs permettant d'interagir en temps réel avec les systèmes du bâtiment. Ainsi, les occupants peuvent, grâce à des équipements connectés et des logiciels dédiés (p. ex. capteurs, smartphones, tablettes), obtenir des informations sur leur consommation énergétique ou la qualité de l'air intérieur, modifier la température d'une pièce, contrôler les volets, voire, dans certains cas avancés, planifier l'utilisation d'équipements électroménagers (p. ex. machine à laver) afin d'améliorer leur confort ou réduire leurs dépenses énergétiques. Les données issues des capteurs ou des actionneurs peuvent être stockées localement ou à distance (p. ex. sur des serveurs externes ou dans le *cloud*) pour une utilisation en temps réel ou ultérieure, permettant leur visualisation ou leur analyses (Heiding et al., 2023). Par conséquent, un bâtiment connecté est avant tout un bâtiment **communicant** pouvant, parfois, être **actif** par action de l'occupant. Étant donné que cette thèse se concentre essentiellement sur l'analyse de données issues de mesures du bâtiment, collectées au moyen d'objets connectés, il apparaît essentiel de considérer l'accessibilité des données. Aussi, on définira un bâtiment connecté comme « *une association diversifiée de systèmes, de logiciels, d'actionneurs et de capteurs, dont les données sont communiquées et accessibles* ».

1.2.1.2. Le bâtiment intelligent, à la fois producteur et consommateur d'informations

Le concept de bâtiment intelligent, n'est pas récent puisqu'il semble trouver ses origines aux alentours des années 80 (Sinopoli, 2010) sous l'appellation anglaise de '*intelligent building*'. Appellation qui évoluera par la suite, aux alentours de 2011, en '*smart building*' afin de tenir compte des avancées et apports technologiques (Buckman et al., 2014; Karimi et al., 2021). En dépit du sens attribué dans la littérature anglaise, une distinction similaire à celle observée pour les objets connectés et intelligents peut être effectuée entre les bâtiments dits connectés et ceux dits intelligents. Dans le cas d'un bâtiment intelligent, les données issues des capteurs ou des actionneurs ne sont pas seulement récoltées et accessibles, elles sont utilisées afin d'influer ou optimiser automatiquement le confort des occupants comme leurs consommations énergétiques (Karimi et al., 2021). Cette distinction dans l'utilisation des données semble également être à l'origine de la différenciation entre les termes anglais '*intelligent building*' et '*smart building*' dans la littérature. Bien qu'actuellement mieux définies, les différences entre ces deux appellations ont pendant longtemps été floues et restent confuses (Buckman et al., 2014). Pour illustrer ce point, Sovacool et Furszyfer Del Rio (2020), comme d'autres auteurs, montrent que sous l'appellation commune de « *smart* » dans la littérature, se

cachent en réalité différents degrés d'automatisation, d'optimisation et de services. Cette confusion est d'autant plus importante dans la littérature française en raison de la traduction conjointe de '*intelligent building*' et de '*smart building*' en « bâtiment intelligent » (Josse, 2017). D'après Buckman et al. (2014), la principale distinction qui peut être faite aujourd'hui entre les bâtiments définis comme intelligents au sens '*intelligent building*' et ceux au sens '*smart building*' tient du fait que les premiers réagissent automatiquement aux sollicitations ou commandes des occupants, tandis que les seconds s'adaptent en fonction des informations ou des mesures dont ils disposent sur une période de temps donnée. Les '*intelligent buildings*' sont considérés comme réactifs contrairement aux '*smart buildings*' qui sont considérés comme adaptatifs.

Alors que les bâtiments connectés sont avant tout définis comme communicants et actifs par contrôle de l'occupant – avec pour principale finalité d'informer ce dernier, et le cas échéant, de permettre un contrôle à distance grâce aux données collectées par les capteurs et les systèmes qui le composent –, les bâtiments intelligents, quant à eux, communiquent, apprennent, réagissent ou s'adaptent automatiquement à leurs occupants. Ils sont donc **communicants**, **autonomes** et **réactifs** ou **adaptatifs**. Aussi, on définira un bâtiment intelligent comme « *une association diverse de systèmes, logiciels, actionneurs et capteurs qui forment un ensemble autonome, réactif ou adaptatif et dont les données sont communiquées et accessibles* ».

Malgré leurs différences évidentes dans les définitions apportées, le caractère communiquant de ces bâtiments est une composante essentielle à prendre en compte. Tant les bâtiments connectés que les bâtiments intelligents génèrent une quantité massive de données variées (Fan et al., 2015). Ainsi, en se basant sur les définitions fournies et en excluant toute considération sur l'utilisation des données, un bâtiment intelligent peut être considéré, au moins, comme un bâtiment connecté. Par conséquent, pour la suite de ces travaux, et uniquement si l'accent est mis sur les données, le terme de *bâtiment connecté* sera utilisé pour faire référence à un bâtiment dont les données sont accessibles.

1.2.1.3. Illustration par l'exemple

Pour illustrer les différentes définitions de bâtiment abordées précédemment, prenons l'exemple d'un studio habité par un locataire en période de chauffe, un mois de décembre. Le locataire, un doctorant studieux, travaille en semaine et n'est généralement pas chez lui de 8h à 18h. Sensible aux questions environnementales et énergétiques, il décide d'appliquer deux profils de chauffage afin de réduire sa consommation sans dégrader son confort : il maintient une consigne de 19 °C s'il est présent et la réduit à 17 °C s'il est absent. À des fins d'illustration, différents profils de bâtiments sont présentés dans la Figure 2. Les niveaux 0 et 1 correspondent à des bâtiments ayant peu ou pas d'objets

connectés exploitables, tels qu'une télévision intelligente, et servent de référence. Dans ces bâtiments, les occupants ne peuvent ni visualiser leurs consommations, ni obtenir d'informations sur leur confort, ni contrôler à distance le bâtiment. Ainsi, un tel bâtiment est dit « primitif » pour le niveau 0 et « isolé » pour le niveau 1. Les niveaux 2, 3 et 4, en revanche, correspondent aux bâtiments connectés et intelligents tels que définis précédemment.

Dans un logement connecté (niveau 2, selon cette thèse), l'occupant détermine lui-même le planning de chauffage ou ajuste simplement le chauffage au grès de ses départs ou retours (selon les équipements installés). Dans un bâtiment intelligent au sens de '*intelligent building*' (niveau 3), le chauffage s'ajuste automatiquement une fois le départ ou le retour du locataire avéré. Enfin, dans un bâtiment intelligent au sens du '*smart building*' (niveau 4), le chauffage s'ajuste automatiquement en prévision du départ du locataire ou de son retour par un retrait progressif du réduit.

Ainsi, le bâtiment connecté met l'occupant au service de son propre confort et de ses économies d'énergie, tandis que le bâtiment intelligent se met au service du confort de l'occupant et de ses économies.

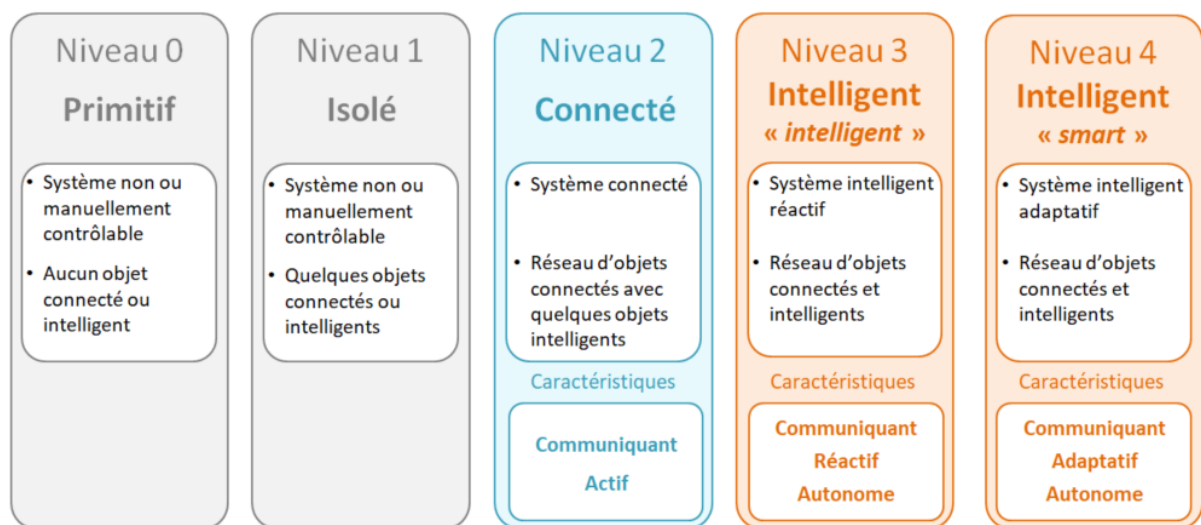


Figure 2 Illustration de différents profils de bâtiment répartis en différents niveaux selon l'évolution de leur connectivité/intelligence. Inspirée des travaux de Buckman et al. (2014), Karimi et al. (2021) et Sovacool et Furszyfer Del Rio (2020).

1.2.2. La donnée, composante centrale des bâtiments connectés comme intelligents

Comme introduit précédemment, les bâtiments connectés et intelligents peuvent être considérés comme des écosystèmes producteurs de données. Données qu'il est possible d'exploiter pour proposer divers services, tels que la visualisation et l'analyse des

consommations ou le contrôle automatique et optimisé des systèmes CVC (Tang et al., 2019). Étant donné que les bâtiments peuvent être constitués d'une multitude d'objets connectés, la collecte et l'exploitation des données reposent sur une structure permettant à des d'objets ou systèmes hétérogènes, utilisant divers protocoles de communications, de communiquer voire d'interagir entre eux ou avec des serveurs distants (p. ex. un capteur de luminosité connecté à des volets) (Park and Rhee, 2018). Cette structure peut être expliquée en détaillant les étapes suivies par le flux de donnée, de sa production (p. ex., par un thermostat d'ambiance dans une pièce) à son utilisation (p. ex., la visualisation de la température ambiante). Chaque étape constitue alors une couche ou un bloc constitutif d'un ensemble qu'on appelle la structure, l'architecture ou même l'infrastructure réseau. À l'aide de cette structure, il devient possible de comprendre le fonctionnement, parfois complexe, des bâtiments connectés et le processus de transformation de l'information, de donnée à service.

La structure d'un bâtiment connecté, tout comme celle de nombreuses autres appliquées à différents domaines, repose sur les nombreuses représentations analogues utilisées pour décrire l'architecture de l'Internet des Objets (Nandan Mohanty et al., 2022). Il n'existe pas de représentation unique et universellement admise d'architecture de l'Internet des Objets (Sethi and Sarangi, 2017). Les représentations les plus simples sont généralement composées de trois couches, tandis que les plus détaillées peuvent dépasser les sept couches. Un constat identique, bien que plus modéré, peut être appliqué aux représentations des structures des bâtiments connectés. Ces représentations oscillent généralement entre trois (Park and Rhee, 2018; Tomat et al., 2020) et quatre couches (Bashir et al., 2022) suivant le degré de détails attendu. À des fins d'illustration, une représentation simplifiée de cette architecture est proposée dans la Figure 3. Cette architecture, adaptée aux bâtiments, est composée de quatre couches par lesquelles les données transitent : perception, réseau, traitement et applicative.

L'ensemble des objets connectés installés dans le bâtiment constitue la couche de *perception*, qui est dédiée à l'acquisition ou à la production de données. La couche *réseau* permet aux objets de communiquer et de transmettre les données récoltées *via* différents systèmes de transmission (comme la CCP ou le Wifi) (Hassan, 2020), alors que la couche de *traitement* permet le stockage et le prétraitement des informations récoltées, que ce soit en local, à distance ou sur le *cloud* (telle que l'extraction, la mise en forme ou la transformation des données) (Trappey et al., 2017). Il convient de noter que ces deux couches peuvent parfois être associées en une seule dans les représentations à trois couches (Sethi and Sarangi, 2017). La dernière couche, dite *applicative*, englobe toutes les applications rendues possibles par l'utilisation des données récoltées par les objets connectés. Ces applications sont variées et vont de la visualisation des données à des systèmes de contrôle et de gestion de l'énergie en temps réel ou de détection d'anomalies.

Cependant, contrairement aux bâtiments intelligents, les applications proposées par les bâtiments connectés sont souvent restreintes à la visualisation, à des analyses *a posteriori* ou au contrôle humain à distance, bien que des applications plus élaborées puissent être mises en place de manière isolée (p. ex., une gestion automatique des ombrages selon l'ensoleillement). Par ailleurs, ces structures ne reposent pas sur une application unique et peuvent prendre différentes formes et être implémentées à différents échelles dans un bâtiment. Par exemple, une Gestion Technique de Bâtiment (GTB) permet de contrôler et de superviser l'ensemble des installations techniques d'un bâtiment (Linder et al., 2017), tandis que des systèmes comme *Google Nest* ou *Alexa Smart Home* permettent de contrôler des lumières, des caméras ou des prises compatibles offrant ainsi un déploiement plus restreint (Tomat et al., 2020).

Cet état des lieux du processus de fonctionnement des bâtiments connectés, de l'acquisition à l'utilisation des données en passant par leur transmission et leur traitement, met en lumière le caractère essentiel des données pour cet écosystème. Sans données, les nombreuses applications et services qui en dépendent ne seraient pas possibles. Ainsi, pour mieux comprendre les enjeux liés à ce domaine, une attention particulière va être apportée dans la suite de ce chapitre aux processus de production et d'exploitation des données, en examinant plus en détails les couches de perception et d'application. L'objectif est de mieux comprendre la nature et le rôle des données : d'où proviennent-elles et comment sont-elles obtenues ? À quelles fins sont-elles utilisées, et de quelle manière ?

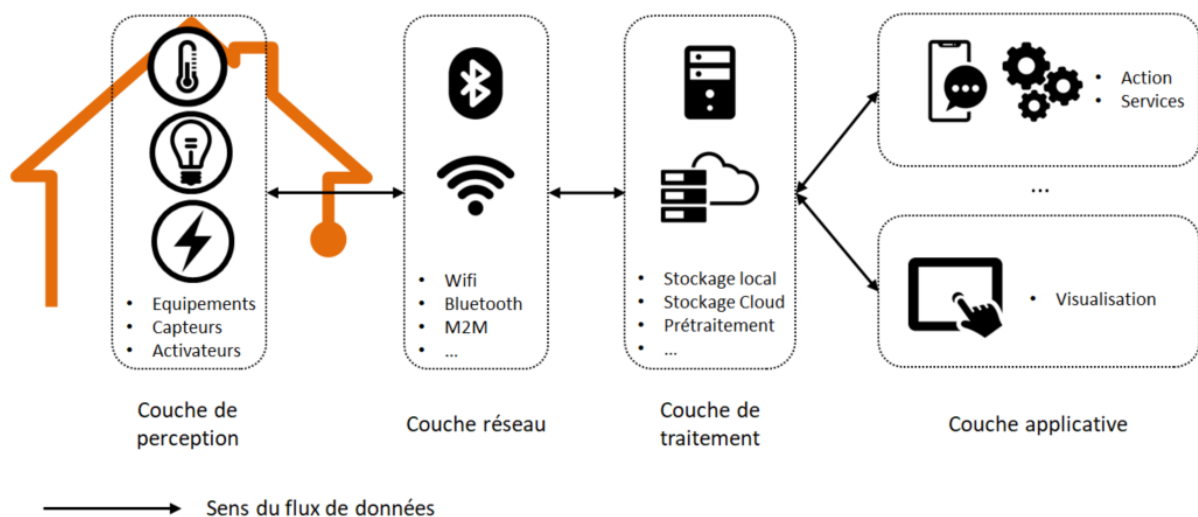


Figure 3 Exemple de structure simplifiée en quatre couches, perception, réseau, traitement et applicative représentant l'écosystème des bâtiments connectés et des bâtiments intelligents.

1.3. Bâtiments connectés : de la production des données...

Comme précédemment exposé, les bâtiments connectés et plus particulièrement les objets, capteurs et actionneurs les composants, génèrent une quantité massive de données, majoritairement issues de mesures, et dont l'extraction d'information peut être utilisée afin de fournir divers services. Cette partie a pour but d'offrir un aperçu sur les origines variées de ces données mais aussi d'en définir et détailler le processus d'acquisition.

1.3.1. La couche de perception : à l'origine des données

La couche de perception, parfois appelée couche physique ou couche des objets (de l'anglais « *things* »), est constituée de différents dispositifs, équipements, capteurs et actionneurs connectés. Selon Moreno et al. (2014), Nabil (2018) et Verma et al. (2019), cet ensemble hétéroclite d'objets connectés peut être classé en plusieurs catégories selon leur champ d'action et leur utilisation. La Figure 4, tirée des travaux de Nabil (2018) en propose une approche simplifiée en six catégories. La catégorie « *smart grid* », pouvant se traduire de l'anglais par « *réseau intelligent* », regroupe l'ensemble des compteurs donnant accès aux consommations et aux productions d'énergie, le cas échéant (p. ex., électrique, gaz). Un compteur d'électricité *Linky* peut être rangé dans cette catégorie, au même titre qu'un compteur mesurant l'autoproduction d'énergie de panneaux photovoltaïques. La catégorie *confort* regroupe les objets qui mesurent et contrôlent la qualité de l'air. Cette catégorie inclut aussi les capteurs d'ambiance intérieure pouvant suivre la concentration de CO₂, la température ou l'humidité ambiante, ainsi que les stations météorologiques extérieures pouvant également mesurer la vitesse du vent ou l'ensoleillement. Dong et al. (2019) proposent d'ailleurs une étude détaillée des éléments constitutifs de cette catégorie. La catégorie des *produits blancs* regroupe, quant à elle, les différents appareils et équipements électroménagers destinés à couvrir les besoins domestiques tels que les tables de cuisson, les fours, les réfrigérateurs, les cafetières, etc. La catégorie de la *domotique* regroupe l'ensemble des équipements, capteurs ou actionneurs permettant d'automatiser, de programmer, de contrôler le bâtiment ou d'apporter des informations sur l'occupant. Elle comprend des capteurs de présence, de lumière ou un système de pilotage de l'ouverture et la fermeture des ouvrants (ou volets). La catégorie de la *sécurité* regroupe les objets liés à la sécurité du bâtiment et de ses occupants, tels que les caméras, les systèmes antieffraction ou les détecteurs de fumées. Enfin, la dernière catégorie regroupe, sous le terme de *smartphone*, les appareils rendant possible les interactions entre le bâtiment et l'occupant ou entre les objets connectés et l'occupant, tels que les smartphones, les tablettes ou ordinateurs.

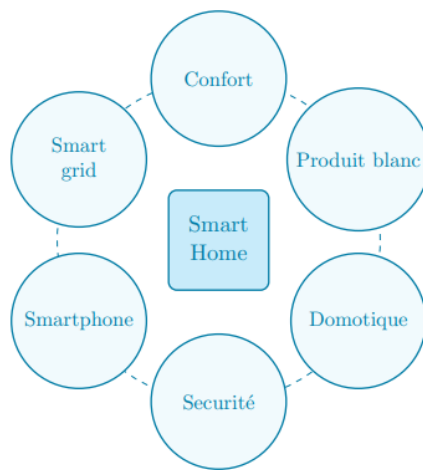


Figure 4 Exemple de classement en six catégories des objets connectés constitutifs d'un bâtiment connecté ou intelligent. Issue de Nabil (2018).

Il est important de noter que l'écosystème composant la couche d'objets est en général constitué majoritairement d'éléments des catégories confort, domotique, sécurité et *smart grid*, qui contiennent essentiellement différents capteurs et actionneurs (Ding et al., 2011; Moudgil et al., 2023). Cette couche de perception est donc en grande partie axée sur l'acquisition ou la production d'une *grande quantité de données variées*, principalement issues de **mesures**, qui pourront être exploitées par la suite. Cette diversité est illustrée par les travaux de Dong et al. (2019), bien qu'ils se concentrent particulièrement sur les aspects liés au contrôle de l'ambiance et aux économies d'énergies.

1.3.2. Mesures, capteurs et incertitudes

D'après le bureau international des poids et des mesures (BIPM⁶), une mesure est le résultat d'un processus de mesurage, qui comprend toutes les opérations visant à évaluer ou à déterminer la valeur d'une grandeur physique (p. ex., une température ambiante ou la concentration de CO₂) (Asch, 2010). En d'autres termes, une mesure est une donnée (ou une valeur) associée à une grandeur physique, obtenue à partir d'instruments de mesure appropriés, parfois appelés chaînes de mesure. Un capteur est, à ce titre, un instrument de mesure. Il en existe de différents types couramment utilisés dans notre quotidien (p. ex., capteurs de température, de pression, de mouvement) pour lesquels Javaid et al. (2021) proposent une étude détaillée. À des fins d'illustration, la Figure 5 présente une représentation schématique d'une chaîne de mesure et du processus de mesurage. Une chaîne de mesure est constituée de l'ensemble des éléments nécessaires à l'évaluation de la valeur de la grandeur physique. Elle comprend *a minima* trois éléments : un capteur (actif ou passif), un conditionneur de signal et une unité d'utilisation. Le capteur traduit les variations de la grandeur physique mesurée en un signal d'une autre grandeur,

⁶ https://www.bipm.org/documents/20126/2071204/JCGM_200_2012.pdf/f0e1ad45-d337-bbeb-53a6-15fe649d0ff1

souvent électrique. Dans le cas d'un capteur passif (i) ce signal est ensuite converti à l'aide d'un conditionneur pour le rendre compatible avec l'unité d'utilisation. Dans le cas d'un capteur actif (ii), le signal est directement exploitable par l'unité d'utilisation (Asch, 2010). L'unité d'utilisation a, quant à elle, pour rôle de rendre le signal obtenu exploitable en permettant de le visualiser ou de l'enregistrer.

Cependant, malgré les récents progrès technologiques qui ont permis d'obtenir des résultats de plus en plus fiables, de nombreuses erreurs ou incertitudes peuvent entacher les mesures obtenues. La valeur réelle de la grandeur physique mesurée ne peut être connue, elle ne peut qu'être estimée. En cause, une incertitude de mesure toujours présente et due aux imperfections des chaînes de mesures, provoquant une dégradation plus ou moins conséquente du signal au cours de son traitement (Asch, 2010). Cette incertitude peut en partie dépendre de la qualité de la chaîne de mesure, de son étalonnage, mais aussi de l'environnement avoisinant. Il est donc important de toujours prendre en considération le fait qu'un instrument de mesure ou un capteur fournit une mesure qui n'est en aucun cas la mesure de la grandeur physique en question. Le résultat du processus de mesure ne constitue qu'une représentation plus ou moins précise de la « réalité » ou de ce qui est considéré comme « vrai ». Cependant, pour des raisons de simplicité, la notion de *mesures réelles* ou *mesures vraies* sera employée dans la suite de ces travaux de thèse pour faire référence aux mesures récoltées *in situ*, qui sont les seules représentations accessibles de la réalité, aussi approximatives soient-elles.

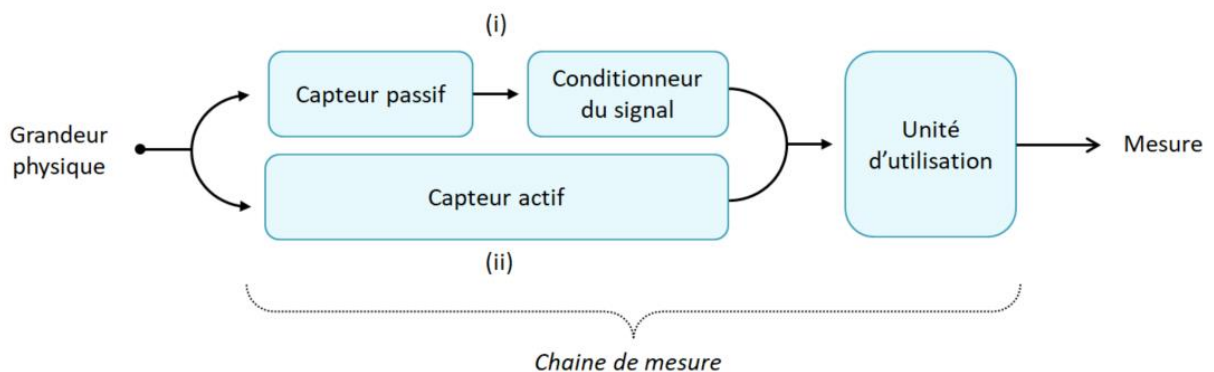


Figure 5 Schématisation d'une chaîne de mesure. Chaîne constituée d'un capteur passif (i) ou d'un capteur actif (ii).

1.3.3. Les mesures issues du bâtiment : quelques notions sur les séries temporelles

Les capteurs du bâtiment mesurent et traduisent des grandeurs physiques en données. Plus précisément, ils détectent et relaient les événements ou changements dans ces grandeurs physiques au cours du temps (Javaid et al., 2021). Ces capteurs transmettent ces évolutions sous forme d'une **série temporelle**, c'est-à-dire une séquence ordonnées

de données qui représente certaines caractéristiques, attributs ou propriétés d'éléments mesurés (p. ex., objets, événements, observations) (Tan et al., 2014).

Dans le domaine du bâtiment et de l'énergétique du bâtiment, les séries temporelles sont composées de grandeurs physiques hétérogènes définies comme des *variables* telles que la température intérieure, la concentration de CO₂, le nombre d'occupants, ou encore le rayonnement extérieur. Ces variables sont collectées de manière continue à pas de temps régulier ou irrégulier par des capteurs, puis stockées dans des bases de données (Djenouri et al., 2020) (*cf.* partie 1.3.2). Chaque point de données, ou observation, correspondant à une variable est défini par une valeur associée à un horodatage, indiquant l'heure de collecte (Brockwell and Davis, 2010). Une série temporelle composée d'une seule variable est dite *univariée* (p. ex., mesure de température seule). En revanche, si elle est composée de plusieurs variables interdépendantes, elle est alors dite *multivariée* (Cheng et al., 2009) comme une série de mesures de température et d'humidité relative collectées par un même capteur. La Figure 6 présente une série univariée avec des mesures de concentration intérieure de CO₂ (i) ainsi qu'une série temporelle multivariée avec des mesures de température, d'humidité et de concentration intérieure de CO₂ provenant d'un même capteur (ii).

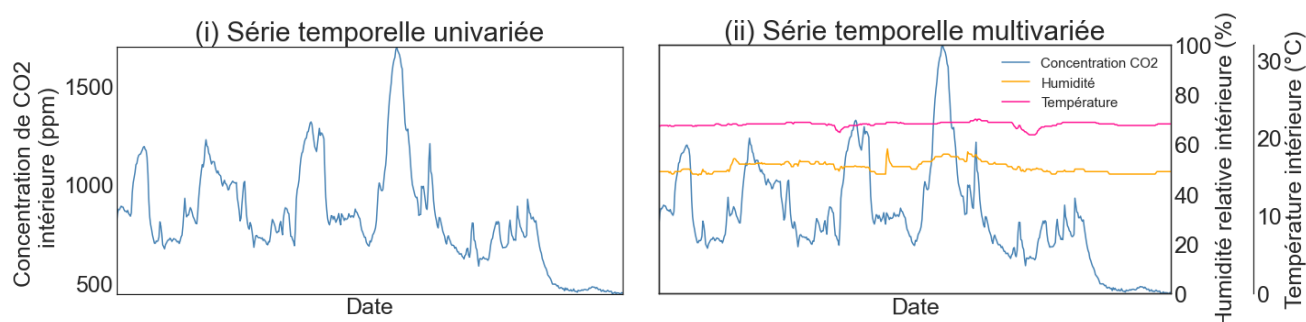


Figure 6 Exemple de séries temporelles univariée et multivariée.

1.4. ... à leur utilisation : enjeux, bénéfices

1.4.1. La couche d'application : des données aux services

Les informations produites par l'ensemble des composants constitutifs de la couche de perception permettent aux bâtiments connectés et intelligents d'offrir une grande variété d'applications potentielles à destination de divers acteurs tels que les occupants, les gestionnaires ou les logiciels et systèmes mêmes les composants (Karagiannis et al., 2015; Sovacool and Furszyfer Del Rio, 2020). Ainsi, la couche applicative peut être vue comme l'étape où de la valeur est créée à partir des données collectées pour fournir des services, à des tiers. Ces services ont pour objectif d'assister les utilisateurs dans la gestion et

l'exploitation du bâtiment au moyen de technologies basées sur l'exploitation des données (telles que les algorithmes, l'intelligence artificielle, l'apprentissage automatique, les statistiques, *etc.*) (Moudgil et al., 2023). Par ailleurs, certains auteurs comme Trappey et al. (2017), font le choix de faire une distinction nette entre l'aspect orienté service de la couche d'application et celui plus orienté sur les logiciels et les calculs, en la scindant en deux sous-couches du même nom.

Les services proposés sont variés et nombreux et peuvent comprendre de l'analyse ou de la visualisation de données, de la détection d'erreurs, des applications d'assistance ou de contrôle à distance, des systèmes d'automatismes, des protocoles de sécurité ou de gestion des ressources, *etc.* (Karimi et al., 2021; Trappey et al., 2017). À ce titre, Djenouri et al. (2020) ont détaillé et recensé plusieurs applications basées sur des techniques d'exploitation des données, notamment d'apprentissage automatique, appliquées aux données issues du bâtiment, comme l'ont fait Jia et al. (2019) dans une étude plus générale. Cette dernière a examiné et recensé différentes applications et les bénéfices qu'elles apportent pour divers types de bâtiments. On assiste ainsi au développement d'un modèle économique et social centré sur l'exploitation des données (Adhikari et al., 2023) sur lequel repose un grand nombre d'applications et de services à destination des utilisateurs ou acteurs du bâtiment (Djenouri et al., 2020).

1.4.2. Enjeux et bénéfices des bâtiments connectés ou intelligents

Ces différents services offerts par les bâtiments connectés et intelligents présentent de nombreux avantages, tant pour les utilisateurs (occupants, prestataires, *etc.*) que pour la réponse aux enjeux écologiques, sociaux, économiques et énergétiques discutés en introduction. Bien que nombreux, une revue de la littérature permet de mettre en lumière les éléments les plus récurrents. Le Tableau 1 présente les principaux bénéfices, répartis en thématiques, associés aux bâtiments connectés ou intelligents, tels qu'identifiés dans plusieurs publications. Pour élaborer ce tableau, seuls les avantages clairement recensés ont été pris en compte, sans considérer les liens de cause à effet éventuels. Par exemple s'il n'est fait mention que d'*économies d'énergies*, seule la thématique « *énergie* » est indiquée même si un lien sous-jacent existe avec d'autres domaines tels que les *finances* ou l'*environnement*. Les publications sélectionnées proposent des points de vue variés, tels que ceux d'occupants ou d'utilisateurs potentiels (Wilson et al., 2017), d'experts (Sovacool and Furszyfer Del Rio, 2020) et de la littérature (Marikyan et al., 2019). Ce tableau met en évidence neuf catégories d'enjeux et de bénéfices récurrents, bien que non exhaustifs : confort, énergie, finance, sécurité, santé, maintenance, social et environnement.

Tableau 1 Thématiques des objectifs et bénéfices régulièrement associés aux bâtiments connectés et intelligents dans la littérature.

Thématiques	(Moreno et al., 2014)	(Sovacool and Furszyfer Del Rio, 2020)	(Marikyan et al., 2019)	(Wilson et al., 2017)	(Karimi et al., 2021)	(Boudell al, 2014)	(Jia et al., 2019)
Confort	✓	✓	✓	✓	✓	✓	✓
Energie	✓	✓	✓	✓	✓	✓	✓
Finances		✓	✓	✓	✓	✓	
Sécurité	✓	✓		✓	✓	✓	
Santé	✓	✓	✓	✓	✓		✓
Maintenance	✓	✓					✓
Environnement		✓	✓				
Aspect social	✓	✓	✓	✓			

Parmi ces thématiques, ce sont celles centrées sur *l'énergie*, plus précisément sur les économies d'énergie ou l'efficacité énergétique, qui reviennent le plus fréquemment. Ce constat est en phase avec le contexte actuel soulevé en introduction. Ces économies peuvent être réalisées par le biais de la visualisation de données (Vellei et al., 2016), le pilotage autonome de systèmes tels que le CVC permettant une meilleure gestion des bâtiments (Stamatescu et al., 2019), l'évaluation de l'occupation pour prédire et ajuster les besoins (S. Kim et al., 2019) ou encore par le biais d'audits énergétiques permettant l'évaluation des performances réelles du bâtiment, par exemple dans le cadre d'un processus de retour d'expérience ou de rénovations (Recht et al., 2018). Mariano-Hernández et al. (2021) proposent par ailleurs une étude recensant différentes stratégies de gestion de l'énergie dans le bâtiment. Cette thématique est souvent étroitement liée à celle de *l'environnement*, bien que cette dernière soit moins régulièrement mise en avant. Un constat similaire peut être réalisé pour la thématique *financière*, même si elle est plus prépondérante. Une gestion plus efficace de l'énergie est souvent synonymes d'économies (Marikyan et al., 2019; Wilson et al., 2017).

Cependant, on retrouve également un niveau élevé d'occurrence pour les thématiques orientées sur le *confort* des occupants, ce qui n'est guère étonnant compte tenu du fait que l'on estime qu'une personne passe environ 80 % de son temps à l'intérieur (Yang and Wang, 2013). L'un des avantages les plus conséquent réside dans la possibilité de contrôler à distance l'intégralité de son logement, y compris les systèmes et objets connectés inclus (Sovacool and Furszyfer Del Rio, 2020). On trouve parfois cette thématique sous les termes de *commodité* ou de *contrôle*. À ce titre, Marikyan et al. (2019) consacrent une partie de leurs travaux à la compilation d'études traitant du confort de l'occupant et du contrôle à distance des équipements et installations.

Les thématiques abordées concernent ensuite généralement celles centrées sur la *sécurité*, notamment la prévention ou la notification d'accidents tels que les départs de feu ou les intrusions (Verma et al., 2019). Ensuite, viennent celles traitant de la *santé*, en

particulier la notification d'incidents (p.ex. chutes) ainsi que le suivi de personnes malades ou âgées (Biswas et al., 2013). On trouve ensuite des thématiques orientées sur la *maintenance* du bâtiment, comprenant la détection d'erreurs (Casado-Vara et al., 2018), la réalisation de diagnostics et les rappels de maintenance. Enfin, il y a des thématiques axées sur des aspects plus *sociaux*, où la connectivité permet de faciliter les interactions tant professionnelles que personnelles (Sovacool and Furszyfer Del Rio, 2020).

Bien que les principales thématiques abordées tournent souvent autour de l'optimisation et la réduction des consommations énergétiques dans le bâtiment ainsi que du confort de l'occupant, de nombreuses publications recensent les différentes utilisations de services pour lesquels des données issues du bâtiment sont utilisées dans la littérature. En complément de celles introduites précédemment on peut également prendre en considération les études de Qiang et al. (2023), Petroşanu et al. (2019) et Moudgil et al. (2023), en plus de celles cités dans le Tableau 1, afin de développer davantage le sujet.

1.5. Les anomalies, un facteur limitant à considérer

1.5.1. De l'importance des données à leurs limites

Comme détaillé dans la partie 1.2.2, le bâtiment connecté représente un système complexe constitué d'un ensemble de couches par lesquelles les données transitent, de leur génération à leur exploitation. Cependant, aussi performant et fiable qu'il puisse être, ce système peut être sujet à des événements anormaux, rares et non désirés, tels qu'un défaut matériel, une perturbation de l'environnement ou du réseau, qui peuvent impacter la quantité et la qualité des données récoltées. Cela peut entraîner des pertes de données (Adhikari et al., 2023) ainsi que la présence de données inhabituelles (Mishra et al., 2021), que l'on désigne sous le terme d'**anomalies**. Sans être exhaustif, les origines de tels événements anormaux peuvent être diverses. Ils peuvent être dus à des erreurs de mesure liées au mauvais étalonnage d'un capteur ou à des perturbations environnementales telles que des vibrations. Ils peuvent également être causés par des défaillances systèmes, provoquées par des erreurs de transmission, d'exploitation ou d'autonomie, comme un arrêt de l'alimentation. De plus, ces anomalies peuvent être la conséquence directe d'erreurs humaines, comme une mauvaise manipulation ou une utilisation inadéquate d'un instrument (Ni et al., 2009). L'impact de ces anomalies sur les données peut également varier considérablement, allant de variations simples et facilement repérables, telles qu'une concentration de CO₂ nulle, à des modifications bien plus complexes et difficilement repérables, comme la notification erronée d'une ouverture de fenêtre.

Cependant, selon leur gravité, la présence d'anomalies dans les données peut empêcher ou compromettre l'utilisation ou le fonctionnement d'applications ou de services pour lesquels quantité et qualité des données sont essentiels, tels que le contrôle à distance ou la visualisation. Cela peut avoir un impact direct sur le fonctionnement du bâtiment, qui de *facto*, peut avoir des conséquences sur l'expérience, le confort, la sécurité et les économies de l'occupant, ainsi que pour d'autres acteurs du bâtiment qui utilisent également ces données, tels que les fournisseurs d'énergie ou les entreprises de maintenance. Pour répondre à cette problématique, il est essentiel d'identifier les événements rares ou anormaux survenant dans les données pour les corriger à terme.

Ainsi, de nos jours, la détection des anomalies est utilisée dans divers domaines à des fins multiples. Par exemple, elle est employée dans le domaine financier pour détecter les fraudes, dans le domaine de la sécurité pour repérer les intrusions, et même dans le domaine médical pour la détection de cancer du sein (Chalapathy and Chawla, 2019; Chandola et al., 2009). Cette approche a évolué au fil du temps, passant d'une simple utilisation de règles définies par des experts à des procédures plus avancées comprenant des analyses statistiques et des techniques sophistiquées d'apprentissage automatique (Zimek and Filzmoser, 2018). Cependant, compte tenu de la quantité et de l'hétérogénéité des données produites par le bâtiment, ainsi que de la variété et la rareté des anomalies, la tâche s'avère complexe, délicate et chronophage. C'est pourquoi, les techniques d'apprentissage automatique apparaissent comme une solution potentielle pour relever ce défi.

1.5.2. Apprentissage automatique pour la détection d'anomalies, une solution ?

L'apprentissage automatique, ou *Machine Learning* (ML) en anglais, est une sous discipline de l'intelligence artificielle qui se base essentiellement sur l'analyse et l'exploitation des données. L'explosion de la quantité de données produites, qui est aujourd'hui massive, a considérablement stimulé le développement de techniques d'apprentissage automatique de plus en plus variées et nombreuses (Foulon, 2020). Leurs champs d'applications se sont diversifiés ces dernières années et couvrent de nombreux domaines et de nombreuses utilisations allant de la finance (Nazareth and Ramana Reddy, 2023), au confort et aux loisirs avec par exemple la reconnaissance et le traitement vocal sur nos téléphones ou télévisions (p.ex., avec les assistants vocaux *Alexa* d'Amazon ou *Siri* d'Apple) (Padmanabhan and Johnson Premkumar, 2015), en passant par la santé (Baloglu et al., 2021) où elles sont mises en œuvre pour la détection et la prédiction de maladies telles que le cancer du sein (Kourou et al., 2015), et même dans le domaine du bâtiment (Djenouri et al., 2020), comme précédemment évoqué.

L'apprentissage automatique part du postulat qu'il est possible d'entraîner des ordinateurs (*Machine*), ou des modèles informatiques, à apprendre (*Learning*) à partir de données collectées afin de résoudre des tâches pour lesquelles ils n'ont pas été explicitement programmés. Ces ordinateurs, ou modèles, « apprennent » à partir de l'expérience et des données passées. L'apprentissage automatique apparaît particulièrement adapté lorsque la réponse à un problème n'est pas connue *a priori*, que l'expertise humaine n'existe pas ou est difficilement applicable, et que des quantités significatives de données sont disponibles (Djenouri et al., 2020). Dans ce contexte, l'apprentissage automatique se présente comme une approche centrée sur les données, ce qui en fait une solution particulièrement adaptée à la détection d'anomalies, laquelle est un sous-domaine majeur de l'apprentissage automatique (Foulon, 2020).

La détection d'anomalies est un domaine de recherche qui a été examiné en profondeur dans la littérature. Cette tâche a donné lieu à de nombreuses études qui diffèrent selon les domaines d'application, les spécificités de données utilisées, leur disponibilité ainsi que les caractéristiques des anomalies recherchées (Chandola et al., 2009). Les travaux de cette thèse se concentreront donc sur la problématique de la détection d'anomalies dans les mesures issues du bâtiment.

1.6. Conclusion

En conclusion de ce chapitre, il est clair que les bâtiments connectés et intelligents offrent des réponses possibles aux enjeux énergétiques, sociaux, financiers et environnementaux actuels. Ces réponses peuvent prendre la forme d'applications et de services variés à destination des occupants, d'autres utilisateurs et du bâtiment lui-même. Toutefois, le bon fonctionnement de ces services dépend pour beaucoup des données générées et collectées par le bâtiment, faisant de l'accessibilité et de la disponibilité des données un enjeu majeur.

Cependant, la présence d'événements rares et anormaux tout au long du processus, de la production à l'exploitation des données, limite souvent la quantité ou la qualité des données disponibles, ce qui entrave la mise en place de nombreux services essentiels. Pour résoudre ce problème, il est essentiel d'identifier ces événements anormaux au sein d'ensembles de données variées afin de les corriger et de disposer finalement de données de haute qualité en quantité suffisante. Néanmoins, cette tâche est complexe, compte tenu de la variété et la rareté de ces événements, ainsi que de la quantité et de l'hétérogénéité des données issues du bâtiment.

L'utilisation de techniques basées sur l'apprentissage automatique apparaît comme une solution prometteuse, car elles peuvent aider à surmonter une partie de ces

difficultés. Le chapitre 2 définira et explorera le principe de la détection d'anomalies appliqué aux données issues du bâtiment, en mettant l'accent sur les séries temporelles. Il abordera également les différentes problématiques, contraintes et enjeux auxquels ce domaine est confronté, tout en développant les objectifs de cette thèse.

Chapitre 2

État de l'art sur la détection d'anomalies et les méthodes d'apprentissage automatique

Ce chapitre vise à présenter la démarche de détection d'anomalies et à mettre en évidence les enjeux qui en découlent. Dans un premier temps, les définitions d'anomalies et de détection d'anomalies seront apportées (section 2.2 et 2.3). Puis, l'état actuel de cette pratique dans différents domaines sera examiné dans la section 2.3. Cette partie sera suivie par l'identification et la définition des différentes méthodes d'apprentissage automatique existantes, leur processus d'évaluation ainsi qu'une description des principales familles de techniques de détection d'anomalies (section 2.4). Un bilan sera ensuite dressé dans la section 2.5 afin de présenter les défis et problématiques liés à la tâche de détection d'anomalies appliquée aux données du bâtiment. Enfin, le chapitre se terminera par la présentation des motivations de la thèse (section 2.6), laquelle sera suivie d'une synthèse (section 2.7).

« Celui qui n'a jamais fait d'erreur, n'a jamais rien fait. »

Extrait traduit de, La légende de Marche-Mort, de David Gemmell

Sommaire du chapitre 2

2.1. Introduction	29
2.2. Qu'est-ce qu'une anomalie ?	29
2.3. Une introduction à la détection d'anomalies.....	31
2.3.1. Contexte et difficultés	31
2.3.2. Illustration par l'exemple.....	32
2.3.3. Domaines d'application	33
2.4. Familles et méthodes d'apprentissage automatique	35
2.4.1. Les méthodes d'apprentissage automatique.....	35
2.4.2. Estimation des performances : métriques d'évaluation.....	37
2.4.3. La famille des méthodes ou techniques de détection d'anomalies.....	40
2.5. La diversité, à la fois une force et une contrainte	49
2.6. Contexte et motivations de la thèse	52
2.7. Conclusion.....	53

2.1. Introduction

Ce chapitre présente les bases de la détection d'anomalies, un champ de recherche vaste et dynamique lié à l'apprentissage automatique et à l'analyse de données. Tout d'abord, nous définirons le terme « anomalie » et discuterons de ses implications sur le processus de détection. Ensuite, nous explorerons les principaux domaines de recherche associés, les approches et les stratégies utilisées afin de mener à bien cette tâche. Pour finir, nous discuterons des problématiques soulevées par cet état de l'art.

2.2. Qu'est-ce qu'une anomalie ?

Dans la littérature, de nombreux termes sont utilisés afin de désigner une anomalie, notamment des synonymes tels que '*discordant observation*' (en français, une observation discordante), '*exception*' (exception), '*aberration*' (aberration), '*surprise*' (surprise), '*peculiarity*' (particularité), '*contaminant*' (contaminant) ou '*novelty*' (nouveau) (Blázquez-García et al., 2022; Kraiem, 2021). Les deux termes les plus communs sont **outlier** et **anomaly**. Ils sont régulièrement employés de manière interchangeable (Aggarwal, 2017; Chandola et al., 2009), et peuvent être respectivement traduits en français par « valeur aberrante » et « anomalie ». Aucun consensus ne semble exister sur le terme à adopter (Carreño et al., 2020), ce qui est également le cas pour leurs définitions, qui varient selon le domaine d'étude et plus particulièrement le cas d'usage (Foulon, 2020). Il existe cependant des définitions plus globales dont les plus communément utilisées sont illustrées chronologiquement dans le Tableau 2. Ces définitions permettent de mettre en évidence deux propriétés régulièrement associées aux anomalies :

- (i) Elles sont rares. La majorité des données étudiées sont constituées de données dites « normales ».
- (ii) Elles possèdent des caractéristiques propres qui les différencient des données « normales » et qui, par opposition, font d'elles des données dites « anormales ».

Sur cette base, une anomalie peut être vue comme : *une observation ou une séquence d'observations (resp. une donnée, un événement) rare avec des caractéristiques significativement différentes des autres observations (resp. données, événements).*

Tableau 2 Définitions d'*outlier* et d'*anomalies* dans la littérature.

Référence	Terme	Définitions
(Grubbs, 1969)	<i>Outlier</i>	<i>'An outlying observation, or "outlier," is one that appears to deviate markedly from other members of the sample in which it occurs'</i>
(Hawkins, 1980)	<i>Outlier</i>	<i>'An observation which deviates so much from other observations as to arouse suspicions that it was generated by a different mechanism'</i>
(Barnett and Lewis, 1994)	<i>Outlier</i>	<i>'An observation (or subset of observations) which appears to be inconsistent with the remainder of that set of data'</i>
(Chandola et al., 2009)	<i>Anomaly</i>	<i>'Anomalies are patterns in data that do not conform to a well-defined notion of normal behavior'</i>
(Sadik and Gruenwald, 2011)	<i>Outlier</i>	<i>'An outlier is a data point which is significantly different from other data points, or does not conform to the expected normal behavior, or conforms well to a defined abnormal behavior'</i>

Cette définition se concentre sur le caractère rare et distinctif des anomalies vis-à-vis des autres observations. Cependant, elle ne tient pas compte de la cause sous-jacente des anomalies. Par exemple, une température extérieure de 30 °C en avril à Bordeaux peut être considérée comme anormale compte tenu des tendances observées ces trente dernières années. Cependant en 2023 ce n'est pas une erreur de mesure qui semble en être la cause mais un ensemble d'autres processus comme le réchauffement climatique ou l'effet d'îlot de chaleur urbain. Cette température extérieure est anormale mais ne remet pas en cause la qualité des mesures affichées. C'est pourquoi il est essentiel de considérer le processus qui génère l'anomalie, ainsi que sa cause sous-jacente, car cela altère le comportement attendu de la variable mesurée (Aggarwal, 2017).

Les anomalies peuvent résulter de processus différents de ceux des autres observations et leur analyse peut fournir, selon les objectifs poursuivis, des informations utiles (Hawkins, 1980). Ainsi et dans un contexte donné, on distingue deux types d'anomalies. Le premier correspond à des anomalies dont l'analyse n'est pas pertinente et qui détériorent la qualité des données en raison de phénomènes parasites tels que des erreurs de mesure ou du bruit. Ces anomalies doivent être corrigées pour obtenir des données exploitables. Une telle anomalie est présentée dans la Figure 7 (i) sous la forme de pics, présentant des variations brusques positives ou négatives, (★) sur une mesure de température intérieure et dont l'origine peut provenir d'un capteur défectueux. Le deuxième type d'anomalies représente des observations qui reflètent des processus ou des phénomènes dignes d'intérêt, mais ne pouvant être déterminé qu'après une analyse plus approfondie (Aggarwal, 2017; Blázquez-García et al., 2022). Une telle anomalie est présentée dans la Figure 7 (ii) sous la forme d'un écart brusque négatif de plusieurs observations (■) sur la même mesure de température intérieure et qui trouve son origine dans l'ouverture d'une fenêtre. Ces anomalies peuvent également être appelées '*weak*

outlier pour le premier type et *strong outlier* pour le second (Aggarwal, 2017; Panjei et al., 2022). Dans les deux cas, ces anomalies se traduisent par des comportements anormaux. La limite entre elles n'est pas clairement définie, aussi l'une peut être interprétée comme étant l'autre. La distinction passe par une analyse plus approfondie de chaque anomalie et nécessite l'apport d'une expertise métier. Dans le cadre de ces travaux ces deux genres anomalies seront considérés puisqu'elles caractérisent toutes deux des phénomènes anormaux.

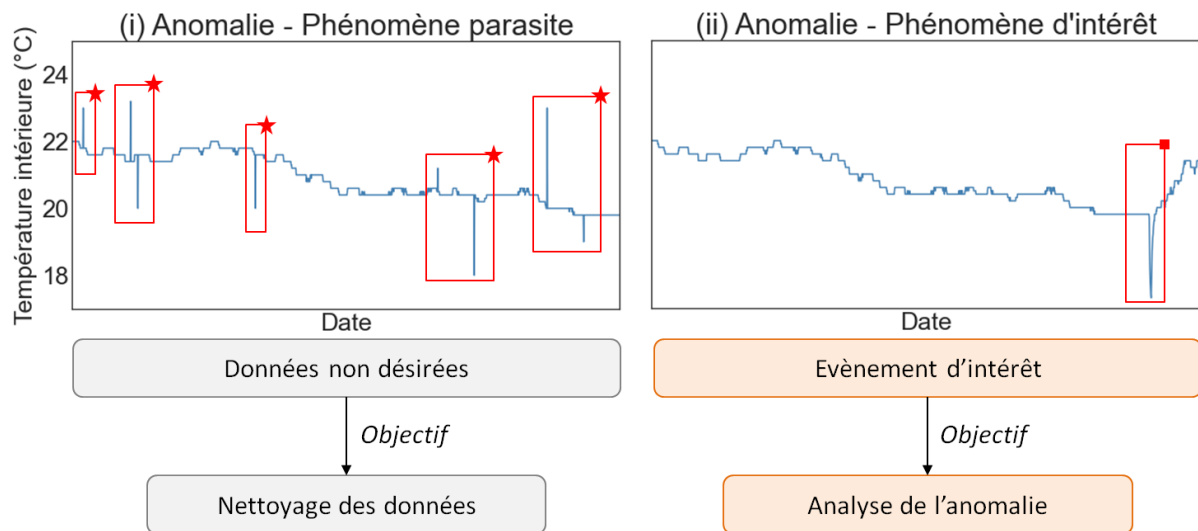


Figure 7 Exemple d'anomalies provoquées par un phénomène parasite (★) et par un phénomène d'intérêt (■).

2.3. Une introduction à la détection d'anomalies

2.3.1. Contexte et difficultés

À première vue, l'approche la plus simple pour détecter des anomalies parmi un ensemble d'observations serait de définir un profil d'observations normales et de considérer toutes les autres comme anormales. L'idée serait d'établir des règles, d'élaborer une méthode, une technique, ou même d'entraîner un modèle dans le but d'effectuer la distinction entre ces deux états et séparer ainsi les observations étudiées en deux classes : normale ou anormale. Cependant, cette approche comporte plusieurs défis comme souligné par Chandola et al. (2009) et Mehrotra et al. (2017) :

- L'absence de définitions précises séparant les observations anormales des normales crée une frontière floue et changeante entre les deux.
- La définition d'une observation anormale varie selon le domaine d'étude, l'application, le contexte ou des objectifs spécifiques déterminés au préalable.

- La présence de bruit dans les données, souvent similaires à des anomalies complique leur distinction et donc leur séparation (Aggarwal, 2017).

Tous ces facteurs rendent la distinction entre observations normales et anormales confuse, variable, et posent diverses problématiques spécifiques aux domaines d'application et aux types d'anomalies rencontrées, influençant ainsi le choix des méthodes utilisées (Carreño et al., 2020; Mehrotra et al., 2017). En effet, il n'est intuitivement pas surprenant de supposer qu'on ne détecte pas un cancer du sein sur des images d'ultrason (Fujioka et al., 2020) comme on détecte des comportements anormaux sur des réseaux sociaux (Savage et al., 2014) ou des erreurs de mesures d'un capteur de CO₂ sur un mur végétal (Liu et al., 2020). D'autres éléments liés aux *données* et propres aux domaines d'application exercent également une influence sur l'approche adoptée, tels que leur nature (Tan et al., 2014), leurs relations (Erhan et al., 2021), leur contexte (Huang, 2018), leur qualité (Tan et al., 2014) et leur dynamique (Foulon, 2020). Ainsi, des problématiques distinctes peuvent également émerger au sein d'un même domaine, qui ne se limite généralement pas à l'étude d'un seul type de données ou d'anomalies.

2.3.2. Illustration par l'exemple

Afin d'illustrer la difficulté d'une approche de détection d'anomalies, posons l'exemple d'une scène dans laquelle un doctorant se rend à une séance de cinéma au cœur de Bordeaux. À quelques rues de sa destination, il aperçoit au détour d'un croisement une grande figure affublée d'un grand chapeau bleu et pointu, une longue cape grise et une écharpe argent, surmontée d'une barbe blanche qui descend jusque sous sa ceinture, ainsi que d'énormes bottes noires. Ce personnage – réplique fidèle d'un magicien bien connu d'une série littéraire et cinématographique à succès – bloque la rue, hurlant à qui voudra bien l'entendre qu'« il ne passera pas ». Un sourire crispé aux lèvres, l'étudiant presse le pas afin de mettre un peu de distance entre lui et ce vieil homme excentrique. Quelques mètres plus loin une nouvelle rencontre insolite prend forme en la personne d'un lapin blanc vêtu d'un gilet bleu et d'une montre à gousset, qui le percute. En lieu et en place d'excuses, l'être à fourrure blanche marmonne en sa direction un rapide « en retard, toujours en retard » avant de s'enfuir en quelques bonds.

Outre la singularité de ces rencontres du troisième type, elles revêtent un caractère anormal du fait de leur localisation en plein le cœur de la ville de Bordeaux. Cela permet de mettre en avant le caractère *spécifique à une application et à un type de données* d'une anomalie ici représenté par la localisation et l'apparence inhabituelle des rencontres. En effet, on s'attendrait plutôt à croiser ces personnages sur un plateau de film, lors d'un festival dédié à leurs univers voire, dans un livre. À ce moment de la scène ces rencontres sont, pour l'étudiant compte tenu des informations dont il dispose, des

anomalies. D'ailleurs, l'histoire aurait pu s'arrêter sur ces conclusions si, choqué, l'étudiant avait décidé de rebrousser chemin.

Malgré tout, et bien qu'un peu chamboulé, il finit par atteindre le cinéma. À l'heure pour la séance qui plus est. Une fois sur place, les posters, le *brouhaha* ambiant mais surtout une dizaine de déguisements d'elfes, de nains et de hobbits lui apprennent sans mal qu'une diffusion des versions longues de la trilogie du « Seigneur des Anneaux » est prévue pour plus tard dans la soirée. Ces nouveaux éléments apportent de nouvelles informations ainsi que de plausibles explications quant à la présence d'un magicien tout de gris vêtu à quelques rues du cinéma... lui retirant potentiellement son statut d'anomalie. Cela permet de mettre en avant l'importance du *contexte* et donc de l'information apportées par *d'autres données* dans la prise de décision mais aussi le caractère *flou et évolutif* de la limite entre normalité et anormalité. Sans ces informations complémentaires, cette rencontre serait restée une anomalie. La rencontre du magicien et du lapin a, de prime abord, été perçues comme similaires. Toutes deux anormales. Seulement, l'apport d'informations complémentaires a permis d'apporter un contexte nouveau quant à la présence du magicien. Ce dernier peut donc, au vue de sa localisation, être plutôt perçu comme un *bruit* tandis que la présence du lapin restera un mystère et par conséquent, une *anomalie*.

On remarquera que le simple changement de quelques éléments à l'histoire suffirait à créer un problème entièrement différent et aux issues qui le sont tout autant. Afin d'illustrer plus formellement les différentes problématiques abordées dans la littérature nous introduirons et discuterons par la suite des domaines mettant en œuvre la détection d'anomalies.

2.3.3. Domaines d'application

De nos jours et comme détaillé précédemment dans la partie 1.5, la détection d'anomalies est utilisée dans de nombreux domaines tels que la finance, la sécurité et même le domaine médical. Les spécificités de chaque domaine, liées à la variété des données et des types d'anomalies, influencent fortement le choix des techniques à appliquer. Par conséquent, la littérature propose un large éventail de techniques de détection d'anomalies. Cette partie, basée sur les travaux de Chalapathy et Chawla (2019), Chandola et al. (2009) et Mehrotra et al. (2017), présente les principaux domaines d'application et de recherche en détection d'anomalies, en fournissant quelques exemples représentatifs.

- **Cyber sécurité.** Le domaine de la cyber sécurité utilise la détection d'anomalies afin de répondre à plusieurs besoins parmi lesquels la détection d'intrusions pour identifier les activités suspectes sur les systèmes informatiques (Phoha, 2002) et la

détection de logiciels malveillants tels que les *malwares* (p. ex. les *ransomwares*) (Ye et al., 2018). L'objectif principal est de protéger des informations sensibles, comme des données bancaires ou médicales, en traitant une variété de données, y compris des textes, des images et des données réseau (Foulon, 2020).

- **Médical.** Le domaine médical applique la détection d'anomalies à des fins d'analyse, de diagnostic et de prédiction (Fernando et al., 2022). Cette approche est utilisée pour reconnaître les signes de pathologies telles que le cancer, les conditions médicales spécifiques (p. ex. une carence) ainsi que les possibles erreurs de mesure (p. ex. une instrumentation défectueuse) ou pour prédire l'état physiologique futur d'un patient, comme p. ex. son taux de rétablissement. Les données traitées dans ce domaine sont diverses, allant de vidéos et imageries médicales telles que l'endoscopie ou la radiographie (Lundervold and Lundervold, 2019; Soffer et al., 2020), aux séries temporelles telle que des électrocardiogrammes (ECG) (Li et al., 2020).
- **Financier.** Le domaine financier se sert principalement de la détection d'anomalies pour lutter contre les fraudes, qui peuvent prendre différentes formes, allant du vol de la carte de crédit à la fraude à l'assurance ou au crédit (Zhou et al., 2018). Les données étudiées dans ce contexte sont généralement variées et peuvent être spatiales (p. ex. un achat dans un autre pays), temporelles, contextuelles (p. ex. un crédit non remboursé) (Mehrotra et al., 2017), ou textuelles (Chandola et al., 2009). L'objectif est de détecter rapidement, voire instantanément, les activités frauduleuses, afin de minimiser les pertes financières (Kraiem, 2021). Cependant, les fraudes ne suivent pas de motifs récurrents propres (ou '*patterns*' en anglais), les rendant donc complexes à détecter (Chalapathy and Chawla, 2019).
- **Industriel.** Le domaine industriel présente une grande variété d'applications pour la détection d'anomalies, allant de la détection d'erreur ou de dommages sur des machines en fonctionnement à la détection de produits défectueux, ainsi qu'à la gestion de stock (Mehrotra et al., 2017). Ce domaine implique un large éventail de données à considérer, notamment différents types de machines, de fluides, de produits aux caractéristiques variées, *etc.*, bien que la plupart de ces données soient principalement temporelles (Purarjomandlangrudi et al., 2014). Par exemple, dans le cas de la détection d'erreurs ou de dommages, l'objectif peut être de détecter l'apparition de motifs spécifiques au moyen de séries temporelles (Martí et al., 2015) ou de prévenir ces anomalies à l'aide d'images (Sanjay Kumar Ahuja et al., 2019), comme l'arrêt inattendu d'une machine pendant une procédure ou l'apparition de corrosion.
- **Réseaux de capteurs.** Dans le domaine de l'Internet des Objets (IdO), qui est lié à notre étude sur les données issues des mesures du bâtiment (Djenouri et al., 2020), la détection d'anomalies est principalement utilisée pour détecter ou prédire des

événements anormaux, tels qu'une consommation anormale d'électricité (Himeur et al., 2021b), ou des erreurs de mesure (Ayadi et al., 2017). Les données étudiées peuvent être de diverses natures, mais sont majoritairement des séries temporelles (Chalapathy and Chawla, 2019).

Dans cette partie, plusieurs cas d'utilisation des techniques de détection d'anomalies ont été identifiés et discutés. On constate qu'il existe un large éventail de données aux utilisations diverses. Bien que chaque domaine ne soit pas limité à un type de données particulier, certaines associations apparaissent plus fréquemment dans la littérature (Foulon, 2020). Par exemple, dans le domaine médical, les applications traitent plus souvent des images, des vidéos ou des séries temporelles que des données textuelles. De même, bien que le domaine du bâtiment puisse inclure plusieurs types de données différents comme des données textuelles, des images ou des vidéos, la majorité des données traitées sont des séries temporelles. Dans la prochaine partie, nous discuterons des différentes familles de techniques et de leurs spécificités.

2.4. Familles et méthodes d'apprentissage automatique

Dans cette partie, nous introduirons les différentes méthodes d'apprentissage que peut suivre un modèle en décrivant les trois principales familles de l'apprentissage automatique ainsi que les méthodes d'évaluation couramment utilisées. Ensuite, nous introduirons les différentes familles de détection d'anomalies qui en découlent.

2.4.1. Les méthodes d'apprentissage automatique

Les méthodes d'apprentissage automatiques sont des méthodes, algorithmes ou ensembles d'algorithmes, qui permettent à un ordinateur « d'apprendre » à partir de données. Ce processus conduit à la création d'un modèle apte à exécuter des tâches spécifiques sur de nouvelles données. Plus précisément, l'apprentissage consiste à entraîner un modèle pour qu'il puisse extraire des informations pertinentes, des caractéristiques, des relations ou des structures à partir de données d'entrée, ce qui lui permet de répondre à des tâches spécifiques (Lawrynowicz and Tresp, 2014). Autrement dit, ces méthodes d'apprentissage permettent au modèle de générer des *réponses en sortie* en fonction des *données d'entrée*. Cela signifie que la performance du modèle dépend fortement des données utilisées pour son entraînement ainsi que de la manière dont il est entraîné (Emmott et al., 2016). Le modèle ne peut exécuter des tâches ou fournir des réponses qu'en fonction de ce qu'il a appris lors de processus d'entraînement. À titre d'exemple on pourra citer la prise de décision, la classification ou même la prédiction comme tâches possibles.

On distingue trois méthodes distinctes d'apprentissage automatique, applicables selon les données à disposition (Aggarwal, 2017), et dont dépendent différentes techniques : l'apprentissage supervisé, l'apprentissage non supervisé et l'apprentissage semi-supervisé. À ces trois méthodes s'ajoute également l'apprentissage par renforcement (Kaelbling et al., 1996; Shakya et al., 2023) qui ne sera pas abordé.

- **Apprentissage supervisé.** Cette méthode nécessite un ensemble de données d'entraînement contenant des exemples préalablement étiquetées. Cela signifie que toutes les observations, qu'elles soient normales ou anormales, sont annotées comme telles. Cet étiquetage permet au modèle d'apprendre au moyen d'exemples concrets, ce qui lui permet par la suite d'effectuer des prédictions ou des classifications sur des données inconnues. Les exemples courants d'apprentissage supervisé incluent des techniques de *classification* (p. ex. détecter l'ouverture d'une fenêtre) et de *régression* (p. ex. prédire la température extérieure qu'il fera dans une heure) (Hastie et al., 2009; Himeur et al., 2021b). Dans le cas de la détection d'anomalies, cela sous-entend de disposer des données normales et anormales étiquetées pour l'entraînement.
- **Apprentissage non supervisé.** Cette méthode est utilisée lorsque les données ne sont pas étiquetées. Dans ce cas, le modèle doit, de manière autonome, identifier ou découvrir les relations, les informations ou les structures « cachées » dans les données (Hastie et al., 2009). Les résultats dépendent largement du choix de la méthode utilisée et de l'objectif prévu pour le modèle (Foulon, 2020). Le processus d'entraînement suivi est similaire, mais c'est à l'utilisateur de donner du sens aux réponses produites par le modèle puisqu'il n'y pas d'étiquettes pour guider l'apprentissage. Cependant, étant donné qu'il est souvent difficile de disposer de données labellisées, l'apprentissage non supervisé est l'une des méthodes les plus utilisées, comme le détaillent Himeur et al. (2021b) et Chandola et al. (2009). Les exemples courants d'apprentissage non supervisé incluent les techniques de *clustering* (p. ex., regrouper des données similaires en amas ou *clusters*), les techniques de *réduction de dimension* (p. ex. réduire la taille d'un ensemble de données tout en conservant autant d'informations que possible) et les techniques basées sur le *voisinage* (Hastie et al., 2009; Himeur et al., 2021b).
- **Apprentissage semi-supervisé.** Cette méthode d'apprentissage se situe entre les méthodes supervisées et non supervisées. Cette approche utilise à la fois des données étiquetées et non étiquetées lors de l'apprentissage, ainsi que des données ne présentant qu'un seul type d'étiquetage, par exemple uniquement des données considérées comme normales. Dans le cas de la détection d'anomalies, l'apprentissage semi-supervisé se concentre principalement sur l'apprentissage des données étiquetées comme normales, ce que l'on appelle souvent l'apprentissage de la « normalité ». Ainsi, toute nouvelle donnée qui s'éloigne de ce que le modèle a appris

comme étant normal constitue une nouveauté et par conséquent une anomalie. Cette approche est souvent appelée « détection de nouveauté » (*'novelty detection'* en anglais) dans la littérature (Markou and Singh, 2003; Miljković, 2020). Les exemples courants d'application de l'apprentissage semi-supervisé incluent la *classification* (p. ex. détecter une intrusion ou une fraude), la *régression*. Dans le cas de la détection d'anomalies, une hypothèse récurrente consiste à considérer l'ensemble des données comme normales, en raison de la rareté des anomalies, afin de s'affranchir du besoin d'étiquetage. Cependant, cette approche est hautement dépendante des données, spécifiques à l'étude, et nécessite une expertise métier conséquente pour différencier *a priori* et *a posteriori* les données normales des données anormales (Foulon, 2020).

En résumé, le choix des techniques d'apprentissage automatique pour la détection d'anomalies est fortement influencé par la disponibilité des données d'entraînement et de leur étiquetage. Cette observation s'ajoute ainsi comme un élément supplémentaire à considérer en plus de ceux liés aux problématiques rencontrées et aux données étudiées (Chandola et al., 2009). Il est également important de souligner que compte tenu de l'ampleur du travail que requiert l'étiquetage des données, et par conséquent de leur rareté, les techniques non supervisées sont les plus couramment utilisées dans la littérature (Kraiem, 2021; Schneider et al., 2016). Néanmoins, qu'il s'agisse d'apprentissage supervisé, non supervisé, ou semi-supervisé, il est essentiel d'évaluer la capacité du modèle à réaliser sa tâche. Cette évaluation intervient pendant et après l'entraînement, et diverses méthodes d'évaluation sont disponibles, que nous explorerons dans la partie suivante.

2.4.2. Estimation des performances : métriques d'évaluation

L'évaluation des performances d'un modèle d'apprentissage automatique est essentielle pour garantir qu'il répond aux attentes. En fonction de la nature de la tâche (p. ex. classification, régression), différentes méthodes peuvent être appliquées pour mesurer les performances. Par exemple, pour les modèles de régression, on utilise souvent des métriques telles que l'erreur quadratique moyenne ou l'erreur absolue moyenne, tandis que pour la classification, on utilise des métriques telles que le taux d'erreurs (Handelman et al., 2019). Ces métriques permettent d'affiner les performances du modèle pendant l'apprentissage et de les évaluer *a posteriori* sur de nouvelles données. Dans le cas de la détection d'anomalies, l'évaluation des performances se concentre sur la capacité du modèle à classer correctement les observations comme normales ou anormales. Cela implique de comparer les prédictions du modèle aux classifications réelles, souvent appelées « vérité terrain ». Selon le contexte et les techniques de détection employées, il existe majoritairement deux types d'approches distinctes pour mener cette évaluation, celles basées sur la matrice de confusion et celles basées sur la courbe ROC (*Receiver Operating Characteristic*) (Tharwat, 2021).

2.4.2.1. La matrice de confusion

La matrice de confusion, telle qu'illustrée dans le Tableau 3, est un outil qui permet d'évaluer la performance d'un système de classification. Elle est composée de quatre éléments qui traduisent les résultats de cette classification : les vrais positifs (VP), les vrais négatifs (VN), les faux positifs (FP) et les faux négatifs (FN).

Tableau 3 Matrice de confusion pour un problème de détection d'anomalie.

		Prédiction	
		Positive	Négative
Réalité	Positive	Vrai Positif (VP)	Faux Négatif (FN)
	Négative	Faux Positif (FP)	Vrai Négatif (VN)

Les *vrais positifs* représentent le nombre de prédictions d'anomalies réalisées par le modèle qui correspondent à la réalité terrain. Les *vrais négatifs* sont l'ensemble des prédictions de données normales qui correspondent à la vérité terrain. Les *faux positifs* quantifient toutes les anomalies prédites par le modèle mais référencées comme normales dans la vérité terrain. Enfin, les *faux négatifs* dénombrent l'ensemble des observations qualifiées comme normales par le modèle mais qui sont en réalité des anomalies. À partir de la matrice de confusion découlent plusieurs métriques couramment utilisées, telles que la précision, le rappel, la spécificité ou encore le F1-score (Johnson and Khoshgoftaar, 2019). D'autres métriques, telles que l'exactitude (de l'anglais '*accuracy*') ou le taux d'erreurs, peuvent également être employées. Cependant, elles sont recommandées principalement lorsque les classes d'observations sont équilibrées, ce qui les rend moins adaptées à la détection d'anomalies, où les anomalies sont rares (Johnson and Khoshgoftaar, 2019). Par conséquent, ces métriques ne sont pas appropriées pour ces travaux et ne seront pas abordées.

- **La précision** est la proportion des identifications positives qui étaient effectivement correctes. Elle se calcule en divisant le nombre de vrais positifs par la somme des vrais positifs et des faux positifs.

$$Précision = \frac{VP}{VP + FP}$$

- **Le rappel**, aussi appelé taux de vrais positifs, correspond à la proportion de vraies instances positives qui ont été correctement identifiées. Il est calculé en divisant le nombre de vrais positifs par la somme des vrais positifs et des faux négatifs.

$$Rappel = \frac{VP}{VP + FN}$$

- **La spécificité**, aussi nommée taux de vrais négatifs, correspond à la proportion de vraies instances négatives qui ont été correctement identifiées. Elle est calculée en divisant le nombre de vrais négatifs par la somme des vrais négatifs et des faux positifs.

$$Spécificité = \frac{VN}{VN + FP}$$

- **Le F₁-score** est une mesure qui représente la moyenne harmonique entre la précision et le rappel. Il varie de 0 à 1, où une valeur de 1 indique une classification parfaite. Un F₁-score de 0,5 signifie qu'il y a deux classifications erronées (faux positifs ou faux négatifs) pour chaque vrai positif.

$$F_1 - score = \frac{Rappel \times Précision}{Rappel + Précision} = \frac{VP}{VP + \frac{1}{2}(FP + FN)}$$

Cependant, dans les situations où les modèles produisent des probabilités ou des distances plutôt que des classifications binaires (normale ou anormale) (Martin Nascimento et al., 2021), l'utilisation d'une matrice de confusion peut devenir complexe car elle dépend d'un seuil de décision clair pour séparer les observations. En l'absence de ce seuil prédéfini, il peut être ambigu de classer une observation comme normale ou anormale. C'est pourquoi une autre approche couramment utilisée pour l'évaluation est la courbe ROC.

2.4.2.2. La courbe ROC

La courbe *Receiver Operating Characteristics* (ROC) a été introduite peu avant l'an 2000 pour des problèmes de classification par Provost et Fawcett (1999). Elle illustre le taux de vrais positifs (sensibilité) en fonction du taux de faux positifs ($1 - spécificité$) pour différents seuils de décision. Elle permet d'évaluer la capacité d'un modèle à distinguer les observations normales de celles anormales sans être liée à un seuil de décision spécifique. Une présentation d'une courbe ROC est illustrée dans la Figure 8 inspirée de Tharwat (2021). Le seuil de décision le plus adapté est souvent choisi comme étant celui correspondant à la plus petite distance d entre la courbe et le point de coordonnées (0, 1). Cependant, la courbe ROC est surtout un atout visuel et ne permet que difficilement de comparer différents modèles entre eux. Par conséquent, une approche plus traditionnelle consiste à calculer l'aire sous la courbe ROC (Tharwat, 2021). L'aire sous la courbe ROC, (AUC-ROC pour *Area Under the Curve ROC*), est une métrique couramment utilisée pour

quantifier la performance globale d'un modèle de classification. Une valeur d'AUC-ROC de 1 indique une performance parfaite, tandis qu'une valeur d'AUC-ROC de 0,5 suggère que le modèle ne fait pas mieux qu'une classification aléatoire.

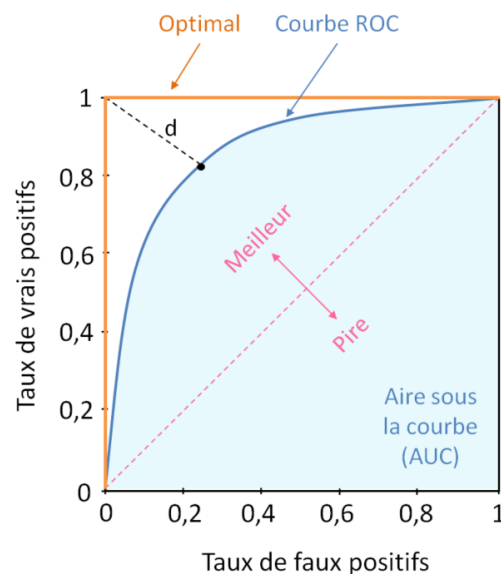


Figure 8 Exemples de représentation de courbes ROC. Une courbe ROC classique (bleu foncé) et l'aire sous la courbe (AUC-ROC) (zone en bleu clair) reflètent les performances d'un modèle pour une tâche de classification. La courbe en orange représente un modèle optimal avec une valeur de AUC-ROC de 1. La ligne rose représente une classification aléatoire. Inspirée de Tharwat (2021).

2.4.3. La famille des méthodes ou techniques de détection d'anomalies

Pour répondre aux besoins spécifiques de chaque domaine d'application, qui ont présentés dans la partie 2.3.3, de nombreuses techniques de détection d'anomalies sont disponibles et utilisées. Ces techniques ont leurs propres spécificités, telles que la façon dont elles présentent les résultats (p. ex. un pourcentage ou l'appartenance à une classe), la méthode d'entraînement (p. ex. supervisée ou non supervisée) ou l'approche scientifique sous-jacente (p. ex. statistique ou basée sur la modélisation). Cependant, elles peuvent être regroupées en plusieurs familles distinctes.

La Figure 9, inspirée des travaux de Chandola et al. (2009) et Himeur et al. (2021b) présente une taxonomie des familles de techniques de détection d'anomalie. Elle recense sept familles principales de techniques basées sur la **classification**, les **statistiques**, le **clustering**, le **voisinage**, la **régression**, la **réduction de dimension et reconstruction** et les **connaissances expertes**. Il est important de noter que cette figure ne prétend pas être exhaustive, car de nombreuses autres techniques existent dans la littérature. Cette partie a pour objectif de présenter succinctement ces sept familles de détection d'anomalies et leurs éventuelles applications dans les domaines des réseaux de capteurs et des mesures issues du bâtiment. Dans le but de proposer un aperçu général des familles de techniques applicables à la détection d'anomalies, cette partie ne se limite pas

exclusivement aux techniques d'apprentissage automatique. Les modèles utilisés dans cette thèse seront présentés et détaillés ultérieurement.

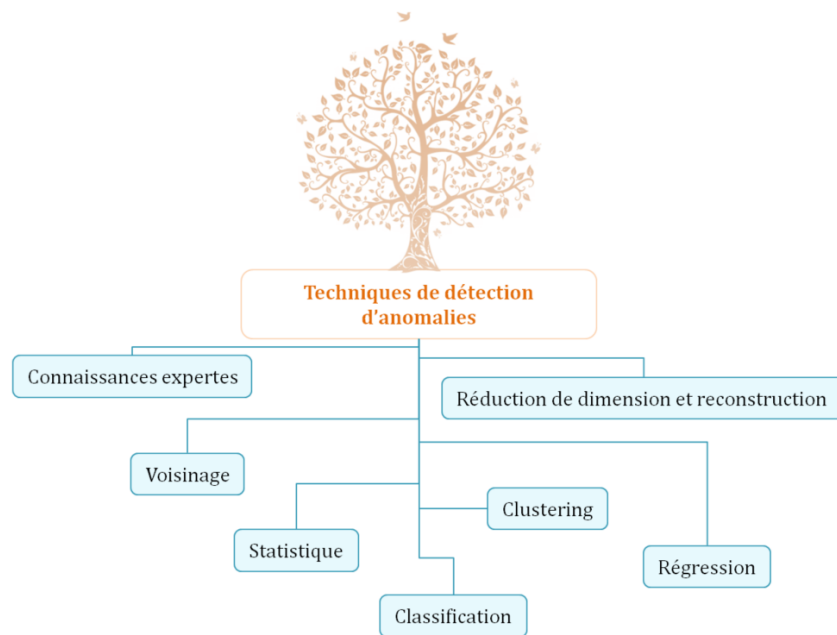


Figure 9 Les principales familles de techniques de détection d'anomalies.

2.4.3.1. Méthodes de classification

Les méthodes de classifications partent du postulat qu'il est possible de séparer les observations ou les séquences d'observations en différentes catégories ou classes. La majorité de ces méthodes se basent sur des techniques d'apprentissage supervisé ou semi-supervisé. En effet, elles impliquent généralement d'entraîner un modèle sur des données pour lesquelles les étiquettes sont connues, puis d'utiliser ce modèle pour classer de nouvelles observations non étiquetées. Pour chaque observation renseignée, ces méthodes peuvent prédire son appartenance à une classe, soit en attribuant une variable discrète définie au préalable (p. ex., 1 pour « pièce occupée » et 0 pour « pièce inoccupée »), soit en fournissant un pourcentage (p. ex., une probabilité de 85 % qu'une pièce soit occupée ou inoccupée). Dans le contexte de la détection d'anomalies, cela revient à estimer ou classer une observation comme étant normale ou anormale.

Dans la littérature, les principales techniques associées aux méthodes de classification comprennent les méthodes basées sur les règles, telles que les arbres de décision, les modèles probabilistes tels que les réseaux bayésiens, les réseaux de neurones artificiels et les Séparateurs à Vaste Marge (SVM). Plusieurs de ces méthodes sont appliquées dans le cadre de cette thèse, et une attention particulière leur sera consacrée ultérieurement. Afin d'illustrer quelques cas d'application, peuvent être cités : les travaux de Phan (2022),

qui utilise des réseaux bayésiens pour classer les activités des ménages selon leur consommation électrique ; Zou et al. (2018), qui appliquent un modèle provenant d'une sous famille des réseaux de neurones pour classer les activités des occupants au regard des changements de propagation des signaux Wi-Fi d'objets connectés ; Haq et al. (2021), qui utilisent une combinaison de réseaux de neurones et de séparateurs à vaste marge pour détecter des consommations électriques suspectes (comme le vol d'électricité) ; enfin Coma-Puig et al. (2016), qui détectent des erreurs dans des consommations électriques *via* plusieurs modèles, dont des réseaux bayésiens, des arbres de décision, des réseaux de neurones et des SVM.

2.4.3.2. Méthodes de clustering

Les méthodes de *clustering*, également appelées méthodes de *partitionnement des données*, ont pour principe de regrouper des observations ou des séquences d'observations similaires en groupes distincts appelés *clusters*. La plupart de ces méthodes se fonde sur des approches d'entraînement non supervisées. Appliqué à la détection d'anomalies, il existe trois modes opératoires distincts pour classer les observations comme normales ou anormales (Kraiem, 2021).

- (i) Le premier mode suppose qu'il est possible de regrouper la totalité des observations normales en un *cluster* unique, et que toute observation n'appartenant pas à ce *cluster* est, de ce fait, anormale.
- (ii) Le deuxième mode opératoire suppose que les observations normales se situent à proximité du centre de gravité du *cluster* le plus proche, contrairement aux observations anormales qui en sont éloignées.
- (iii) Enfin, le troisième mode opératoire suppose que les observations normales sont regroupées en forte concentration dans des *clusters* dits « denses », contrairement aux observations anormales qui se regroupent en petit *clusters* de faible densité.

Dans la littérature, les principales techniques associées aux méthodes de *clustering* comprennent les méthodes hiérarchiques, telles que la *Balanced Iterative Reducing and Clustering using Hierarchies* (BIRCH), les méthodes par partitionnement, comme la très utilisée *K-Means*, et les méthodes par densité, souvent représentées par la *Density-Based Spatial Clustering of Applications with Noise* (DBSCAN) (Ariyaluran Habeeb et al., 2022; Mehrotra et al., 2017). Cependant, il existe plusieurs autres approches possibles, comme le détaillent Xu et Tian (2015). Pour illustrer ces trois méthodes, leur fonctionnement peut être abordé comme suit :

- **Les méthodes hiérarchiques** reposent sur des mesures de distances comme la distance euclidienne ou la distance de Manhattan. Initialement, chaque observation

est considérée comme un *cluster* individuel qui fusionne progressivement avec le *cluster* le plus proche jusqu'à ce que toutes les données soient regroupées en un seul même *cluster* ou qu'elles soient trop éloignées. Cette méthode produit un dendrogramme permettant de classer les observations (Madhulatha, 2012).

- **Les méthodes par partitionnement** visent à diviser un ensemble d'observations en K_c *clusters*, où K_c est un nombre prédéfini. L'objectif est d'optimiser un certain critère (par exemple, minimiser la distance intra-cluster tout en maximisant la distance inter-cluster) pour former ces partitions. Le cas particulier de l'approche des centroïdes est utilisé (Kraiem, 2021), à titre d'exemple, pour en illustrer le processus. Au départ, K_c observations sont sélectionnées aléatoirement comme centroïdes des *clusters*, puis chaque observation est assignée au centroïde le plus proche, lequel est actualisé à chaque nouvel ajout. Le processus prend fin lorsque les centroïdes ne changent plus ou lorsque les observations restent dans les mêmes *clusters* après plusieurs itérations.
- **Les méthodes basées sur la densité** commencent par choisir une observation aléatoirement, puis un cluster se forme si un nombre suffisant d'observations se trouvent à une certaine distance, notée d , de cette observation initiale. L'observation initiale est alors considérée comme à *forte densité*. Plusieurs observations à forte densité peuvent être proches les unes des autres créant ainsi un *cluster* qui s'étend. Ainsi toute observation qui n'est pas à forte densité et qui n'en comporte pas dans son entourage proche est considérée anormale (Xu and Tian, 2015).

Comme chaque méthode forme et évalue différemment les clusters, ce choix peut avoir un impact significatif sur les résultats, comme le montre la Figure 10, issue des travaux de Bataineh (2022), qui compare les méthodes *K-means* et DBSCAN.

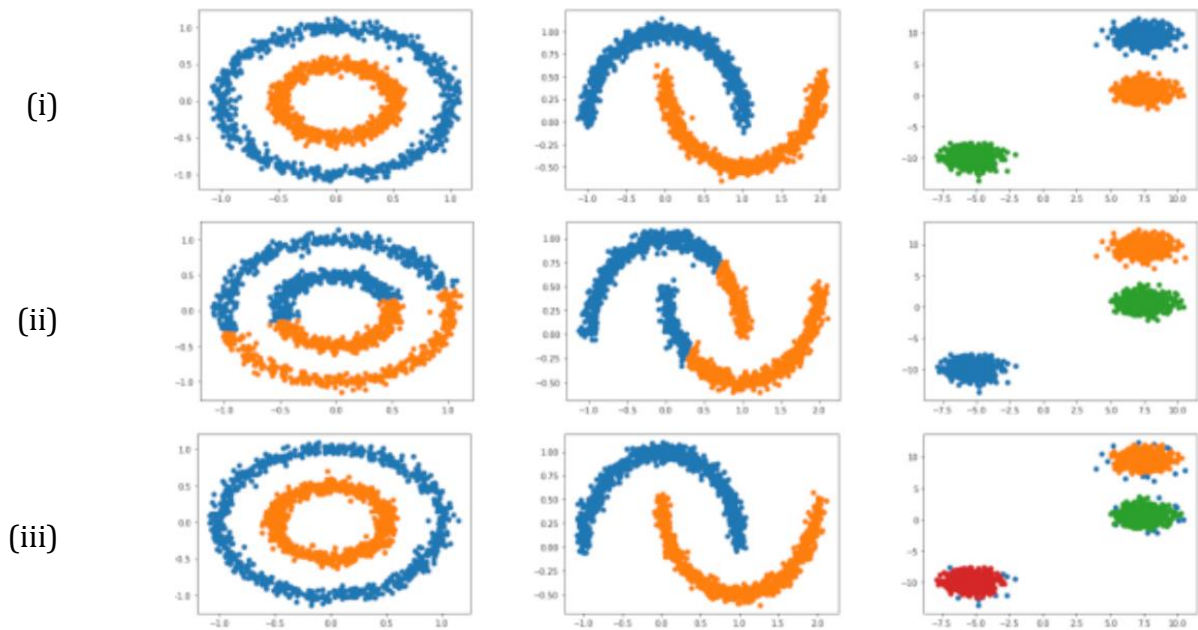


Figure 10 Résultats visuels proposés par différentes méthodes de clustering, (i) la vérité terrain, (ii) les résultats de la méthode K-means et (iii) les résultats de la méthode DBSCAN. Chaque couleur, bleue, orange, verte et rouge représente un cluster ou des observations hors cluster. Issue de Bataineh (2022).

Des d'applications concrètes peuvent être illustrées par quelques exemples. Gunay et Shi (2020) ont utilisé des méthodes hiérarchiques et de partitionnement pour détecter les défaillances dans un système de régulation CVC. Zhang et al. (2020) ont appliqué la méthode des *K-means* centroïdes conjointement aux taux de distorsion harmonique pour identifier des équipements ménagers selon leur appel de puissance. Enfin, Prabhakaran et al. (2022) ont aussi appliqué la méthode des *K-means* centroïdes sur des consommations électriques pour estimer l'occupation d'un bureau.

2.4.3.3. Méthodes de régression

Les méthodes de régression visent à comprendre les tendances et les comportements récurrents des observations passées, ce qui permet au modèle d'établir des relations de cause à effet pour prédire les valeurs futures des observations (Ducharlet et al., 2022). Contrairement aux autres techniques abordées dans cette section 2.4.3, les modèles de régression ne prédisent pas directement l'appartenance à une classe (normale ou anormale), mais ils estiment la valeur prédite de l'observation à venir (p. ex., une valeur discrète pour une température intérieure). Cette valeur prédite est considérée comme la référence de ce qui est « normal ». Ainsi, l'écart entre la valeur prédite par le modèle et la valeur réelle de l'observation est étudié pour déterminer si une observation est anormale ou non. Plus cet écart est important, plus l'observation a de chances d'être considérée comme anormale. La majorité de ces méthodes se basent sur des techniques d'entraînement supervisées ou semi-supervisées (Foulon, 2020).

Dans la littérature, les méthodes de régression sont associées à plusieurs techniques, notamment les méthodes de régression linéaire, les méthodes autorégressives comme l'*Autoregressive Integrated Moving-Average* (ARIMA), les réseaux de neurones et les séparateurs à vaste marge. Plusieurs de ces méthodes sont appliquées dans ces travaux, et des explications détaillées seront fournies ultérieurement. Plusieurs études peuvent être présentées en exemple afin d'illustrer certaines applications de ces méthodes tels que Candelieri (2017), qui utilise des modèles à supports à vaste marge pour prédire les niveaux de consommation d'eau à différentes échelles, urbaine ou individuelle, afin d'optimiser les cycles de fonctionnement des pompes et détecter les fraudes ou les anomalies. Chou et Telaga (2014) utilisent plusieurs méthodes, dont des réseaux de neurones et ARIMA, pour détecter les anomalies de consommation électrique. Enfin, Alduailij et al. (2021) appliquent plusieurs méthodes de régression, notamment la régression linéaire, la régression dynamique, ARIMA, ainsi que des modèles de réseaux de neurones complexes pour prédire des pics de consommation énergétique et assister l'occupant dans l'optimisation de sa gestion de l'énergie.

2.4.3.4. Méthodes statistiques

Les méthodes statistiques partent du postulat qu'il est possible d'évaluer le degré d'anomalie d'une observation ou d'une séquence d'observations par rapport à un modèle stochastique (ou statistique) préalablement établi. Ce modèle représente un comportement normal ou attendu, associé à un intervalle de confiance. Ainsi, plus une observation a de faibles probabilités d'être générée par le modèle (c'est-à-dire plus elle s'éloigne de l'intervalle de confiance), plus elle a de chances d'être considérée comme anormale. Ces méthodes attribuent un score d'anomalie à chaque observation, caractérisant ainsi cette déviation au modèle. Une observation avec un score élevé est considérée comme anormale tandis qu'une observation avec un score faible est considérée comme normale. Les méthodes statistiques peuvent être généralement séparées en deux approches : paramétriques et non-paramétriques (Zhang, 2013).

- **L'approche paramétrique** suppose que la distribution des observations est connue *a priori* (p. ex., une distribution normale) et construit un modèle en se basant soit sur les paramètres de distribution s'ils sont connus, soit en les estimant à partir des observations. Les observations qui ne correspondent pas à cette distribution sont considérées comme anormales.
- **L'approche non-paramétrique**, en revanche, ne suppose aucune connaissance préalable de la distribution des données. Elle construit le modèle uniquement en fonction des observations étudiées, ce qui la rend moins limitée dans son utilisation (Desforges et al., 1998). Dans cette approche, les observations qui se situent dans des régions de faible densité sont considérées comme des anomalies.

Ces méthodes peuvent être utilisées directement sur les ensembles d'observations ou conjointement à d'autres méthodes telles que celles de régression. Dans la littérature, les principales techniques associées aux méthodes statistiques incluent celles basées sur les modèles gaussiens, comme les boîtes à moustaches (ou *box-plot*) (Aggarwal, 2017), les méthodes basées sur des histogrammes comme l'*Histogram-based Outlier Score* (HBOS) ou encore les méthodes basées sur des noyaux comme la méthode *Parzen windows observation* (Bishop, 2006; Zhang, 2013). Pour illustrer quelques cas d'application, on peut citer l'exemple de Jakkula et Cook (2010) qui utilisent l'écart-type des données pour détecter des anomalies dans des consommations électriques ; Zhang et al. (2023) qui appliquent la méthode des histogrammes HBOS conjointement à de nombreuses autres méthodes (telles que la régression et les méthodes de voisinage) pour détecter des anomalies dans des consommations électriques ; ou encore l'exemple de Martin Nascimento et al. (2021), qui utilisent plusieurs méthodes statiques comme la règle des trois sigmas, la déviation absolue moyenne ou des boîtes à moustaches pour détecter des données anormales dans les mesures de consommation électrique après l'utilisation d'un modèle de régression.

2.4.3.5. Méthodes basées sur le voisinage

Les méthodes basées sur le voisinage (ou proximité) fonctionnent sur le même principe que les méthodes de *clustering* et suivent un approche d'entraînement principalement non supervisée. Elles partent du principe que les observations normales sont généralement proches les unes des autres, tandis que les observations anormales sont éloignées. Ainsi, une observation entourée de plusieurs observations voisines proches est considérée comme normale, tandis qu'une observation entourée d'observations éloignée est considérée comme anormale. Pour la détection d'anomalies, il existe deux modes opératoires distincts utilisés pour classer les observations comme normales ou anormales.

- (i) Le premier mode opératoire consiste à évaluer la distance entre une observation donnée et toutes les autres observations pour déterminer ses k plus proches voisins. En d'autres termes, cela revient à mesurer la disparité entre l'ensemble des observations étudiées (Chouchane, 2018) et à considérer comme anormale toute observation distante de la majorité des autres observations. Cette distance peut être calculée de différentes manières, telles que la distance euclidienne ou la distance de Manhattan.
- (ii) Le deuxième mode opératoire consiste à évaluer la densité locale d'une observation en calculant les distances entre cette observation et ses k plus proches voisins. Une observation est considérée comme anormale si sa densité locale est significativement faible vis-à-vis des autres observations.

Dans la littérature, les principales techniques associées aux méthodes de voisinage incluent les méthodes basées sur la proximité, comme les K plus proches voisins (ou KNN pour *K-Nearest Neighbours*), ainsi que les techniques basées sur la densité telles que le *Local Outlier Factor* (ou LOF) (Chandola et al., 2009; Ducharlet et al., 2022). De plus, l'approche par *Isolation Forest* (ou IF) (Foulon, 2020), bien qu'elle puisse être considérée comme une famille à part entière (Schmidl et al., 2022), est également largement utilisée. Plusieurs de ces méthodes sont appliquées dans ces travaux, et des explications détaillées seront fournies ultérieurement. Des d'applications concrètes peuvent être illustrées par quelques exemples. Borhani Afuosi et Zoghi (2020) utilisent plusieurs variantes de modèles des K plus proches voisins pour estimer la position spatiale d'occupants en se basant sur le signal Wi-Fi, dans le but d'optimiser la gestion de l'énergie d'un bâtiment. Himeur et al (2021a) évaluent plusieurs modèles, dont des KNN et des LOF, pour la détection d'anomalies dans les consommations électriques. Enfin, Zhang et al. (2023) appliquent conjointement les méthodes KNN, LOF et IF pour détecter des anomalies dans les consommations électriques d'un bâtiment.

2.4.3.6. Méthodes basées sur les connaissances expertes

Les méthodes basées sur les connaissances expertes reposent sur des règles définies par des experts du domaine. Ces règles sont conçues pour capturer des motifs spécifiques ou des comportements considérés comme anormaux, en se basant sur la compréhension experte du processus ou du contexte analysé (Kraiem, 2021). Elles peuvent être utilisées directement sur l'ensemble des observations disponibles ou conjointement à d'autres méthodes pour filtrer les résultats ou préparer les données (p. ex., en utilisant la proximité ou le *clustering*) (Himeur et al., 2021b). À ce titre, Kraeim (2021) identifie deux approches différentes basées sur les règles ou les machines à état finis :

- **Les approches à base de règles** utilisent un ensemble de règles de type « *SI-ALORS* » (de l'anglais 'IF-THEN'). Ces règles représentent des conditions et des actions correspondantes. Elles permettent de formaliser des connaissances de manière simple et sont faciles à interpréter. Elles peuvent être simples, imbriquées avec des sous-conditions, ou combinées pour former des ensembles plus complexes, tels que des chaînes de règles ou des arbres de décision.
- **Les machines à états finis** ou *Finite State Machines*, sont des modèles mathématiques utilisés pour modéliser des comportements séquentiels. Elles sont composées d'un ensemble d'états, dont un état initial, et d'une fonction de transition qui détermine le passage d'un état à un autre en fonction d'une entrée donnée (Ben-Ari and Mondada, 2018). Une transition est activée lorsque l'entrée satisfait certaines conditions liées à l'état courant. Les machines à états finis peuvent être *déterministes*, ce qui signifie qu'il existe une transition unique pour chaque état et chaque entrée, ou *non déterministes*,

avec zéro ou plusieurs transitions possibles pour chaque état et entrée. Cette approche permet de modéliser des ensembles complexes de règles. Un exemple d'application dans la détection d'anomalies, tel que l'utilisation sur des *clusters* Kubernetes, est présenté par Cao et al. (2022).

Quelques cas d'application des méthodes basées sur les connaissances expertes sont illustrés dans les travaux de Peña et al. (2016), qui ont mis en place un ensemble de sept règles *SI-ALORS* basées sur leur expertise dans les systèmes CVC, les mesures d'ambiance et l'historique des données disponibles. Ces règles sont utilisées pour détecter des comportements anormaux et, *in fine*, améliorer la gestion énergétique d'un bâtiment. Wan Yen et al. (2019) ont appliqué trois règles *SI-ALORS* sur les tensions de phases de lignes pour détecter des observations anormales de courte durée dans un réseau électrique intelligent (*smart grid*). De plus, Himeur et al. (2020) ont extrait des micro-moments à l'aide d'un modèle basé sur des règles, à partir de données de consommation électrique, en se concentrant sur des activités liées aux équipements ménagers, comme la mise en marche ou une consommation excessive. Ils ont ensuite entraîné un algorithme de classification à détecter ces micro-moments.

2.4.3.7. Méthodes de réduction de dimension et de reconstruction

Les méthodes de réduction de dimension, également appelées méthodes d'analyse spectrale, partent du postulat qu'il est possible de réduire l'espace vectoriel où sont représentées les observations étudiées en un sous-espace plus restreint tout en préservant l'essentiel de leurs informations. Appliquées à la détection d'anomalie, elles visent à déterminer un sous-espace (*via* des projections, des rotations, *etc.*) dans lequel les observations normales et anormales sont clairement distinctes et identifiables (Bishop, 2006; Duda et al., 2001). Ces méthodes sont souvent utilisées pour prétraiter des données de grande dimension et sont régulièrement couplées à d'autres techniques de détection d'anomalies, bien qu'elles puissent également être utilisées de manière indépendante (Himeur et al., 2021b). Les méthodes de réduction se basent sur des techniques d'entraînement supervisées, semi-supervisées ou non supervisées. Parallèlement, une autre approche consiste à compresser l'information en une représentation réduite, puis à reconstruire les données initiales à partir de cette représentation. Si une observation peut être reconstruite avec une faible erreur, elle est considérée comme normale (Chalapathy and Chawla, 2019). En revanche, une erreur de reconstruction élevée suggère qu'une observation pourrait être anormale. Ces méthodes, dites de reconstruction (Schmidl et al., 2022), se basent majoritairement sur des techniques d'apprentissage semi-supervisé, bien que certaines applications puissent se faire de façon non-supervisée (Norlander and Sopasakis, 2019).

Dans la littérature, les principales techniques paramétriques associées aux méthodes de réduction de dimension comprennent l'Analyse Discriminante Linéaire (ADL), l'Analyse en Composante Principale (ACP) et les réseaux de neurones (Himeur et al., 2021b). Pour les méthodes de reconstruction, les réseaux de neurones artificiels sont les plus utilisés, notamment les Auto-Encodeur et les Encodeur-Décodeur (Chalapathy and Chawla, 2019). Plusieurs de ces méthodes sont appliquées dans cette thèse, et une attention particulière leur sera consacrée ultérieurement. Parmi des exemples d'applications figurent, l'utilisation de l'ADL par Himeur et al. (2021a) pour classer en cinq catégories différentes les consommations électriques d'équipements ménagers. Li et Wen (2014) qui ont appliqué une méthode d'ACP aux mesures de température, d'humidité et de pression d'air pour détecter des erreurs dans une centrale de traitement d'air, tandis que Burgas et al. (2021) ont utilisé l'ACP pour réduire la dimension d'un nombre conséquent de mesures issues de capteurs, couplée à des méthodes statistiques, dans le but de détecter des erreurs ou des comportements anormaux relatifs à la gestion d'un bâtiment. Enfin, Fan et al. (2018) et Pereira et Silveira (2018) ont tous deux utilisés des variantes d'Auto-Encodeur pour détecter des anomalies dans consommations énergétiques en adoptant des approches semi-supervisées et non supervisées.

2.5. La diversité, à la fois une force et une contrainte

Comme détaillé tout au long de ce chapitre, la littérature offre une multitude de techniques de détection d'anomalies, chacune possédant ses propres caractéristiques et usages. Ces techniques répondent à divers besoins et contraintes. Dans le premier chapitre, nous avons souligné que les bâtiments connectés et intelligents génèrent une grande variété de données, caractérisées par des profils, des relations et des utilisations variées. Cette diversité, qui découle de l'hétérogénéité des données étudiées (Foulon, 2020), génère de nombreuses problématiques pouvant complexifier tant le choix que la mise en œuvre des techniques de détection d'anomalies.

La Figure 11, issue de la synthèse des travaux de Chandola et al. (2009) et de Chalapathy et al. (2019), illustre ce défi. Elle a été construite en associant les familles de techniques de détection d'anomalies recensées dans la littérature par ces deux études aux domaines d'applications correspondants. Du côté gauche, en se concentrant uniquement sur un domaine spécifique, tel que les réseaux de capteurs (représentés en bleu), on constate qu'il y a déjà un grand nombre de famille de détections d'anomalies à considérer, couvrant la quasi-totalité des familles recensées. En outre, du côté droit, on peut voir qu'en se limitant à une seule famille, comme les réseaux de neurones, il existe un nombre encore plus important de techniques ou de méthodes spécifiques liées au domaine d'application qui nous intéresse. Ce constat est également valable pour d'autres familles de techniques,

comme le montrent les travaux de Schmidl et al. (2022). Ils ont, en effet, évalué plus de 70 méthodes de détection d'anomalies différentes sur des séries temporelles, sélectionnées parmi plus de 150 techniques répertoriées dans de nombreuses publications. De plus, bien que le choix des techniques soit en partie intrinsèque à un domaine d'étude particulier ou à des données spécifiques, il n'est pas rare d'observer l'application d'un nombre conséquent techniques variées sur des données ou mesures similaires, voire identiques. On peut citer, par exemple, l'étude d'Himeur et al. (2021b) qui répertorie plus d'une trentaine de techniques de détection d'anomalies différentes appliquées aux seules consommations énergétiques du bâtiment.

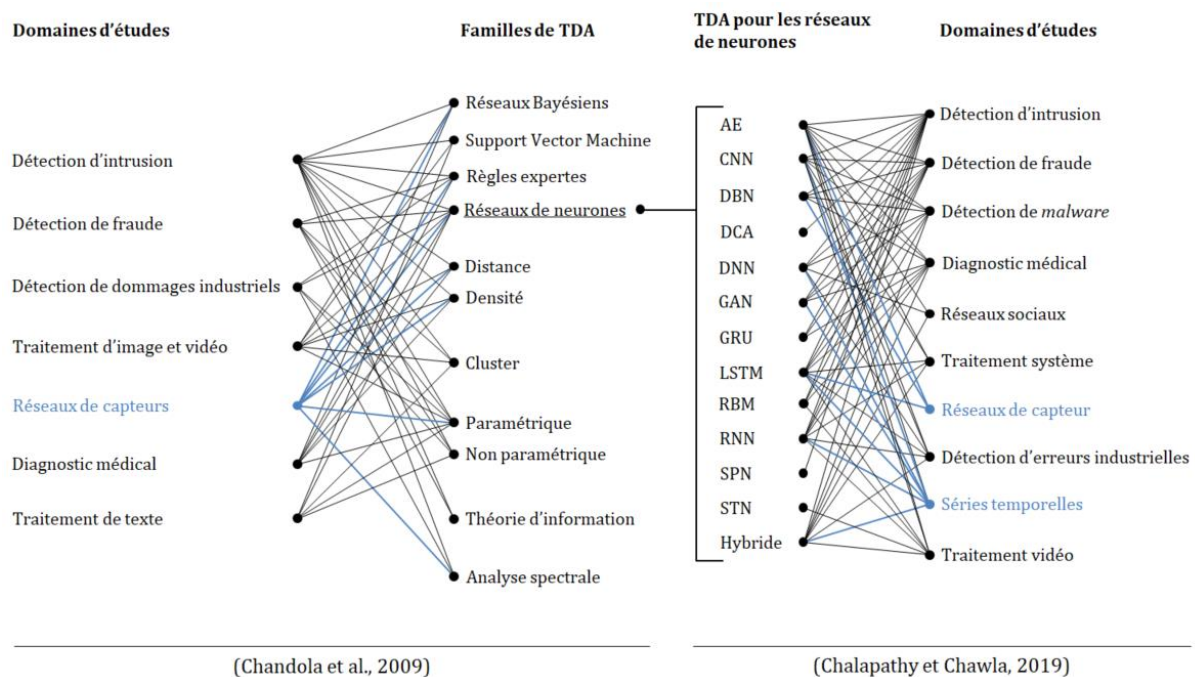


Figure 11 Répartition de l'utilisation des techniques de détection d'anomalies selon les domaines d'études, inspiré de Chalapathy and Chawla (2019) et Chandola et al. (2009). TDA : Technique de Détection d'Anomalies, AE : Autoencoder, CNN : Convolutional Neural Network, DBN : Deep Belief Network, DCA : Dilated Convolutional Autoencoder, DNN : Deep Neural Network, GAN : Generative Adversarial Network, GRU : Gated Recurrent Unit, LSTM : Long Short-Term Memory, RBM : Restricted Boltzmann Machine, RNN : Recurrent Neural Network, SPN : Sum Product Network, STN : Spatial Transformer Network, Hybride: modèle hybride composé de modèles différents appariés.

Il est également important de noter qu'une même méthode, paramétrée et entraînée à l'identique sur deux ensembles de données différents, qu'ils partagent des variables semblables ou non, peut produire des résultats qui diffèrent sensiblement. Cela est résumé de manière plus générale par Gundersen et al. (2023) : « *Different datasets lead to different outcomes* ». Ainsi, une technique utilisée pour détecter des anomalies dans les mesures de concentration de CO₂ d'un mur végétal (Liu et al., 2020) pourrait produire des résultats sensiblement différents si elle était appliquée aux données provenant d'un capteur mesurant la concentration de CO₂ dans une pièce. Pour cet exemple, de telles différences peuvent être attribuées à plusieurs facteurs, notamment des équipements de

mesures différents, des dynamiques de concentration de CO₂ distinctes ou la présence d'anomalies différentes. Un constat similaire peut également être observé si, l'entraînement effectué, ou les paramètres des modèles diffèrent (Pham et al., 2020). En résumé, de nombreux facteurs peuvent contribuer aux variations des résultats de détection d'anomalies, même lors de l'application d'une méthode similaire sur des ensembles de données d'apparence proches. Il est donc essentiel de prendre en compte ces facteurs et d'adapter les méthodes en fonction du contexte spécifique et des données disponibles.

La diversité des niveaux d'instrumentation dans le domaine du bâtiment ajoute une contrainte significative à la détection d'anomalies (Dong et al., 2019). En effet, chaque bâtiment ou pièce peut être équipé de capteurs de manière différente, rendant la qualité et la variété des données disponibles très inégales pour chaque étude. Par exemple, la détection d'anomalies dans les consommations électriques d'une maison peut produire des résultats très différents selon la disponibilité des données, qu'il s'agisse uniquement des consommations globales ou de celles de chaque pièce individuelle. Sélectionner des modèles *a priori* devient donc une tâche complexe, d'autant plus lorsqu'une publication ne présente les performances que d'un seul modèle appliqué à un cas donné, qui peut différer sensiblement des objectifs d'une autre étude. La tâche peut se révéler encore plus délicate si les métriques d'évaluation utilisées sont différentes d'une publication à une autre, telles que la précision, le F₁-score, ou l'AUC-ROC.

De manière plus générale, la reproductibilité des méthodes et des résultats observés dans la littérature présente souvent des défis. Gundersen et al. (2023) ont identifié 41 causes potentielles de ce phénomène, qu'ils ont réparti en six catégories distinctes, et liées (1) à la conception de l'étude, (2) aux algorithmes, (3) à leur implémentation, (4) aux données utilisées ainsi qu'au (5) processus d'évaluation et (6) de documentation. Plus précisément, certaines « lacunes » fréquentes dans les études de détection d'anomalies, telles qu'une formulation floue du problème, un manque d'informations quant au paramétrage, la structure et l'entraînement des méthodes ainsi qu'une documentation insuffisante sur la génération, le prétraitement, le processus de sélection et de transformation des données, ont été relevées et concordent avec celles d'autres auteurs comme Schmidl et al. (2022). Les travaux de Wu et Keogh (2021) renforcent cette constatation en illustrant certains défauts présents sur des ensembles de données populaires utilisés dans de nombreuses études, notamment en ce qui concerne l'étiquetage des anomalies et leur densité, allant même jusqu'à titrer '*Current Time Series Anomaly Detection Benchmarks are Flawed and are Creating the Illusion of Progress*'. Ces éléments impactent directement la reproductibilité des études et peuvent être considérés comme des freins cachés difficiles, voire impossibles, à prendre en compte lors du choix des méthodes.

En résumé, la sélection des méthodes les plus appropriées pour résoudre un problème donné peut être complexe, compte tenu du vaste choix de techniques et d'études disponibles, chacune adaptée à des cas d'usage variés, ainsi qu'à des contraintes inhérentes. Ce défi est d'autant plus prononcé dans le cadre de cette thèse, où des données aux profils et aux usages variés sont considérées.

2.6. Contexte et motivations de la thèse

À notre connaissance, il n'existe actuellement aucune étude évaluant différentes techniques de détection d'anomalies en utilisant une base de données commune contenant des anomalies provenant de mesures dans des bâtiments. Bien que certains travaux récents s'en rapprochent, comme celui de Schmidl et al. (2022), qui s'intéresse à un ensemble de séries temporelles, ils présentent cependant plusieurs limitations courantes dans la littérature. Limitations compréhensibles au regard de la quantité de techniques et de jeux de données évalués, qui sont respectivement au nombre de 71 et 958. Ainsi, la grande majorité des données utilisées dans ces études sont soit artificielles (comme le jeu de données *GutenTAG*), soit proviennent d'autres domaines d'application avec des caractéristiques spécifiques. Par exemple, avec le jeu de données *Daphnet* (Bachlin et al., 2010) qui contient des données de capteurs d'accéléromètres de patients atteints de la maladie de Parkinson, tandis que les ensembles de données *MITDB* (Moody and Mark, 2001) et *SVDB* (Greenwald et al., 1992) contiennent des données d'électrocardiogrammes. Il y a également un manque récurrent d'explications détaillées concernant l'implémentation et la représentation des anomalies, ce qui peut rendre leur interprétation difficile. De plus, et de façon analogue à certaines études, peu ou pas d'informations sont apportées sur le processus de sélection (Tan et al., 2014) ou sur les techniques de transformation des données appliquées (*feature engineering*), qui sont pourtant essentielles pour optimiser les performances des méthodes utilisées (Bengio et al., 2014). On dénombre aussi plusieurs travaux comme ceux de Liu et al. (2020), qui comparent plusieurs méthodes et procédures, mais en ne se concentrant que sur une unique variable mesurée et un nombre limité d'anomalies. D'autres travaux comme ceux de Kraiem (2021), appliquent un nombre plus restreint de modèles mais sur un éventail varié d'anomalies et de données. Il existe donc actuellement un flou quant aux performances potentielles des techniques de détection d'anomalies lorsqu'elles sont appliquées à des données et des anomalies variées propres au domaine du bâtiment.

Ainsi, cette thèse vise à développer une méthodologie de sélection de techniques de détection d'anomalies basée sur l'apprentissage automatique. Cette méthodologie a pour but de comparer différentes techniques de détection d'anomalies, provenant de plusieurs familles existantes, en les évaluant à l'aide de mesures et d'anomalies issues du domaine

du bâtiment. Comme mentionné précédemment, il est illusoire de penser définir les performances générales d'une ou plusieurs techniques à partir d'une seule étude, et donc de les classer ou les catégoriser de manière absolue. Par conséquent, ces travaux ne visent pas à déterminer la technique la plus performante, mais plutôt, proposer une méthode permettant d'acquérir une vision éclairée des forces et faiblesses de ces techniques au regard des spécificités apparentes des mesures étudiées. L'objectif final est de fournir des orientations pour choisir les techniques les plus appropriées en fonction des données disponibles ou sélectionnées, des types d'anomalies présentes ou ciblées, et des exigences des modèles envisagés. Finalement, il s'agit de contribuer à une compréhension globale des différentes problématiques liées à l'étude des mesures du bâtiment, tout en apportant une aide quant au choix des techniques adaptées à des problèmes donnés.

2.7. Conclusion

En conclusion de ce chapitre, nous avons exploré les divers aspects de la détection d'anomalies. Nous avons souligné la nature variable et subjective du concept d'anomalie, ainsi que les nombreux défis auxquels sont confrontés les processus de détection. En présentant divers domaines d'applications, nous avons mis en évidence la grande variété des problèmes, des données et des anomalies traitées dans la littérature. Cette variété influence fortement le choix et le nombre de familles et de techniques à considérer, chacune ayant ses propres avantages et inconvénients.

Nous avons également pu montrer que cette diversité, combinée aux contraintes liées à l'adaptation des méthodes aux spécificités des données, au manque de normes de comparaison entre les techniques et à l'hétérogénéité des études, pose de nombreux défis dans le contexte de la détection d'anomalies appliquée le domaine du bâtiment.

Le prochain chapitre se concentrera sur la méthodologie proposée pour aborder ces défis et évaluer de manière approfondie les techniques de détection d'anomalies dans le contexte du bâtiment connecté.

Chapitre 3

Méthodologie de sélection de techniques de détection d'anomalies appliquée aux mesures du bâtiment

Ce chapitre vise à présenter une méthodologie de sélection de techniques de détection d'anomalies appliquées aux mesures provenant du bâtiment. Cette approche consiste à, dans un premier temps, formaliser divers cas d'application représentatifs des données (section 3.2) et des anomalies (section 3.3) spécifiques au domaine du bâtiment, créant ainsi une taxonomie. Cette taxonomie, présentée dans la section 3.4, servira de base pour comparer et classer différentes techniques en fonction des problématiques uniques du domaine du bâtiment. Cette classification permettra par la suite de sélectionner les techniques les plus adaptées pour des mesures et des anomalies données. Le chapitre se terminera par une synthèse dans la section 3.5.

« Tu m'as acceptée comme professeur. Quelle leçon aurais-je pu t'enseigner si je t'avais donné la réponse avant que tu sois en mesure de formuler la question ? »

Extrait traduit de, L'Alliance de fer : le paria, d'Anthony Ryan

Sommaire du chapitre 3

3.1. Contexte et démarche	57
3.2. La nature des données.....	57
3.2.1. La nature des données issues de mesures du bâtiment	58
3.2.2. Caractéristiques des données considérées dans la littérature	62
3.2.3. Caractéristiques considérées pour la thèse	63
3.3. Les types d'anomalies	64
3.3.1. Contexte	64
3.3.2. Les types d'anomalies rencontrées dans les mesures.....	64
3.3.3. La caractérisation des anomalies au travers les typologies	67
3.3.4. Mise en place d'une typologie d'anomalies adaptée	70
3.4. Méthodologie de sélection des techniques de détection.....	75
3.4.1. La taxonomie des problématiques rencontrées dans les mesures du bâtiment	75
3.4.2. Problématique du contexte dans les cas d'applications réels.....	77
3.5. Conclusion.....	79

3.1. Contexte et démarche

De nos jours, les techniques de détection d'anomalies sont largement utilisées dans divers domaines, chacune ayant ses avantages et inconvénients. Le choix des techniques appropriées pour des cas d'étude spécifiques nécessite une analyse conséquente de la littérature, ce qui peut être complexe, en particulier dans des domaines comme celui du bâtiment, où l'hétérogénéité des données représente un défi.

Comme évoqué dans le chapitre précédent, le choix et la comparaison des techniques de détection d'anomalies dépendent de la formulation spécifique du problème à résoudre. Chaque domaine d'application possède ses propres problématiques, mais celles relatives aux *données* et aux *types d'anomalies* sont des éléments récurrents (Blázquez-García et al., 2022; Chandola et al., 2009). Dans le domaine du bâtiment, nous avons déjà souligné la diversité des données et des anomalies, ainsi que les défis associés à leur détection. Ces éléments suggèrent l'existence de plusieurs problèmes distincts à considérer pour rendre l'application et l'évaluation des techniques de détection d'anomalie pertinentes. En effet, certaines méthodes – ou familles entières de méthodes – pourraient être considérablement impactées, en bien comme en mal, par les caractéristiques propres aux données ou aux anomalies. Il est donc essentiel de mettre en place une approche permettant de distinguer différents cas d'applications pertinents.

C'est pourquoi ce chapitre discutera en premier lieu des différentes caractéristiques des données et des anomalies issues de mesures provenant du bâtiment. Un état de l'art permettra ensuite d'identifier celles qui sont souvent considérées dans la littérature comme ayant un impact sur la détection d'anomalies et, donc, sur la sélection des techniques. Les éléments pertinents pour nos travaux, serviront à la construction d'une taxonomie⁷ des problématiques rencontrées dans notre domaine d'étude. Cette taxonomie servira de cadre commun et contrôlé, facilitant la comparaison d'un large éventail de techniques de détection d'anomalies en fonction des données et anomalies rencontrées. Elle constituera la base de la méthodologie de sélection.

3.2. La nature des données

Cette partie vise à introduire et discuter des caractéristiques constitutives des données provenant de mesures du bâtiment. En comprenant mieux ces caractéristiques, il est possible non seulement de mieux appréhender les défis inhérents aux données

⁷ Ici, au sens d'une théorie de classification suivant plusieurs critères

rencontrées, mais aussi identifier les caractéristiques susceptibles d'influencer le choix des techniques de détection d'anomalies.

3.2.1. La nature des données issues de mesures du bâtiment

Les données utilisées dans le contexte de la détection d'anomalies, notamment les mesures, jouent un rôle central. Elles sont régies par un ensemble de principes, de processus et de règles qui peuvent, dans certains cas, être approchées par des modèles, lesquels sont alors en capacité de différencier les éléments anormaux qui tendent à s'en éloigner (Mehrotra et al., 2017). De nombreuses caractéristiques relatives aux données, à leur récolte et au processus de génération des données, auront un impact significatif sur le choix et le succès, ou l'échec, de l'application de certaines techniques de détection. Cet impact doit être considéré, que les données soient la cible ou la source de la tâche de détection (Huang, 2018). Bien que le domaine du bâtiment englobe plusieurs types de données différentes comme des images ou des vidéos, la majorité des données traitées sont des séries temporelles (comme expliqué dans la partie 1.3). Par conséquent, ces travaux seront, exclusivement concentrés sur ce type de données.

Pour faciliter la suite des propos, le terme « *données* » sera utilisé pour désigner l'ensemble des *variables* ou des *mesures* issues du processus de mesurage, lesquelles sont constituées d'une succession d'*observations*. Selon la littérature (Tan et al., 2014), il est possible de définir la nature des données par les caractéristiques suivantes : le type des variables, le type de relation, la dynamique des variables et la présence de bruit. Ces caractéristiques sont présentées dans la suite de cette partie.

3.2.1.1. Le type des variables

Le type des variables se réfère aux valeurs que peuvent prendre les observations d'une variable, ce qui influe sur les opérations ou les méthodes d'analyse appropriées (Tan et al., 2014). Pour les mesures issues du bâtiment, on distingue généralement deux types de variables, les variables *quantitatives* (ou numériques) et les variables *qualitatives* (ou catégorielles) (Huang, 2018). Les variables quantitatives désignent des données dont les caractéristiques sont mesurables et quantifiables et le deuxième des données qui désignent des catégories ou des classes descriptives, ordonnées ou non (Tan et al., 2014). Par exemple, une mesure de la température extérieure en degrés Celsius est quantitative tandis qu'une mesure d'ouverture-fermeture de fenêtres (fermée – 0 ou ouverte – 1) est qualitative. Le Tableau 4 illustre la diversité des types de variables dans quelques exemples de données issues de mesures du bâtiment (Dong et al., 2019) réparties en deux catégories couramment utilisées (Yu et al., 2013).

Tableau 4 Exemples de données quantitatives et qualitatives dans le bâtiment.

Catégories de mesures	Quantitatives	Qualitatives
Conditions intérieures ou extérieures	Température (°C)	État des fenêtres (0-1)
	Humidité relative (%)	Occupation ou mouvement (0-1)
	Concentration de CO ₂ (ppm)	État des luminaires (0-1)
Gestion du bâtiment	Consommation ou production	
	électrique (kWh)	État des équipements (0-1)
	Consommation d'eau (m ³)	Position des vannes (0-1)
	Consigne de température (°C)	

3.2.1.2. Le type des relations

Les relations, aussi appelées dépendances (Foorthuis, 2021), font référence aux multiples corrélations qui existent tant au sein d'une même variable (intra-variable) qu'entre plusieurs (inter-variable). Les *relations intra-variables* représentent les dépendance entre les observations d'une même variable (Chandola et al., 2009) et sont parfois appelées « contexte » dans la littérature. Elles sont temporelles dans notre cas d'étude mais peuvent aussi être, spatiales, ou spatio-temporelles (Huang, 2018; Tan et al., 2014). Les *relations inter-variables*, quant à elles, représentent les dépendances entre les observations de différentes variables (Erhan et al., 2021; Hair, 2010). Une variable sans relations avec d'autres variables est usuellement appelée *univariée*, tandis que dans le cas contraire, elle sera dite *multivariée*. En revanche, une variable composée d'observations sans aucune relation (p. ex., un tirage aléatoire) est considérée comme indépendante (Foorthuis, 2021).

Les variables sont rarement indépendantes les unes des autres et peuvent, dans certaines situations, rendre compte de phénomènes anormaux qui ne seraient pas détectables si elles étaient examinées individuellement. Par exemple, un capteur mesurant plusieurs variables ou plusieurs capteurs situés dans un même espace peuvent rendre compte de tels phénomènes. La Figure 12 illustre un exemple de mesure de température intérieure en décembre, dont la valeur reste anormalement constante (★) malgré une ouverture de fenêtre, ce qui est clairement anormal en raison de la baisse significative et simultanée de la concentration de CO₂, ainsi qu'en tenant compte des comportements passés et futurs observés pour ces deux variables.

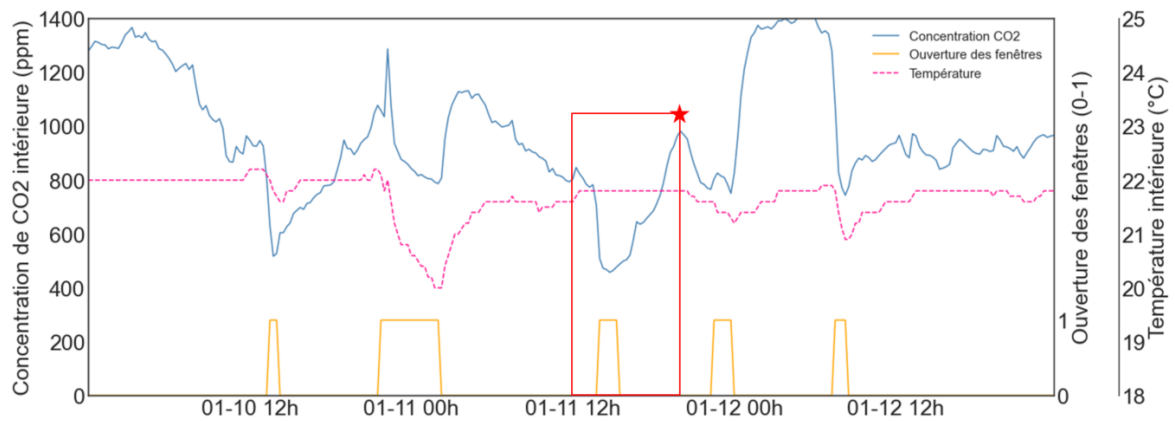


Figure 12 Exemple d'une anomalie (★) détectée par corrélation de variables.

3.2.1.3. La présence de bruit

Le terme *bruit* fait référence aux variations aléatoires ou données manquantes qui peuvent apparaître dans des données réelles issues de mesures (Tan et al., 2014) et qui ne sont pas liées à des phénomènes parasites ou d'intérêt (Aggarwal, 2017). Le bruit est inhérent au processus de mesurage et peut varier en fonction de la méthode utilisée pour collecter les données ainsi que de la variable mesurée (Foulon, 2020). Son impact est illustré dans la Figure 13, issue des travaux de Aggarwal (2017) et tiré de Foulon (2020). On observe qu'il rend la frontière entre les données normales et anormales moins claire, ce qui complique le processus de détection.

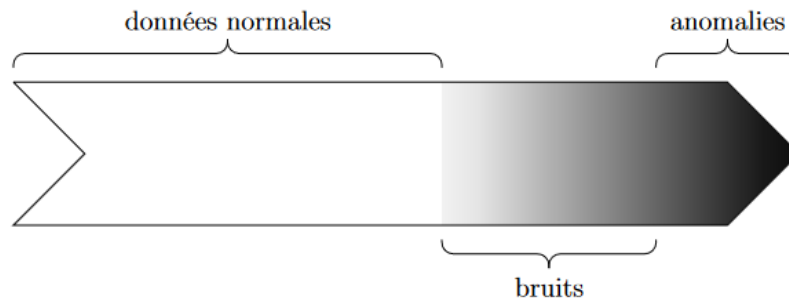


Figure 13 Spectre des données normales aux données normales avec présence de bruit. Issue de Aggarwal (2017) et tiré de Foulon (2020).

Le bruit est une problématique ancienne et récurrente en analyse de données, et il existe aujourd'hui de nombreuses approches de prétraitement pour réduire son impact, telles que le lissage ou l'agrégation (Foorthuis, 2021; Hird and McDermid, 2009; Kostelich and Yorke, 1990; Tan et al., 2014). Un exemple de lissage est illustré dans la Figure 14 avec une mesure de concentration de CO₂. Les données bruitées, telles que récoltées par le capteur sont présentées à gauche (i) alors que les mêmes données prétraitée au moyen

d'un lissage par moyenne mobile exponentielle (Thibault et al., 2023) sont présentées en (ii).

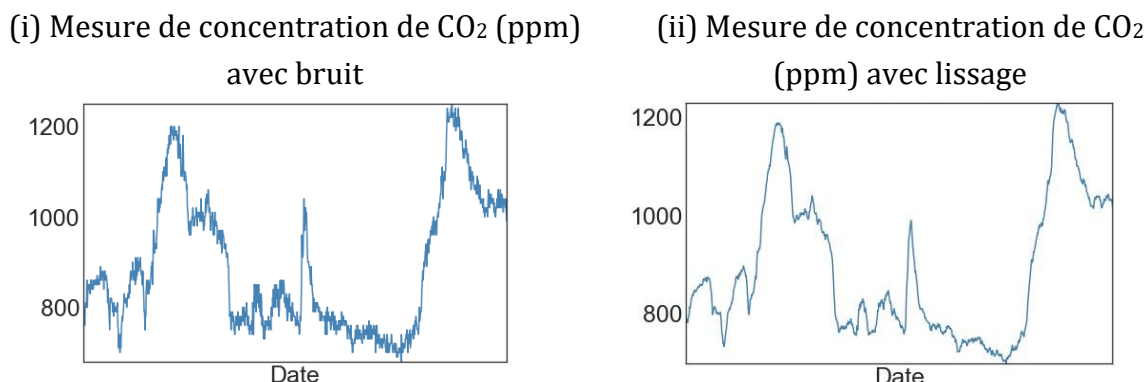


Figure 14 Exemple de lissage par moyenne mobile exponentielle de mesure de concentration de CO₂.

3.2.1.4. La dynamique des variables

La notion de dynamique fait référence aux variations temporelles ou contextuelles des observations au sein d'une variable. Les données réelles sont rarement statiques et peuvent évoluer avec le temps ou en réponse à différents phénomènes. La dynamique est intimement liée au processus qui génère les données. Elle peut se manifester par des motifs spécifiques qui se répètent de manière cyclique, périodique ou saisonnière, ainsi que par des changements temporaires de tendance ou de régime lorsque les changements sont définitifs (Foulon, 2020). Ces variations ont un impact sur la distribution des données, que ce soit pour une même variable ou pour plusieurs (Erhan et al., 2021). Elles influencent donc directement la définition de ce qui peut être considéré comme un comportement normal. Un comportement considéré normal à un moment donné peut devenir non représentatif de la normalité dans le futur (O'Reilly et al., 2014). Les dynamiques sont généralement difficiles à modéliser ou à généraliser car elles peuvent dépendre de facteurs externes rares. Dans certains cas, elles peuvent être confondues à des anomalies, ce qui complexifie le processus de détection (Foulon, 2020). Par conséquent, leur prise en compte nécessite l'utilisation de techniques ou d'approches adaptées (Tran et al., 2019).

Plusieurs exemples de dynamiques propres aux mesures du bâtiment sont illustrés dans la Figure 15. On y observe des variations naturelles journalières sur quelques jours de mesures de températures extérieures (i), un changement de tendance dans les mesures de concentration de CO₂ suite au départ des occupants sur plusieurs jours (ii), et enfin un changement de régime dans les consommations électriques suite à la remise à zéro du compteur le 1^{er} janvier (iii).

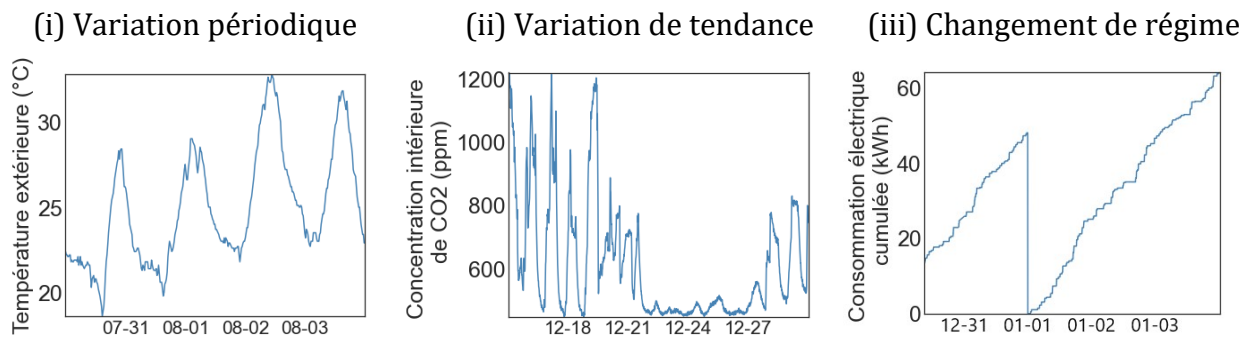


Figure 15 Exemple de dynamiques de données.

3.2.2. Caractéristiques des données considérées dans la littérature

Parmi les différentes caractéristiques liées à la nature des données discutées en amont, le *type des variables* semble être la plus souvent considérée dans la littérature lors du choix des techniques de détection (Chandola et al., 2009; Foorthuis, 2021; Huang, 2018). Cela s'explique par son influence à la fois sur la pertinence de certaines techniques (Chandola et al., 2009) et sur la manière dont les anomalies sont présentes et représentées (Foorthuis, 2021). Il est logique de constater que le type des variables a un impact significatif, car il détermine si certaines techniques peuvent être appliquées. Par exemple, des méthodes basées sur des calculs de distances, comme les K plus proches voisins, reposent exclusivement sur cette caractéristique pour évaluer si une observation est anormale ou non, comme l'illustrent les travaux de Chandola et al. (2009) ou encore Kalinichenko et al. (2014). Par ailleurs, il convient de noter que la littérature tend à mettre davantage l'accent sur le traitement des données quantitatives au détriment des données qualitatives, comme le révèlent les travaux de Taha et Hadi (2020). Cette disparité s'explique par le fait que de nombreuses techniques sont basées sur des approches difficiles à appliquer aux données qualitatives (p. ex., celles à base de distance ou de densité), bien que des adaptations existent (Boriah et al., 2008; Chandola et al., 2012).

Les *types de relations* entre les variables sont également régulièrement pris en compte, car ils constituent une source conséquente d'informations à exploiter, pouvant faciliter le processus de détection d'anomalies (Blázquez-García et al., 2022; Chandola et al., 2009; Foorthuis, 2021; Hair, 2010). Cependant, les capacités d'exploitation de cette information dépendent grandement des techniques utilisées. Certains modèles vont seulement considérer les relations inter-variables (p. ex., les séparateurs à vaste marge), tandis que d'autres tiennent compte uniquement des relations intra-variables (p. ex., la méthode ARIMA) ou des deux (p. ex., les réseaux de neurones) (Chandola et al., 2009; Schmidl et al., 2022). Ainsi, la présence, l'absence et la manière de prendre en compte ces relations influent directement sur le choix des techniques. Peut être citée, à titre d'exemple, l'étude de Blázquez-García et al. (2022) qui porte sur l'analyse des pratiques de détections

d'anomalies non supervisées appliquées aux séries temporelles dans la littérature. Cette étude se concentre uniquement sur les types de relations inter-variables pour classifier les données comme univariées ou multivariées, dans le but de constituer une caractéristique inhérentes aux données permettant de différencier l'applicabilité des techniques de détection recensées. Cependant, il convient de noter que les relations intra-variables (dans ce cas, temporelles) et le type de données (dans ce cas, quantitatives) sont implicitement définis par le cadre de l'étude.

La *dynamique* des données et le *bruit* sont deux aspects particuliers qui ne figurent pas directement parmi les critères principaux guidant le choix d'une technique de détection. Ils sont davantage perçus comme des défis compliquant la tâche de détection (Chandola et al., 2009; Foulon, 2020). Les deux génèrent une variabilité ou des variations propres à chaque mesure, ce qui peut être interprété comme une anomalie. Leur dépendance apparente à chaque mesure, application ou cas d'étude suggère qu'ils doivent être considérés de manière spécifique, au cas par cas. De plus, certaines techniques de détection d'anomalies sont particulièrement sensibles au bruit, comme celles basées sur la théorie de l'information (Kraiem, 2021) ou certains réseaux de neurones (Chawla and Wang, 2017).

3.2.3. Caractéristiques considérées pour la thèse

Au regard de ce qui a été détaillé dans cette partie, il apparaît difficile, à ce stade, de proposer une caractérisation générale qui reprendrait l'ensemble des caractéristiques évoquées et qui serait applicable à l'ensemble des données pouvant être extraites du bâtiment. En effet, bien que certaines caractéristiques telles que le type des variables ou la présence de bruit puissent séparer les données en deux groupes distincts, d'autres semblent fortement corrélées aux spécificités des cadres posés (ou permis) par une étude et par conséquent, aux données utilisées. Les dynamiques et les types des relations, notamment celles entre les variables, sont autant de spécificités pour lesquelles une généralisation *a priori* semble illusoire compte tenu du nombre de cas à envisager. Par conséquent, nous proposons d'adopter une approche individuelle pour chaque variable étudiée dans la suite de ces travaux.

Chaque mesure sera caractérisée par son **type**, ses **relations** potentielles, la présence de **bruit** ou de **dynamiques** spécifiques. Cette approche nous permettra de sélectionner plusieurs variables de natures variées et d'évaluer *a posteriori* l'impact sur le choix des techniques de détection d'anomalies.

La grande hétérogénéité des données rend la tâche de détection d'anomalies complexe, car la définition de la normalité et le choix des techniques peuvent varier d'une variable à

l'autre en fonction de leur nature (Foulon, 2020). L'**expertise métier**, qui repose sur la connaissance préalable des processus de génération des données, des corrélations intra et inter variables, ainsi que des complications potentielles, est donc essentielle pour mener à bien ce processus. Cependant, cela ne constitue qu'un aspect du problème, car l'anomalie en constitue un autre, qui sera développé dans la prochaine partie.

3.3. Les types d'anomalies

Les anomalies peuvent prendre différentes formes, valeurs ou tailles, et elles peuvent suivre certains motifs récurrents. Par conséquent, la connaissance *a priori* des anomalies et de leurs caractéristiques est un atout essentiel pour les détecter, et de ce fait, pour sélectionner la technique adaptée (Braei and Wagner, 2020). Dans cette partie, nous introduirons et discuterons les diverses représentations des anomalies issues de mesures réelles, ainsi que leur caractérisation, telles qu'elles sont présentées dans la littérature.

3.3.1. Contexte

De façon analogue à la définition d'une anomalie, il n'existe pas de consensus établi dans la littérature concernant les types d'anomalies rencontrées. Par conséquent, il n'est pas rare de trouver différentes approches visant à les qualifier ou à les répertorier. Certaines approches prennent la forme de typologies générales, conçues pour s'adapter à des données variées, regroupant les anomalies en classes ou familles selon des caractéristiques prédéfinies (Chandola et al., 2009; Foorthuis, 2021). D'autres proposent des typologies moins générales, spécifiques à un type de données unique (Bhuyan et al., 2014; Blázquez-García et al., 2022). Enfin, d'autres encore, généralement pour des applications plus spécifiques, répertorient ou utilisent des listes d'anomalies aux caractéristiques définies, adaptées à une application ou un genre de données particulier (Ni et al., 2009; Schmidl et al., 2022; Sharma et al., 2010). Cette multitude d'approches s'explique majoritairement par l'influence des données et des problématiques, propres au domaine d'application, sur les anomalies et leurs représentations (Foorthuis, 2021).

3.3.2. Les types d'anomalies rencontrées dans les mesures

Il est courant de trouver dans la littérature des recensements variés de types d'anomalies, principalement observés dans des données provenant de mesures réelles et souvent causés par des défaillances des capteurs (Kraiem, 2021; Ni et al., 2009; Sharma et al., 2010). Les causes et les effets de ces défaillances peuvent être variés. Par exemple, Ni et al. (2009) et Bosman (2016) répertorient une série de dysfonctionnement de capteurs pouvant induire des anomalies tels que des erreurs de calibration, des pannes matérielles, des problèmes du système ou de communication, un manque d'alimentation,

des attaques malveillantes, ou une inadéquation entre l'environnement mesuré et le capteur utilisé. Le Tableau 5, inspiré des travaux de Kraiem (2021), offre un aperçu non exhaustif des types d'anomalies régulièrement mentionnées dans la littérature, qui sont aussi représentés dans la Figure 16.

Tableau 5 Types d'anomalies issues de mesures dans la littérature (capteurs ou séries temporelles). Les références suivies d'une * représentent des travaux qui recensent et introduisent des types issus de plusieurs publications. Inspiré de Kraeim (2021).

Anomalies (types)	(Ni et al., 2009)	(Sharma et al., 2010)	(Kaiser and Maravall Herrero, 1999)*	(Kraiem, 2021)*	(Erhan et al., 2021)*	(Kullaa, 2013)
Pics	<i>Outlier</i>	<i>Short</i>	<i>Additive outlier</i>	Pic positif ou négatif	<i>Spike</i>	<i>Gain</i>
Valeurs constantes	<i>Stuck-at fault Rooftop</i>	<i>Constant</i>	<i>Level shift</i>	Constante (plateaux)	<i>Constant</i>	<i>Constant</i>
Bruits	<i>Excessive noise</i>	<i>Noise</i>		Bruit	<i>Noise</i>	<i>Noise</i>
Écarts	<i>Spike</i>	-	<i>Innovative outlier</i>	-	-	<i>Bias</i>
Dérives	<i>Drift</i>	-	-	-	<i>Drift</i>	<i>Drift</i>
Autres	-	-	<i>Transitory changes</i>	Changement de niveau	-	<i>Precision degradation</i> <i>Constant noise</i> <i>Bottom noise</i>

Il convient de souligner que certaines études ne fournissent pas systématiquement des descriptions détaillées des anomalies, rendant leur classification – et de ce fait leur représentation – sujette à interprétation. Cinq principaux types d'anomalies se dégagent :

- **Les pics.** Ce type d'anomalies est caractérisé par sa très courte durée, souvent une seule observation, dont la valeur s'écarte significativement de celles des observations avoisinantes (Foorthuis, 2021; Ni et al., 2009; Noshad et al., 2019). Il s'agit de l'anomalie la plus fréquemment étudiée dans la littérature (Blázquez-García et al., 2022) et est souvent désignée comme une « *anomalie ponctuelle* ». La Figure 16 illustre deux exemples de pics, l'un présentant une variation brusque positive et l'autre négative, mis en évidence par des encadrés marqués d'une étoile (★).
- **Les valeurs constantes.** Ce type d'anomalies peut varier en taille et consiste en général en plusieurs observations dont les valeurs ne changent pas ou varient très peu. La valeur adoptée diffère selon les cas et peut correspondre à celle de l'observation précédente, être aléatoire ou dépendre d'une valeur de référence telle qu'une valeur minimale ou maximale (Kraiem, 2021; Ni et al., 2009; Yao et al., 2010). Ces anomalies

se caractérisent par une variance nulle et, selon les cas, une moyenne sensiblement différente de celle des observations avoisinantes. Elles sont illustrées deux fois dans la Figure 16 par un encadré marqué d'un carré (■).

- **Les bruits.** Ces anomalies varient en taille et comprennent généralement plusieurs observations qui présentent une dispersion plus importante que celle attendue par rapport à l'ensemble des observations voisines (Ni et al., 2009; Sharma et al., 2010). Ces anomalies sont caractérisées par une variance plus élevée que celle prévue. La Figure 16 en illustre un exemple mis en valeur par un encadré marqué d'un triangle (▲).
- **Les écarts.** Ces anomalies varient en taille et comprennent généralement plusieurs observations dont les valeurs changent brusquement d'une quantité fixe, qu'elle soit positive ou négative (Kullaa, 2013; Liu et al., 2023; Noshad et al., 2019). Elles peuvent aussi être désignées sous le nom de « *changements temporaires* » (Foorthuis, 2021). Selon les cas, le retour à la normale peut être progressif ou brutal. Ces anomalies se caractérisent principalement par un changement soudain de la moyenne par rapport aux observations voisines. Un exemple d'écart est présenté dans la Figure 16 et mis en évidence par un encadré marqué d'un rond (●).
- **Les dérives.** Ces anomalies, souvent de très grande taille, sont constituées d'observations dont la distribution évolue dans le temps (Wang et al., 2021). Elles peuvent inclure l'ensemble d'un jeu d'observations et se caractérisent par un changement de tendance permanent, qui n'est pas lié à des facteurs environnementaux, entraînant ainsi une modification de la moyenne des observations (Pereira and Glisic, 2023). Un exemple de dérive avec une tendance volontairement exagérée est illustré dans la Figure 16 et est mis en évidence par un encadré marqué d'un losange (◆). Un cas particulier des dérives pourra également être pris en compte, celui des **changements de niveaux** (Kraiem, 2021). Ils se distinguent des dérives par le fait qu'ils induisent un changement abrupt et non progressif, bien que permanent.

De plus, il convient de souligner que peu de travaux présentent des exemples ou traitent d'anomalies issues de mesures composées de données qualitatives, en particulier binaires. Cependant, Foorthuis (2021) illustre deux exemples caractéristiques : la présence d'éléments inhabituels ou de groupes d'éléments inhabituels. Dans le cas d'une mesure de mouvement utilisée pour évaluer l'occupation d'une pièce, cela reviendrait à observer une indication de mouvement (ou d'absence de mouvement) alors qu'il n'y a personne (ou qu'il y a des occupants présents).

En résumé, les anomalies issues de mesures présentent une grande diversité de profils. Elles varient en matière de **taille**, pouvant être composées d'une à plusieurs observations, voire inclure l'ensemble des données dans le cas des dérives. De plus, elles se différencient

par leurs **dynamiques**, présentant des caractéristiques uniques telles que des changements de variance, de moyenne ou de structure, qui les différencient des observations normales et d'autres anomalies. Cependant, cette partie vise principalement à introduire empiriquement les types d'anomalies issues de mesures et leurs représentations. Une approche plus globale et discriminante sera abordée dans la partie suivante.

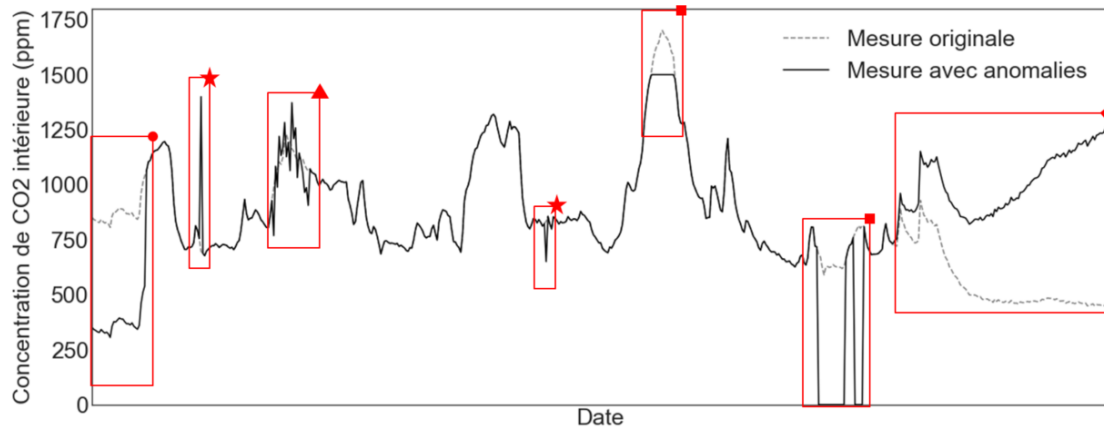


Figure 16 Illustration de types d'anomalies. Pics (★), bruit (▲), valeurs constantes (■), écarts (●) et dérives (◆).

3.3.3. La caractérisation des anomalies au travers les typologies

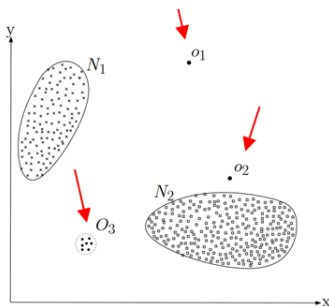
Les typologies sont des outils utilisés pour distinguer des concepts (Bailey, 1994) en se basant sur un ou plusieurs principes classificatoires, aussi appelés dimensions. Elles permettent de fournir une description concise de concepts parfois complexes tout en proposant une vision d'ensemble exhaustive et structurée (Foorthuis, 2021). Ces typologies sont généralement conçues dans le but de classer et comparer plusieurs études entre elles. Dans le contexte de notre application, elles sont surtout répandues dans les travaux d'analyse de la littérature et regroupent plusieurs anomalies en fonction de caractéristiques communes, ce qui en fait une source d'information précieuse pour le choix des techniques de détection d'anomalies.

La littérature propose de nombreuses typologies d'anomalies. Chandola et al. (2009) ont introduit trois catégories d'anomalies, pour des données en deux dimensions, qui ont été largement reprises dans d'autres travaux (Benkabou, 2018; Braei and Wagner, 2020; Chalapathy and Chawla, 2019; Zimek and Filzmoser, 2018). Les représentations de ces catégories, telles qu'introduites par les auteurs, sont illustrées dans la Figure 17 par les flèches rouges. La première, *l'anomalie ponctuelle* (i), fait référence à une observation anormale par rapport à l'ensemble des observations dont elle est issue, par exemple, une mesure d'humidité intérieure relative de 120 %. La deuxième, *l'anomalie contextuelle* (ii),

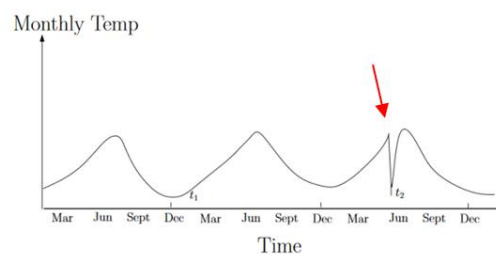
fait référence à une observation anormale – voire une série d’observations anormales selon certaines publications (Liu et al., 2020; Mulongo et al., 2020) – par rapport à un contexte donné, qu’il soit spatial ou temporel. Ainsi, l’observation en question peut être considérée comme normale dans un contexte donnée et anormale dans un autre, tout en appartenant au même ensemble d’observations (p. ex., une température extérieure mesurée à 30 °C en hiver, ce qui est plus typique de l’été). Les anomalies contextuelles sont les plus complexes à détecter puisqu’elles dépendent de l’information disponible pour la contextualisation. Enfin, la dernière catégorie, appelée *anomalie collective* (iii), fait référence à une série anormale d’observations par rapport à l’ensemble complet des observations. Prises individuellement, les observations peuvent être considérées comme normales, mais leur succession les rend anormales (p. ex., la saturation du capteur de CO₂ qui entraîne une série de valeurs mesurées identiques).

En se basant sur ces trois catégories, deux composantes principales ressortent : une anomalie peut (1) se caractériser par sa **taille** et (2) dépendre d’un **contexte** donné, qui peut couvrir toutes ou seulement une partie des observations disponibles. Cependant, il est important de noter qu’il existe un flou entourant les définitions et les représentations des anomalies contextuelles et collectives, ce qui induit une potentielle confusion entre ces deux catégories.

(i) Anomalie ponctuelle



(ii) Anomalie contextuelle



(iii) Anomalie séquentielle

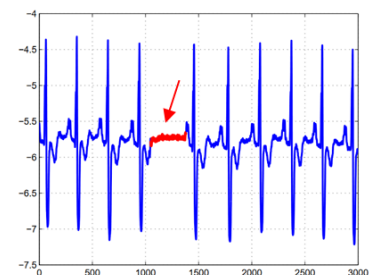


Figure 17 Représentation des types d'anomalies présentées. Issue de Chandola et al. (2009).

Face au même constat, Huang (2018) propose alors quatre catégories d’anomalies distinctes réparties sur deux axes, (1) leur taille et (2) l’apport de contexte. Il introduit, d’une part, les *anomalies ponctuelles* et les *groupes d’anomalies*, qui font respectivement référence à une ou plusieurs observations anormales par rapport à l’ensemble complet des observations dont elles sont issues. La détection s’effectue donc sans qu’il soit nécessaire d’apporter un contexte particulier (p. ex., temporel ou spatial). D’autre part, les *anomalies ponctuelles contextuelles* et les *groupes d’anomalies contextuels* dans le cas où le contexte fait part intégrante du processus de détection.

Dans son article d'analyse de la littérature, Blázquez-García et al. (2022) approfondissent les notions précédemment abordées tout en se concentrant sur les séries temporelles. Une anomalie est présentée comme (1) dépendante de la taille de l'observation anormale. Ainsi, une anomalie peut être une observation anormale unique, une séquence d'observations anormales ou la totalité des observations d'une série temporelle. Ces différentes catégories sont respectivement désignées comme une *anomalie ponctuelle*, une *sous-séquence anormale* et une *série temporelle anormale*. À la taille de l'anomalie considérée s'ajoute (2) la taille des observations nécessaires à sa détection et donc, du contexte apporté. Ainsi, l'anomalie est dite *globale* si elle est considérée anormale par rapport à l'ensemble du corpus d'observations, et *locale* si elle l'est uniquement par rapport aux observations environnantes. Une analogie peut ici être établie avec la catégorisation de Huang (2018) selon l'apport ou non de contexte dans le processus de détection. Cependant, une composante nouvelle est introduite avec la prise en compte de (3) l'influence de l'anomalie sur d'autres variables. Une anomalie à un instant donné qui n'est observable que sur une variable est dite *univariée*, et *multivariée* si elle affecte plusieurs variables (p. ex., une erreur de transmission survenant sur un capteur intérieur mesurant la température, l'humidité relative et la concentration de CO₂ constitue une anomalie multivariée). À des fins d'illustration, la Figure 18, tirée des travaux des auteurs, présente deux anomalies ponctuelles (i) et séquentielles multivariées (ii) représentées par les indices (O1) et (O2), ainsi qu'une anomalie univariée en (O3). Ainsi, Blázquez-García et al. (2022) présentent une typologie d'anomalies plus complète et spécifiques aux séries temporelles.

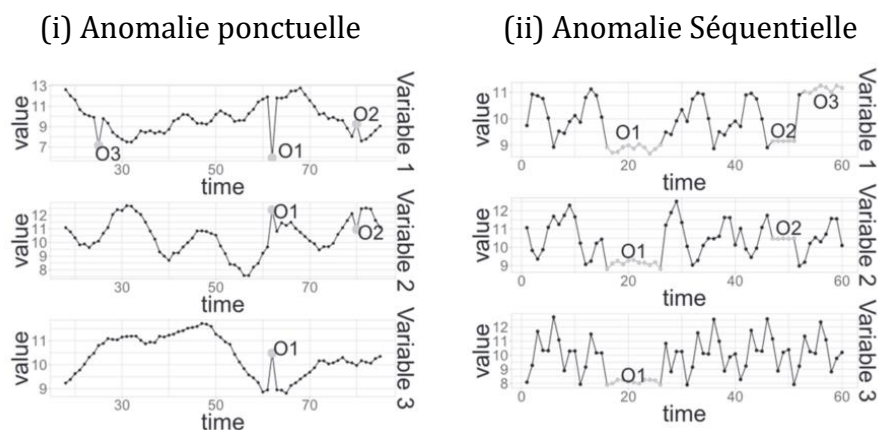


Figure 18 Représentation d'anomalies univariées et multivariées. Tirée de Blázquez-García et al. (2022).

Contrairement aux classifications précédemment abordées, Foorthuis (2021) se concentre davantage sur la nature des anomalies que sur leur détection, bien que ce dernier sujet soit brièvement discuté. L'auteur propose une typologie d'anomalies construite en cinq dimensions et adaptée à plusieurs types de données. Cette classification se distingue par sa complexité, en faisant l'une des plus complètes parmi celles observées

dans la littérature dans le cadre de cette thèse. Cette typologie générique caractérise initialement une anomalie par (1) le type des variables considérées, qu'elles soient *qualitatives*, *quantitatives* ou une *combinaison des deux*. De plus, elle prend en compte (2) la cardinalité des relations, c'est-à-dire le nombre de relations qui déterminent la dépendance et l'influence d'autres variables sur l'anomalie et son processus de détection. Une anomalie est qualifiée d'*univariée* si elle peut être détectée uniquement en utilisant des éléments de la variable à partir de laquelle elle est issue, et de *multivariée* si des informations supplémentaires sous forme d'autres variables sont nécessaires. Ces caractéristiques font écho aux notions de contexte abordées précédemment, mais ici, elles concernent le contexte entre les variables. De plus, la typologie tient également compte de (3) la taille de l'anomalie, qu'elle soit *ponctuelle* ou qu'elle constitue un *agrégat* (ou *groupe*) d'observations. Ces trois éléments permettent de créer une structure principale composée de neuf catégories d'anomalies. À cela s'ajoute deux niveaux de détail supplémentaires, ou sous-types, qui dépendent des types de données considérées, à savoir (4) leur structure et (5) leur distribution.

Ces différentes typologies issues de diverses publications permettent de mettre en lumière les principaux éléments qui semblent caractériser une anomalie dans la littérature. Deux caractéristiques semblent être reprises par l'ensemble des études présentées et apparaissent donc comme pertinentes à considérer dans ces travaux. D'une part, la **taille** de l'anomalie, qui peut être une *observation unique*, une *succession d'observations* ou *l'ensemble des observations*. Cette dernière semble être la caractéristique commune à l'ensemble des typologies considérées mais également aux exemples de types d'anomalies issues de mesures présentés dans la partie 3.3.2. D'autre part, le **contexte nécessaire à la détection**, qui peut prendre différentes formes ou définitions selon les études, mais qu'il est possible de considérer comme la *quantité d'informations complémentaires à apporter* dans le but de détecter une anomalie. Cette information complémentaire peut varier en taille et en provenance, qu'elle soit issue du jeu d'observations initial, d'autres jeux d'observations avec des variables différentes, ou qu'elle soit temporelle. Plus l'apport de contexte – et donc d'information – est important pour détecter une anomalie, plus elle peut être considérée comme complexe à détecter. Dans ce cas, la notion de « complexité » semble suggérer que l'anomalie partage des similitudes avec les observations normales de la mesure en matière de profil, d'apparence, de dynamique ou de valeur.

3.3.4. Mise en place d'une typologie d'anomalies adaptée

Pour englober une grande diversité d'anomalies aux caractéristiques diverses et les évaluer sur différentes données, le développement d'une typologie d'anomalies spécifique au domaine du bâtiment, est proposée. Notre objectif n'est pas de créer une

nouvelle typologie à partir de zéro, mais plutôt d'adapter celles qui existent déjà en conservant les caractéristiques des anomalies jugées pertinentes. La typologie proposée repose sur deux axes, chacun représentant l'une des deux caractéristiques identifiées comme les plus couramment mentionnées:

- (i) **La taille de l'anomalie.** Cet axe concerne le nombre d'observations composant l'anomalie. Elle est dite *ponctuelle* lorsqu'elle concerne une seule observation, *collective* lorsqu'elle englobe plusieurs observations successives et *globale* lorsqu'elle concerne la totalité (ou presque) des observations d'une variable. Cette caractéristique est commune à toutes les typologies étudiées bien qu'incluant aussi, comme suggéré par Blázquez-García et al. (2022), la possibilité de considérer une variable dans son ensemble comme anormale.
- (ii) **La nature de l'anomalie.** Cet axe concerne le contexte et donc la quantité d'informations nécessaire pour détecter l'anomalie. Cette caractéristique reprend en partie celles proposées par Huang (2018) et Foorhuis (2021), en se concentrant exclusivement sur la source de l'information. Ainsi, dans le cas où aucune information complémentaire n'est nécessaire, l'anomalie est dite de nature *indépendante*. Si du contexte est nécessaire, et que l'information requise provient de la même variable que celle dont provient l'anomalie (p. ex., un contexte temporel), elle est considérée de nature *contextuelle univariée*. Dans le cas où l'information complémentaire provient également d'une ou plusieurs autres variables, elle est alors dite de nature *contextuelle multivariée*. Cet axe permet de retranscrire graduellement les défis que présentent certaines anomalies en matière de détection.

De manière plus détaillée, cette typologie permet de caractériser les anomalies en neuf types distincts :

- **Les anomalies indépendantes**, qu'elles soient **ponctuelles**, **collectives** ou **globales**, se traduisent par une observation, une succession d'observations ou la totalité des observations d'une variable donnée, déviant remarquablement des critères prédéfinis. Elles dévient tellement des connaissances préétablies de la variable, ou de l'ensemble des autres observations du jeu de données, que leur présence à elle seule est suffisante pour les rendre détectables sans apport de contexte. La Figure 19 illustre ces trois types d'anomalies avec des exemples de mesure d'humidité relative intérieure. Elles peuvent se traduire par des valeurs aberrantes (par exemple, une d'humidité relative mesurée de 120 % comme en (i), des valeurs seuils (par exemple, des valeurs nulles) comme en (i) et (ii), ou une échelle inadaptée (comme un mauvais calibrage du capteur) comme en (iii). Ces anomalies peuvent correspondre, dans la

littérature (partie 3.3.2), à des *pics* conséquents pour les anomalies ponctuelles et à des *valeurs constantes* nulles ou à des *écarts* importants pour les anomalies collectives.

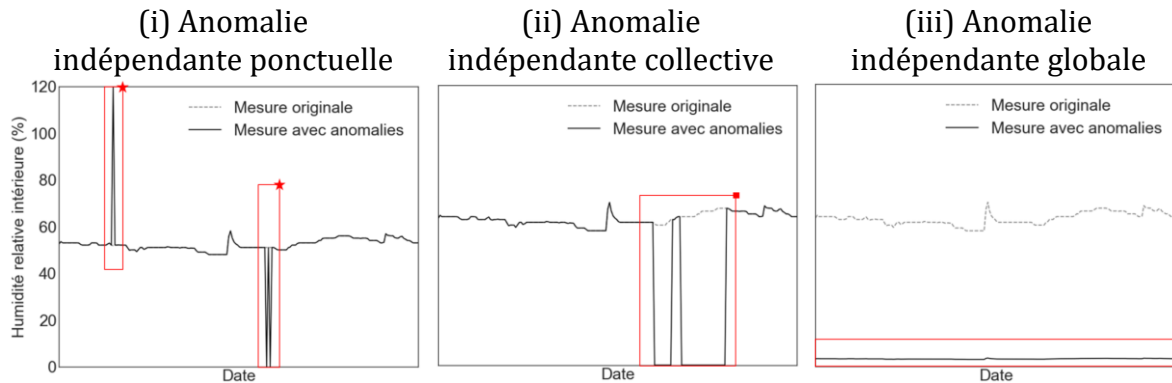


Figure 19 Exemples d'anomalies indépendantes ponctuelles (i.★), collectives (ii.■) et globales (iii.●).

- **Les anomalies univariées**, qu'elles soient **ponctuelles**, **collectives** ou **globales** se traduisent par une, une série ou la totalité des observations d'une variable donnée. Cependant, contrairement aux anomalies indépendantes, elles ne peuvent être détectées que lorsque des informations complémentaires issues de la même variable sont prises en compte pour les mettre en contexte. La Figure 20 illustre des exemples de ces trois anomalies dans le cadre d'une mesure de température intérieure. Ce type d'anomalies peut se traduire par des valeurs inhabituelles au regard des observations avoisinantes, comme présenté en (i) avec plusieurs pics liés à une faible variation positive ou négative, ou (ii) avec une séries de valeurs constantes. Elles peuvent aussi se manifester par des tendances inhabituelles par rapport à l'ensemble des observations précédentes, comme une dérive (iii). Ces anomalies univariées ponctuelles, collectives et globales peuvent respectivement correspondre à des *pics*, des *valeurs constantes*, des *bruits*, des *écarts*, des *dérives* ou des *changements de niveau*, comme présenté dans la partie 3.3.2.

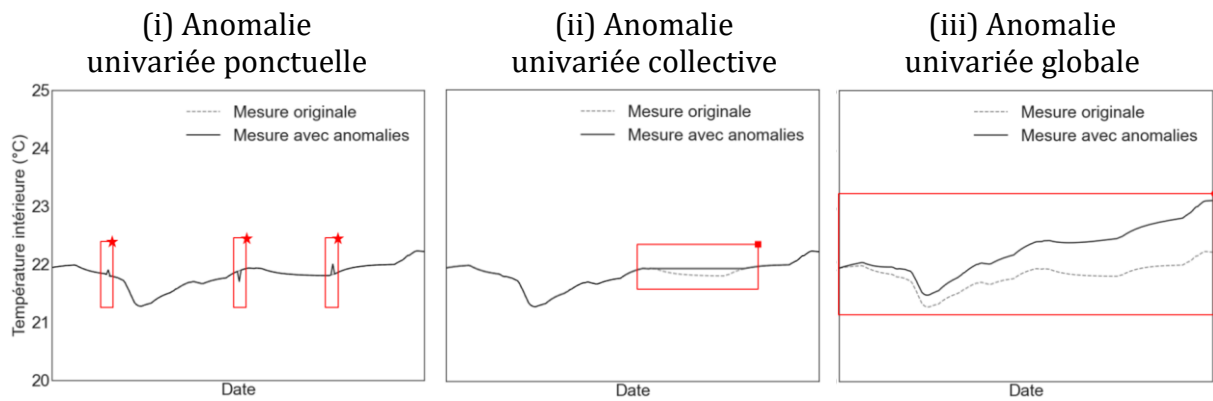


Figure 20 Exemples d'anomalies univariées ponctuelles (i.★), collectives (ii.■) et globales (iii.●).

- Les anomalies multivariées**, qu'elles soient **ponctuelles**, **collectives** ou **globales**, se traduisent par une, une série, ou la totalité des observations d'une variable donnée. Elles se démarquent en déviant notablement de ce qui est attendu, grâce à des informations provenant d'une ou plusieurs autres variables. Leur détection requiert cet apport d'informations supplémentaires ou mise en contexte. La Figure 21 illustre ces trois anomalies avec des exemples utilisant une mesure binaire d'occupation intérieure comme variable principale et la concentration de CO₂ comme variable complémentaire. Ce type d'anomalies peut se traduire par des valeurs inhabituelles au regard des observations obtenues *via* d'autres sources de données, comme présenté en (i) avec plusieurs valeurs ponctuellement nulles d'occupation alors que la concentration de CO₂ est croissante et élevée, ou en (ii) avec une absence d'occupation sur toute une période malgré un profil de concentration CO₂ contradictoire. De plus, cela peut aussi se manifester par l'absence d'une quantité importante de données ou par un renversement du profil d'occupation, généralement lié à une mauvaise paramétrisation ou à une lecture incorrecte du capteur (iii). Par rapport aux types d'anomalies observées dans la littérature et présentées dans la partie 3.3.2, les anomalies multivariées ponctuelles peuvent correspondre à des *pics*, les anomalies multivariées collectives à des *valeurs constantes*, des *bruits* ou des *écarts* peu prononcés, tandis que et les anomalies multivariées globales sont principalement associées à des *dérives* ou à des *changements de niveau* plus faibles. Elles se différencient essentiellement des anomalies univariées par la provenance de l'information nécessaire à la détection.

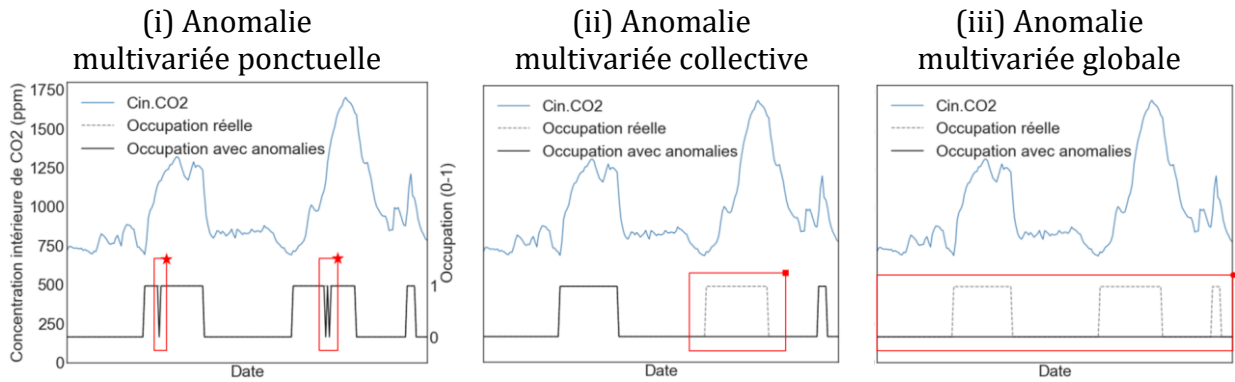


Figure 21 Exemples d'anomalies multivariées ponctuelles (i.★), collectives (ii.■) et globales (iii.●).

L'ensemble de la typologie proposée est résumée dans le Tableau 6. Il est également important de noter que, à ce stade de l'étude, la caractéristique de **dynamique** des anomalies recensée et observée dans la partie 3.3.2 n'a pas été prise en compte. Son effet sur les données, et donc sur la représentation de l'anomalie est supposé, *a priori*, avoir un impact significatif sur les performances de la tâche de détection et, par conséquent influencer le choix des techniques. Cependant, à ce stade, cet impact apparaît difficilement quantifiable. C'est pourquoi, seules les caractéristiques dont l'impact a été généralement recensé dans la littérature pour construire la typologie d'anomalies furent sélectionnées. Néanmoins, dans le chapitre 5, nous prendrons en compte la dynamique lors de l'évaluation des différentes anomalies. Si possible, nous approcherons ou évaluerons son impact, ce qui pourrait conduire à des modifications de la typologie.

Tableau 6 Typologie d'anomalie adaptée aux anomalies et mesures issues du bâtiment

Nature	Contextuelle	Taille		
		Ponctuelle	Collective	Globale
	Indépendante	Anomalie indépendante ponctuelle	Anomalie indépendante collective	Anomalie indépendante globale
		Sa présence est suffisante pour la détecter		
	Univariée	Anomalie univariée ponctuelle	Anomalie univariée collective	Anomalie univariée globale
		Une ou plusieurs informations complémentaires provenant de la même variable sont nécessaires pour la détecter		
	Multivariée	Anomalie multivariée ponctuelle	Anomalie multivariée collective	Anomalie multivariée globale
		Une ou plusieurs informations complémentaires provenant d'une ou plusieurs autres variables sont nécessaires pour la détecter		

3.4. Méthodologie de sélection des techniques de détection

3.4.1. La taxonomie des problématiques rencontrées dans les mesures du bâtiment

En se basant sur les caractéristiques des données (cf. partie 3.2.3) et des anomalies (cf. partie 3.3.4), il est possible de construire une taxonomie pour identifier diverses problématiques, induites par ces caractéristiques, à résoudre et pouvant s'adapter à plusieurs mesures issues du bâtiment. Cette taxonomie est composée de trois axes : deux fixes et un modulable. La Figure 22 présente un aperçu schématique de cette taxonomie. Les deux axes fixes reprennent la **taille de l'anomalie** (i) positionnée en abscisses (x) et la **nature de l'anomalie** (ii) sur les ordonnées (y). L'axe modulable (ou variable) représente **l'ensemble des mesures** (iii) pouvant être considérées, ce qui dépend des différentes natures de données qu'il est possible de rencontrer. Cet axe est positionné sur la cote (z). La taxonomie peut donc être vue comme une grille en trois dimensions, où les axes des abscisses et des ordonnées représentent respectivement les trois éléments constitutifs de la *taille* (ponctuelle, collective ou globale) et de la *nature* (indépendante, contextuelle univariée ou contextuelle multivariée) des anomalies, et dont la cote (z) contient M mesures. Ainsi, cette taxonomie est de dimensions (3, 3, M).

(i) Taille des anomalies (x) (ii) Nature des anomalies (y) (iii) Mesures considérées (z)

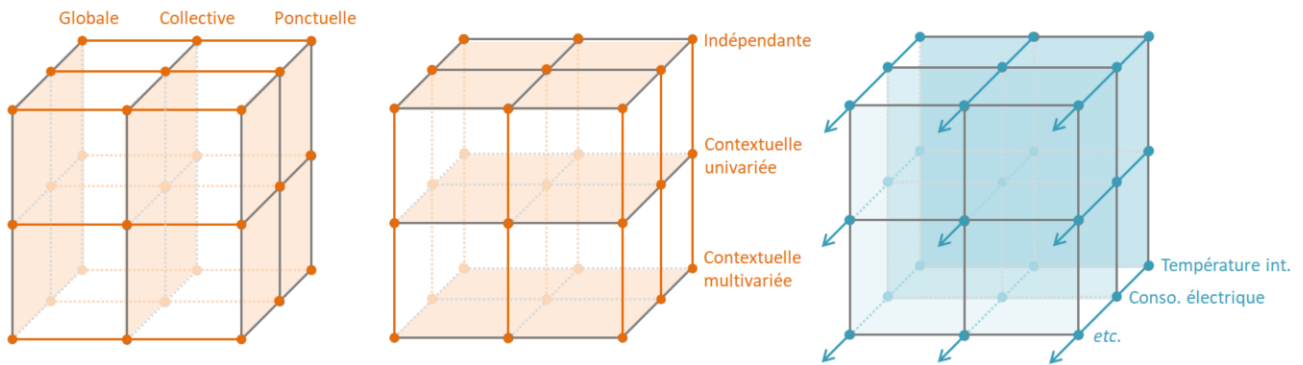


Figure 22 Schématisation de la taxonomie des problématiques en trois dimensions proposée. L'exemple proposé est de dimensions (3, 3, 3) avec deux axes fixes de trois éléments chacun correspondant à la taille (i) et la nature (ii) des anomalies respectivement sur les abscisses (x) et ordonnées (y). Le troisième axe est de dimension variable mais est composé de trois éléments dans cet exemple. Il correspond aux mesures considérées et est positionné sur la cote (z). Les aires représentent visuellement les zones où les caractéristiques d'un axe x , y , ou z sont constantes.

En utilisant la taxonomie, il est possible de représenter **un problème donné** (ou une problématique donnée) par un point, un vecteur, un plan ou même une zone définis dans l'espace. Chaque représentation est alors composée d'un ensemble de coordonnées, chacune dépendant des caractéristiques de chaque axe. La Figure 23 illustre quelques exemples sur une taxonomie appliquée aux mesures de températures intérieures, de consommations électriques et de concentrations de CO_2 intérieures (notées respectivement T_{int} , $C_{\text{électrique}}$ et $\text{CO}_{2,\text{int}}$). Dans ce contexte, la taxonomie des problématiques prend la forme d'une grille de dimensions (3, 3, 3).

Par exemple, en (i), est présenté un problème de détection d'anomalies collectives d'une nature contextuelle univariée sur les mesures de consommations électriques. Le cas (ii) est représenté sous la forme d'un vecteur qui décrit le problème sous la forme d'une détection d'anomalies globales de nature contextuelle univariée sur des données de concentration de CO_2 intérieure et de consommation électrique. En (iii), le plan englobe l'ensemble des anomalies ponctuelles, indépendamment de leur nature, pour toutes les mesures.

Cette taxonomie, bien que relativement simple avec ses trois axes, offre une approche structurée, exhaustive et modulable pour différencier les différents problèmes pertinents ou cas d'applications sur lesquels évaluer les techniques de détection d'anomalies. Sa souplesse permet d'ajouter de nouvelles dimensions si nécessaire (p. ex., la dynamique d'une anomalie), d'en modifier la structure (comme considérer uniquement les données en fonction du type de la variable plutôt que de leur nature), ou d'en retirer des dimensions (p. ex., si aucune différence liée à la taille des anomalies n'est recensée). Cette taxonomie servira de base pour évaluer plusieurs méthodes de détection d'anomalies sur

différentes problématiques, pour *in fine* sélectionner celles qui conviennent le mieux à une **mesure**, une **taille** ou une **nature** d'anomalie donnée.

Cependant, dans des conditions réelles d'application, l'identification de la nature d'une anomalie, et en particulier de l'apport de contexte nécessaire à sa détection, constitue une problématique à part entière, qui sera discutée dans la partie suivante.

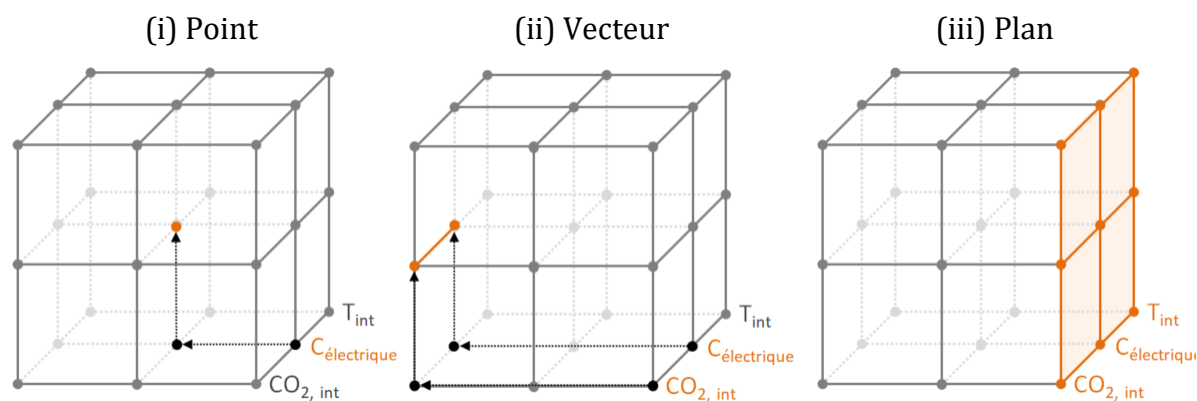


Figure 23 Exemple de la représentation des problématiques au travers une taxonomie de taille (3, 3, 3). Y sont représentés un problème sous forme d'un point dans l'espace de la taxonomie (i), un problème sous forme de vecteur (ii) et un problème sous forme d'un plan (iii).

3.4.2. Problématique du contexte dans les cas d'applications réels

3.4.2.1. Illustration par l'exemple

Pour illustrer les défis liés à l'apport de contexte, prenons l'exemple de la détection des ouvertures de fenêtres à l'aide de mesures d'ambiance. Chaque ouverture de fenêtre se traduit sur ces mesures par une série d'observations à la dynamique bien différente des autres. Bien que des détails plus précis seront apportés dans le chapitre 4, il est possible de considérer cet exemple comme un cas de classification binaire : la fenêtre est ouverte (anormal) ou fermée (normal). Les résultats présentés dans la Figure 24 ont été obtenus en entraînant un réseaux de neurones à détecter et à classer de façon supervisée les ouvertures de fenêtres sur une période de chauffe (durant l'hiver 2019-2020) pour une chambre d'appartement. Le modèle a ensuite été testé sur les données issues d'un mois d'hiver une année plus tard (2020-2021). Les mesures de température intérieure servent de repère visuel pour rendre compte du comportement observé, tandis que les ouvertures réellement mesurées pendant cette période constitue la vérité terrain et sont représentées en bleu. Trois entraînements distincts ont été réalisés pour ce cas. Le premier, en orange (i), s'est basé uniquement sur les mesures de température intérieure comme unique information. Le deuxième, en rouge (ii), a inclus les mesures en concentration de CO_2 intérieure de la pièce en plus de la température intérieure. Enfin, le dernier entraînement, en violet (iii), a combiné six variables recensées dans la littérature

comme ayant un impact sur la détection d'ouvertures de fenêtres, telles que la température intérieure et extérieure, la concentration de CO₂, l'humidité relative ambiante (Dai et al., 2020). Les traits représentent les ouvertures de fenêtres, et leur épaisseur indique leur durée.

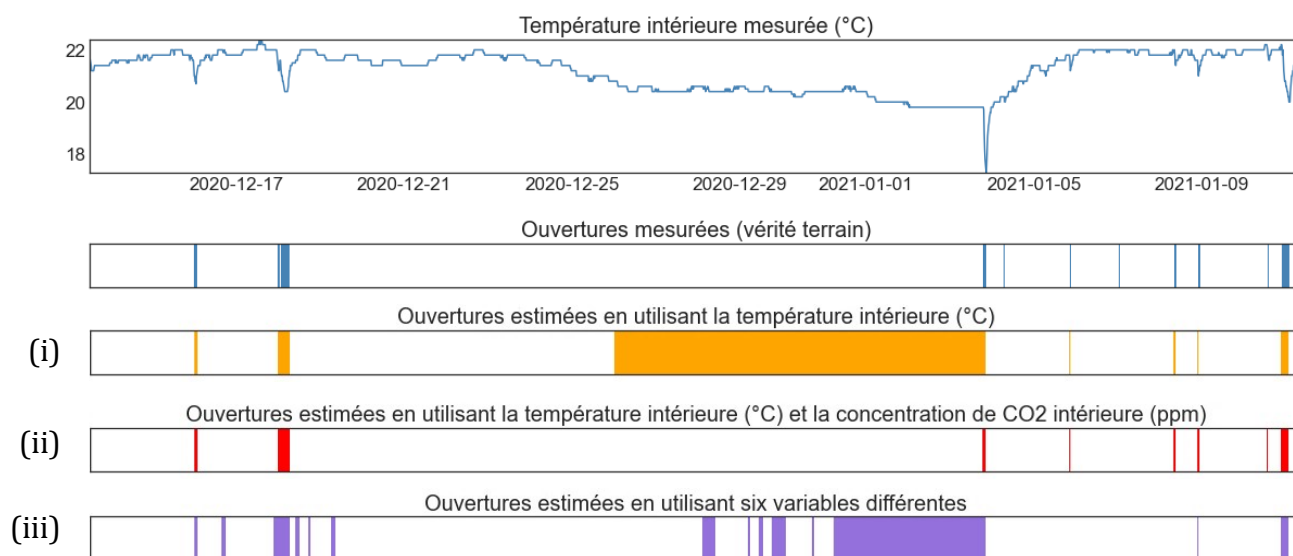


Figure 24 Exemple de détections d'ouvertures de fenêtres selon différentes combinaisons de variable.

Il est évident dès le départ que le premier modèle (i) offre les performances les moins satisfaisantes. Cela est principalement dû à une estimation éronnée des ouvertures de fenêtres sur plusieurs jours, qui ne correspond pas à la réalité (au centre de la figure). Cette erreur peut s'expliquer par la dynamique particulière de la température à ce moment-là, avec une baisse moyenne de la température de 2 °C sur plusieurs jours en raison de l'absence des occupants et donc de chauffage pendant cette période. Étant donné que cette dynamique particulière n'apparaît jamais dans les données utilisées pour l'entraînement, il est probable que le modèle ait associé les baisses significatives de température à des périodes d'ouvertures. Cette hypothèse est écartée grâce à l'apport de nouvelles informations suite à l'intégration de la variable de concentration de CO₂ (ii). Il semble que ces informations supplémentaires aient été utilisées par le modèle pour établir d'autres corrélations, qui, à la lumière des résultats obtenus, semblent mieux refléter le processus sous-jacent qu'il était nécessaire d'identifier. En revanche, le dernier modèle, qui a été entraîné avec un nombre plus important de variables, pourtant considérées comme pertinentes, ne semble pas offrir de meilleurs résultats, bien au contraire. On peut supposer que le modèle n'a pas été en mesure d'extraire les informations les plus pertinentes parmi toutes celles à sa disposition.

3.4.2.2. Le contexte, une caractéristique devant être modulable

Cet exemple illustre plusieurs points importants qui seront discutés de manière plus formelle. Premièrement, le processus de sélection des variables est primordial et doit s'adapter à la tâche ou à l'objectif visé. Dans le but de résoudre un problème donné, il est courant d'utiliser toutes les données jugées pertinentes à disposition en raison de leur potentiel informatif (Emmott et al., 2016). Cependant, cette approche peut augmenter significativement la quantité d'informations à traiter, ainsi que le risque d'y inclure des données parasites qui complexifient et faussent le processus d'apprentissage ou de détection (Su et al., 2019). Par conséquent, la sélection et la représentation appropriées des variables, toutes deux pouvant être corrélées à l'expertise métier, peuvent avoir un impact considérable sur les performances de la plupart des modèles (Bengio et al., 2014). La meilleure approche consiste donc à trouver un équilibre entre l'apport d'informations pertinentes et d'informations non pertinentes. Ainsi l'impact du **choix des variables**, comme celui de leur **représentation**, sur les performances des méthodes de détection d'anomalies, seront pris en compte dans ces travaux.

Cependant, il est important de noter que, même avec une bonne connaissance des variables, de leurs relations et de leurs caractéristiques, il existe des limitations propres aux cadres définis ou proposés par une étude. En effet, le processus de sélection des variables dépend essentiellement de la disponibilité et la qualité des mesures. Si la concentration de CO₂ n'avait pas été disponible, des résultats tels que ceux en (ii) n'auraient peut être pas pu être obtenus. C'est une contrainte importante à prendre en compte lorsqu'il s'agit d'un cadre d'application réaliste. Il n'est pas surprenant de supposer qu'une méthode de détection d'ouvertures de fenêtres, et par extension de détection d'anomalies, basée sur deux variables, pourrait être préférable à une méthode basée sur six variables, même si elle est meilleure.

Par conséquent, nous proposons de caractériser une anomalie en fonction d'un contexte modulable. Pour cela, le **contexte nécessaire à la détection**, qui détermine la nature de l'anomalie, peut être considéré comme la *quantité minimale d'informations complémentaires nécessaires* pour détecter l'anomalie considérée.

3.5. Conclusion

En conclusion de ce chapitre, il convient de souligner que la démarche proposée vise à comparer et à classer diverses techniques de détection d'anomalies sur des mesures et des anomalies issues du domaine du bâtiment. Cette démarche, sous forme de taxonomie, repose sur la combinaison de caractéristiques distinctives des données et des anomalies du domaine du bâtiment, caractérisées comme ayant un impact sur la tâche de détection

d'anomalies. Cependant, il est important de noter que cette approche, qui croise la nature des données et les types d'anomalies, reste une approche « *classique* » pour la classification et la sélection des techniques de détection d'anomalies (Blázquez-García et al., 2022; Chalapathy and Chawla, 2019; Chandola et al., 2009), et elle a été adaptée d'approches similaires observées dans la littérature pour correspondre aux spécificités du contexte de cette thèse.

Cependant, il est à noter que le développement de cette taxonomie a d'abord permis d'explorer les caractéristiques des données et des anomalies, ainsi que les différents obstacles qui compliquent la détection d'anomalies en raison de la nature des données et de la disponibilité des informations contextuelles. De plus, il a été proposé une méthodologie permettant d'identifier des problèmes pertinents sur lesquels évaluer diverses méthodes de détection d'anomalies.

La suite de ces travaux consistera donc à évaluer un large éventail de techniques de détection d'anomalies sur ces problèmes pertinents. À ce titre, le chapitre 4 présentera une preuve de concept sous forme d'une application de techniques de détection supervisée sur des anomalies collectives, qu'elles soient contextuelles univariée ou multivariées.

Chapitre 4

Détection supervisée d'anomalies : application aux ouvertures de fenêtres

Ce chapitre vise à appliquer la méthode de comparaison et de sélection de techniques de détection d'anomalies développée dans le chapitre précédent. Il s'agit d'une preuve de concept utilisant différentes techniques d'apprentissage supervisé pour détecter un type spécifique d'anomalies étiquetées, en l'occurrence la détection d'ouvertures de fenêtres, laquelle correspond à un problème de classification binaire similaire à la détection d'anomalies. Un état de l'art sur la détection d'ouvertures de fenêtres est réalisé dans un premier temps pour comprendre les enjeux spécifiques de cette thématique (section 4.2). Puis les modèles et le processus d'évaluation employés sont décrits dans la section 4.3. Les données employées pour ce cas d'étude ainsi que le protocole d'expérimentation suivi sont abordés dans la section 4.4, laquelle sera suivie de la présentation des résultats dans la section 4.5. Ce chapitre se terminera par une synthèse dans la section 4.6.

« En fait, tout était une affaire de perspectives. Quand on cherchait l'aventure, on ne voyait pas les choses de la même façon que lorsqu'on la subissait. »

Extrait traduit de, La Roue du Temps : l'œil du monde, de Robert Jordan

Sommaire du chapitre 4

4.1. Contexte et démarche	83
4.2. État de l'art : la détection d'ouvertures de fenêtres.....	85
4.2.1. Détection ou prédiction du comportement des occupants.....	85
4.2.2. Modélisation des ouvertures de fenêtres	86
4.2.3. Apport de l'expertise métier pour l'optimisation des performances	88
4.3. Description du protocole expérimental.....	89
4.3.1. Processus d'évaluation des performances	89
4.3.2. Présentation des modèles.....	92
4.3.3. Configuration et entraînements des modèles.....	97
4.4. Description et prétraitement des données	100
4.4.1. Collecte et préparation des données.....	100
4.4.2. Analyse des données de test	102
4.4.3. Prétraitement des données par transformation et association	103
4.4.4. Combinaisons et association des données	107
4.5. Résultats	107
4.5.1. Évaluation basée sur le F_1 -score.....	108
4.5.2. Évaluation avec métriques additionnelles	114
4.5.3. Synthèse appliquée à la détection d'ouvertures des fenêtres	118
4.6. Conclusion.....	119

4.1. Contexte et démarche

Comme précédemment mentionné, les techniques de détection d'anomalies sont affectées par plusieurs caractéristiques des données et des anomalies. En simplifiant le problème de détection en fixant certaines de ces caractéristiques, c'est-à-dire en ne considérant pas l'ensemble des dimensions de la grille de comparaison des techniques, il devient possible d'évaluer l'impact des obstacles identifiés lors du précédent chapitre et de juger de la pertinence de la méthodologie proposée sur les axes restants. Dans cette optique, ce chapitre présente une approche basée sur la détection des ouvertures de fenêtres dans les mesures du bâtiment.

L'ouverture d'une fenêtre est une action susceptible d'entraîner des modifications significatives des conditions intérieures. Elle impacte l'équilibre thermique, hygrométrique et la qualité de l'air d'une pièce, ce qui engendre des observations déviant sensiblement des autres observations, rappelant ainsi le comportement d'une anomalie (Hawkins, 1980). Une ouverture de fenêtre se distingue au sein d'une variable par des motifs spécifiques qui s'écartent de la tendance générale, tout comme une anomalie. Elles correspondent donc à des anomalies dites « d'intérêt » comme définies dans la section 2.2. Des techniques d'apprentissage automatique peuvent être utilisées pour détecter et classer ces motifs, permettant ainsi de déterminer si une observation est normale ou anormale, et par conséquent, si une fenêtre est fermée ou ouverte. Quelques exemples sont présentés dans la Figure 25, avec à gauche (i) le profil de température, d'humidité et de concentration de CO₂ d'une pièce sans ouverture pendant deux jours, et à droite (ii) la même pièce avec plusieurs ouvertures d'une durée similaire. Dans ce cas précis, les ouvertures de fenêtres surviennent en novembre et se caractérisent par des baisses conséquentes et récurrentes de la température intérieure, de la concentration de CO₂ ainsi que des variations moins marquées mais notables de l'humidité relative intérieure.

Selon le degré d'instrumentation des habitations, il peut arriver que les états d'ouvertures et de fermetures des fenêtres soient monitorés (Dong et al., 2019) et donc qu'on dispose de données étiquetées. Le cas échéant, la détection des ouvertures de fenêtres peut être abordée comme un problème de détection d'anomalies supervisé, présentant plusieurs atouts :

- Premièrement, cette approche permet d'évaluer des techniques supervisées sur des « anomalies » produites par un processus réel distinct de celui générant les variables dont l'étiquetage est accessible. Cette approche, préconisée par Emmott et al. (2016), bien que complexe à mettre en pratique dans des conditions réelles comme le soulignent Chandola et al. (2009), se concentre sur des anomalies de taille collective

présentant une dynamique spécifique et récurrente. Elle permet donc de fixer ces deux caractéristiques et de recentrer la tâche de détection sur une problématique unique.

- Deuxièmement, elle apporte une alternative à la détection d'anomalies sur des données qualitatives binaires en les modélisant au moyen d'autres mesures. Le but final est de comparer aux mesures réelles pour détecter d'éventuelles anomalies, telles qu'une ouverture de fenêtre non répertoriée à cause d'un défaut ou de l'arrêt des capteurs. Un tel cas d'application revient à effectuer une tâche de détection sur des anomalies de nature contextuelle multivariée.
- Troisièmement, cette approche permet de modéliser l'un des aspects du comportement des occupants, qui a un fort impact sur le confort et la consommation énergétique des bâtiments. Cependant, ces comportements ne sont généralement pas mesurés sur site ou ne peuvent pas être exploités directement. Ils sont plutôt approchés à l'aide de techniques expertes ou stochastiques (Markovic, 2018), (Markovic, 2018), ce qui, selon les applications, peut conduire à des écarts conséquents entre les estimations et la réalité (Bonte et al., 2014; Recht et al., 2018).

Dans un premier temps nous réalisons un état de l'art centré sur la détection d'ouvertures de fenêtres, ce qui nous permettra d'aborder les problématiques spécifiques à ce domaine. En fonction des leçons tirées de cette revue de la littérature, nous adapterons notre approche quant au choix des modèles à évaluer, des mesures à considérer, et des méthodes à employer. Par la suite, nous évaluerons les capacités de détection de plusieurs modèles sur différentes mesures, en tenant compte du contexte variable selon les données disponibles. Ces évaluations seront effectuées sur des données réelles issues de mesures.

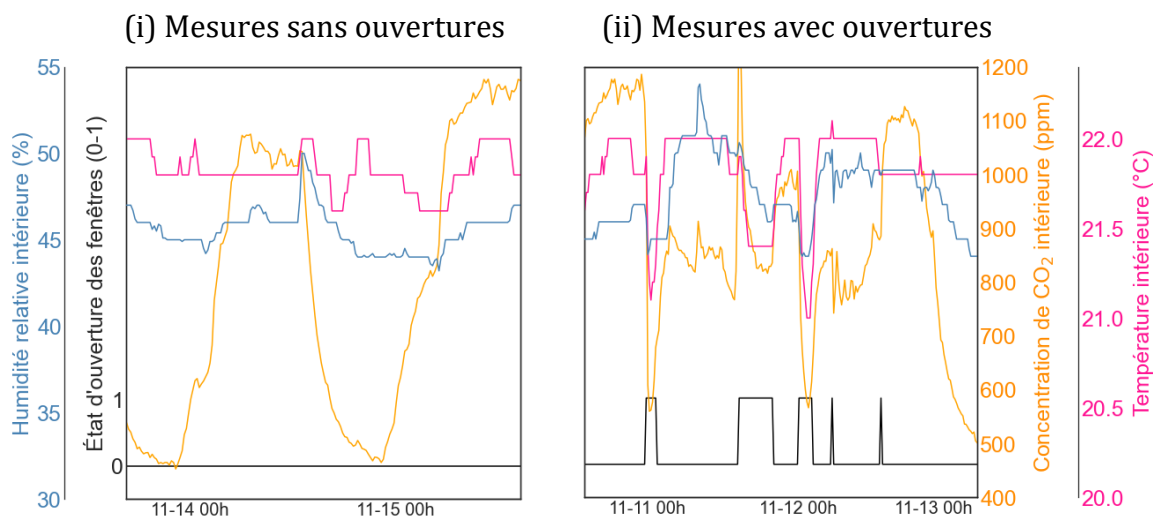


Figure 25 Exemples de motifs récurrents anormaux liés aux ouvertures de fenêtres. À gauche (i) : deux jours sans ouverture dans une pièce avec des mesures de température, d'humidité relative et de concentration de CO₂ intérieure. À droite (ii) : la même pièce avec plusieurs ouvertures sur une période de deux jours de mesure.

4.2. État de l'art : la détection d'ouvertures de fenêtres

4.2.1. Détection ou prédiction du comportement des occupants

Le comportement des occupants a des conséquences tant sur leur confort thermique que sur la consommation énergétique des bâtiments (Lopes et al., 2017). Environ 25 % de la consommation totale d'énergie en Europe entre 1990 et 2014 était consacrée aux usages domestiques (*European Environment Agency*, 2015). Pour comprendre, contrôler ou minimiser ces impacts, de nombreuses techniques d'apprentissage automatique peuvent être utilisées pour modéliser ces comportements, tels que l'occupation et l'ouverture des fenêtres. Ces techniques peuvent être appliquées dans le domaine du bâtiment pour accomplir diverses tâches, notamment la détection d'anomalies, l'analyse et la visualisation de données, la mise en place de systèmes autonomes, le contrôle à distance, la gestion des ressources, *etc.* (Dai et al., 2020; Fan et al., 2015; Karimi et al., 2021; Trappey et al., 2017).

Dans cette optique, la Figure 26 présente trois applications différentes des modèles de comportement des occupants : les modèles de **détection *a posteriori*** (passés), en **temps réel** (présents) et **prédictifs** (futurs). Les modèles de détection *a posteriori* (i) peuvent utiliser toutes les données mesurées pour détecter ou classer *a posteriori* les comportements des occupants. Cette approche peut être employée, entre autres, pour la détection de défauts, la détection et la correction d'anomalies en comparant les comportements modélisés aux comportements mesurés, la réduction de l'écart entre la consommation d'énergie simulée et la consommation d'énergie réelle pour la vérification de la performance énergétique (Gallagher et al., 2018; Recht et al., 2018), ou encore pour évaluer des actions de rénovation (Grillone et al., 2020; Hong et al., 2020). Les modèles de détection en temps réel (ii) se focalisent sur la détection en temps réel des comportements au moment où ils se produisent, tandis que les modèles de prédiction (iii) visent à prévoir les comportements avec un ou plusieurs pas de temps à l'avance. Ces deux approches sont les plus couramment utilisées dans les études (Dai et al., 2020) et peuvent être appliquées en temps réel pour détecter des défauts (Mahmoud et al., 2016), optimiser le contrôle des systèmes pour améliorer le confort ou réaliser des économies d'énergie (Anastasiadou et al., 2021; Ardakanian et al., 2016; Djenouri et al., 2020). Il convient de souligner que le principal intérêt de ces deux approches réside dans leur capacité à fournir des résultats peu de temps après l'acquisition des données. Cependant, cet avantage pose de nombreux défis techniques, tels que le stockage, la gestion et le traitement des ressources (p. ex., modèles, données) en temps réel, sur site ou en ligne (Ayadi et al., 2017; Erhan et al., 2021). Leur implémentation constitue donc une problématique distincte nécessitant des moyens adaptés.

Outre les spécificités propres à chaque technique utilisée et à leur processus d'entraînement respectif, ces trois applications peuvent être mises en place en utilisant les mêmes données et dépendent principalement des utilisations prévues ou souhaitées. Par conséquent, afin d'éviter toute surcharge d'informations et compte tenu de l'orientation de cette thèse, cette étude se concentre uniquement sur la **détection *a posteriori***, dont l'approche est similaire au processus de détection d'anomalies. Toutefois, il convient de noter que l'ensemble de la démarche développée peut également être appliquée, de manière intégrale, à la détection en temps réel et à la prédiction.

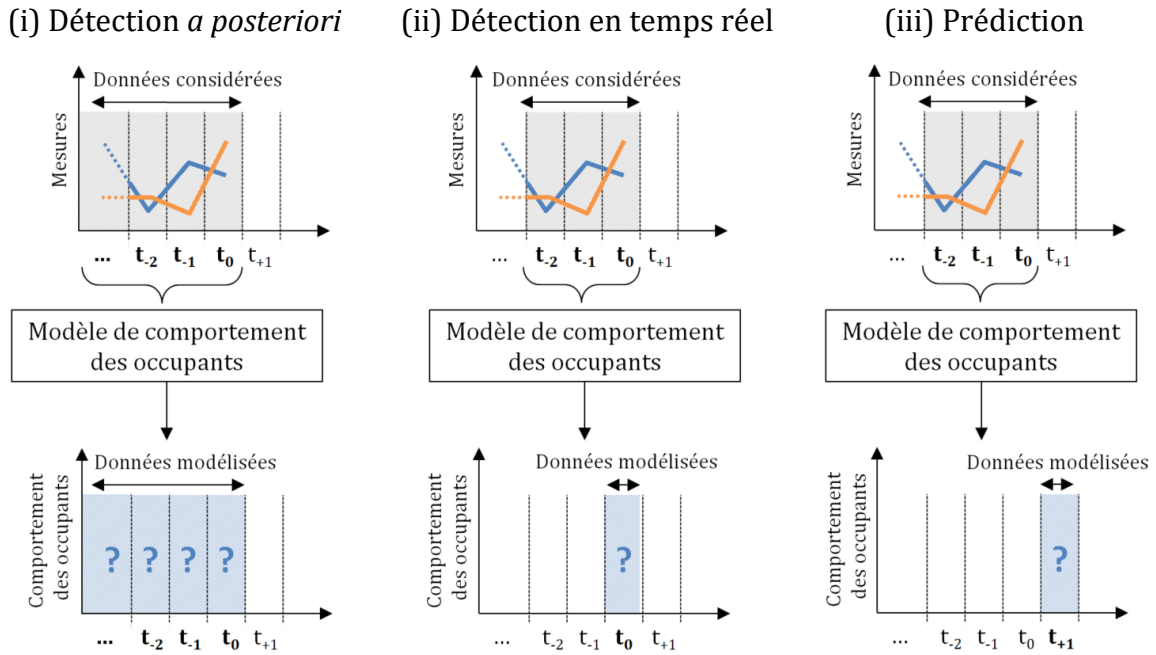


Figure 26 Détection du comportement des occupants selon les données utilisées et les objectifs visés : (i) *a posteriori*, (ii) en temps réel et (ii) par prédiction.

4.2.2. Modélisation des ouvertures de fenêtres

L'analyse approfondie de la littérature sur la modélisation des ouvertures de fenêtres au moyen de techniques d'apprentissage automatique révèle deux méthodologies distinctes identifiées par Dai et al. (2020). La première se concentre sur les actions des occupants vis-à-vis des fenêtres, en tenant compte de l'environnement spécifique (p. ex., la température intérieure et extérieure, la vitesse du vent) pour modéliser les actions d'ouverture et de fermeture (H. Kim et al., 2019; Naspi et al., 2018). Cela revient à définir la probabilité qu'un occupant ouvre (ou ferme) une fenêtre en fonction de son environnement. La seconde méthodologie, sur laquelle repose cette étude, se concentre plutôt sur la modélisation de l'état de la fenêtre en tant qu'*ouverte* ou *fermée*, en fonction des caractéristiques de l'environnement observées à un moment donné (Markovic, 2018; Markovic et al., 2017; Stazi et al., 2017). Par exemple, il s'agit de déterminer si la fenêtre

est ouverte ou fermée en se basant sur la température intérieure mesurée. Une analyse de la littérature met en lumière trois tendances clés (Ahmed et al., 2023; Dai et al., 2020) :

- (i) La plupart des travaux se concentrent sur un seul modèle d'apprentissage automatique à la fois et comparent rarement différents modèles dans la même étude.
- (ii) Les modèles de Régression Logistique (RL) (Cramer, 2003) sont les plus couramment utilisés pour déterminer le statut des fenêtres (Haldi and Robinson, 2009; Rijal et al., 2018; Shi et al., 2018; Stazi et al., 2017), suivis, dans les études plus récentes, par les Réseaux de Neurones (RN) (Markovic, 2018; Markovic et al., 2019; Wei et al., 2019). La grande majorité des modèles présentés repose sur une approche supervisée, impliquant l'utilisation de données étiquetées.
- (iii) En dehors des modèles génériques spécifiquement entraînés à cette fin, à l'aide d'une quantité extrêmement variée de données (Markovic, 2018), les modèles d'apprentissage automatique qui modélisent l'état des fenêtres sont souvent spécifiques aux bâtiments, aux occupants ou aux saisons (Markovic, 2018; Shi et al., 2018). De plus, leurs performances peuvent être évaluées selon différentes métriques telles que la précision, le F1-score, le taux de vrais positifs, l'aire sous la courbe, *etc.* Ainsi, l'évaluation et la comparaison de différents modèles dans le cadre d'études multiples s'avèrent complexes.

Ces tendances illustrent des problématiques similaires à celles observées dans cette thèse en ce qui concerne la modélisation des ouvertures de fenêtres. Par conséquent, il est essentiel d'explorer différents modèles afin de comparer leurs performances. Outre les modèles de régression logistique et les réseaux neuronaux, dont l'emploi semble significatif dans ce domaine, d'autres modèles moins couramment utilisés dans la détection des ouvertures de fenêtres sont examinés. Le choix de ces modèles est basé sur leur popularité dans d'autres domaines d'application ainsi que leur capacité à accomplir diverses autres tâches telles que la reconnaissance de motifs récurrents, la classification, la prédiction ou la détection d'anomalies.

Les modèles retenus sont les suivants : les **Réseaux de Neurones Récurrents** (RNR) (Chalapathy and Chawla, 2019; Huchuk et al., 2019; Liu et al., 2020), l'**Analyse Discriminante Linéaire** (ADL) (Candanedo and Feldheim, 2016; Chen et al., 2016), les **Séparateurs à Vaste Marge** (SVM) (Chen et al., 2016; Liu et al., 2020) et la **forêt aléatoire** (Candanedo and Feldheim, 2016; Chen et al., 2016; Huchuk et al., 2019). Dans la littérature, ces modèles sont souvent désignés par leurs acronymes anglais respectifs : *Recurrent Neural Network* (RNN), *Linear Discriminant Analysis* (LDA), *Support Vector Machines* (SVM) et *Random Forest Classifier* (RFC).

Il convient de noter que les modèles SVM et les forêts aléatoires ont été appliqués avec succès dans certaines études de modélisation de l'état des fenêtres (Markovic et al., 2017). En revanche, les modèles de RNR et d'ADL semblent peu représentés dans le contexte de la modélisation de l'état des fenêtres, bien qu'ils aient montré leur efficacité sur des tâches similaires telles que la détection d'occupation (Candanedo and Feldheim, 2016; Huchuk et al., 2019) ainsi que dans d'autres domaines (Hochreiter and Schmidhuber, 1997) tels que le médical pour la détection d'anomalies (Fernando et al., 2022) ou l'énergétique pour l'optimisation des performances (Skrobek et al., 2020). Les modèles d'apprentissage automatique utilisés dans cette étude seront décrits plus en détail par la suite

4.2.3. Apport de l'expertise métier pour l'optimisation des performances

Le processus de sélection des variables est primordial et doit s'adapter à la tâche fixée ou à l'objectif visé (Emmott et al., 2016). Dai et al. (2020) classifient les variables utilisées pour modéliser les ouvertures de fenêtre en deux groupes distincts : les variables *environnementales* et les variables *non-environnementales*. Les variables environnementales proviennent de mesures de l'ambiance intérieure ou extérieure, telles que la température de l'air, la luminosité, la vitesse du vent, le rayonnement solaire, le bruit, *etc.* Les variables non-environnementales englobent des informations relatives aux bâtiments, aux occupants ou à l'horodatage, telles que le type de pièce, le sexe, l'âge ou l'heure de la journée.

Comme l'analysent Dai et al. (2020), quelle que soit la technique ou l'approche employée, les données les plus fréquemment utilisées proviennent des mesures d'ambiance, notamment la température de l'air extérieur et intérieur, l'humidité, la concentration de CO₂ intérieures et la vitesse du vent. Cependant, il demeure complexe d'évaluer précisément l'influence des variables sélectionnées sur les modèles d'ouvertures car cette influence semble fluctuer selon les techniques utilisées, voire les cas d'applications. Par exemple, dans le cas des modèles de régression logistique, la plupart des études indiquent que la température intérieure et extérieure sont les variables les plus influentes (Haldi and Robinson, 2009, 2008; Rijal et al., 2007). Cependant, pour les réseaux de neurones, une étude détaillée, telle que celle de Markovic et al. (2019), mettant en évidence l'impact des mesures pour les applications d'ouverture de fenêtres, apporte des conclusions différentes. Ces travaux analysent les poids acquis par chacun des neurones d'un modèle entraîné sur plus de vingt variables différentes et soulignent surtout l'importance des données environnementales intérieures, notamment la concentration de CO₂. Au regard de ces résultats, il semble donc préférable de ne pas limiter ces travaux à l'étude de certaines variables par rapport à d'autres. Une approche large incluant différentes mesures de l'environnement intérieur et extérieur est donc privilégiée pour cette étude.

Cependant, il est important de noter que la plupart des données utilisées dans les études recensées n'ont subi ni transformation (comme la dérivation ou le lissage) ni combinaison (comme les différences), malgré les avantages potentiels pour les performances des modèles (Bengio et al., 2014). C'est pourquoi, et compte tenu des points abordés dans les chapitres précédents, nous examinerons et appliquerons diverses transformations basées sur une expertise métier à plusieurs variables et modèles afin d'évaluer leur contribution et leur pertinence. Ces transformations seront discutées en détail ultérieurement.

4.3. Description du protocole expérimental

Cette section présente les méthodes d'évaluation des performances utilisées pour cette étude. Les modèles sélectionnés seront également introduits et plusieurs informations relatives à leur construction et leur entraînement seront apportées.

4.3.1. Processus d'évaluation des performances

4.3.1.1. Le F_1 -score

La détection d'ouverture de fenêtre est une tâche de classification binaire courante en apprentissage automatique, où les deux classes, ouverte et fermée, présentent un déséquilibre significatif en matière de répartition. Comme illustré par Dai et al. (2020), plusieurs métriques peuvent être employées pour évaluer les performances de classification de l'état d'une fenêtre. Dans cette étude, le F_1 -score est utilisé pour évaluer les performances des modèles. Cette métrique permet d'évaluer de manière globale les performances des modèles, en offrant une mesure commune utilisée dans d'autres études. Elle est particulièrement adaptée aux tâches de classification confrontées à des problèmes de déséquilibre de classes (Johnson and Khoshgoftaar, 2019). Cependant, il est important de noter que le F_1 -score présente certaines limitations. Comme le montre Fourure et al. (2021), il est fortement influencé par le taux de répartition des classes, aussi appelé « taux de contamination » dans le cas de classes normales et anormales. Cela peut rendre complexe la comparaison des performances entre deux ensembles de données différents. En outre, son usage peut être limité pour différencier les modèles dans cette application spécifique, notamment lorsque leurs performances sont similaires ou identiques.

Comme l'illustre la Figure 27, un état d'ouverture de fenêtre est défini, dans ces travaux, comme une séquence d'états ouverts (cellules pleines) entrecoupée par des états fermés (cellules vides). Les deux évaluations des modèles, (i) et (ii), aboutissent au même F_1 -score de 0,72, bien que leurs résultats diffèrent : (i) détecte seulement la moitié des ouvertures de manière précise, tandis que (ii) détecte toutes les ouvertures mais sous-

estime leur durée. Cela met en évidence les limites du F_1 -score lorsqu'il est employé pour évaluer les performances de détection des ouvertures de fenêtres. Ces limites s'appliquent également à d'autres métriques couramment utilisées dans ce contexte, comme la précision, le rappel, le taux de faux positifs ou le taux de vrais positifs (Dai et al., 2020).

L'introduction de métriques d'évaluation complémentaires (Seliya et al., 2009) semble être un atout essentiel pour comparer les résultats avec d'autres études, les interpréter, et orienter le choix des modèles en fonction de besoins spécifiques. Cela permet, par exemple, de mettre l'accent sur le nombre d'ouvertures détectées ou sur la durée totale de l'ouverture, en plus de la détection des vraies et fausses ouvertures. Dans la partie suivante, nous présenterons et discuterons des métriques plus adaptées à ce domaine, à savoir la classification des ouvertures.

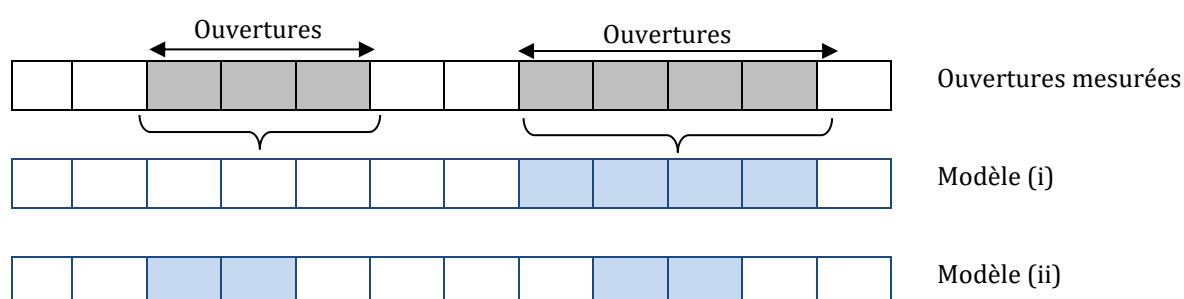
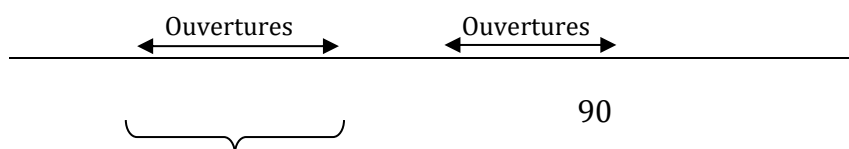


Figure 27 Exemple de modélisation d'ouvertures de fenêtres pour deux modèles.

4.3.1.2. Les vraies et fausses ouvertures

Une *vraie ouverture* correspond à une ouverture de fenêtre détectée par un modèle qui coïncide avec une ouverture de fenêtre réelle (ou mesurée) dans une plage de temps donnée. En d'autres termes, une vraie ouverture est un vrai positif ou une séquence continue de vrais positifs pouvant inclure un ou plusieurs faux positifs. *A contrario*, une ouverture de fenêtre détectée qui ne satisfait pas à cette exigence est considérée comme une *fausse ouverture*. Par exemple, pour une plage de temps de deux pas de temps utilisée dans cette étude, les résultats du modèle (i) et (ii) comprennent respectivement deux et une vraie ouverture, comme illustré dans la Figure 28. Cette restriction a pour objectif de pénaliser les modèles qui n'identifient pas de façon appropriée les ouvertures en pénalisant ceux qui représentent les ouvertures de fenêtre par un ensemble discontinu d'état ouverts et fermés, ainsi que ceux qui les surestiment ou les sous-estiment trop. Six métriques d'évaluation sont créées à partir de ces définitions: le nombre total de vraies et de fausses ouvertures, la durée totale des vraies et des fausses ouvertures, et la précision moyenne des vraies ouvertures.



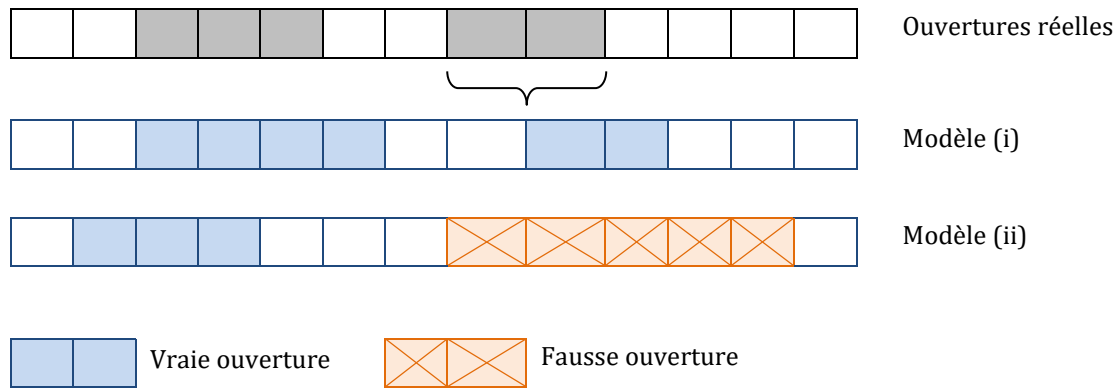


Figure 28 Exemples de vraies et de fausses ouvertures.

4.3.1.3. Métriques d'évaluation : le nombre total et la durée totale de vraies et de fausses ouvertures

Le *nombre total de vraies et de fausses ouvertures* évalue la capacité d'un modèle à détecter les ouvertures de fenêtres indépendamment de leur durée, en comptabilisant le nombre total de vraies et de fausses ouvertures détectées par le modèle.

La *durée totale de vraies et de fausses ouvertures* évalue la capacité d'un modèle à quantifier la durée des ouvertures de fenêtres. Ces métriques correspondent à la somme des durées de toutes les vraies et fausses ouvertures détectées par un modèle. Dans la Figure 28, nous observons que les résultats du modèle (i) affichent une durée totale de vraies ouvertures de six pas de temps sans aucune fausse ouverture, tandis que les résultats du modèle (ii) montrent une durée totale de vraies ouvertures de trois pas de temps pour une durée totale de fausses ouvertures de cinq pas de temps.

4.3.1.4. Métriques d'évaluation : la précision moyenne des vraies ouvertures

Pour cette métrique, un score est attribué à chaque action d'ouverture détectée (représentée par le premier état ouvert d'une vraie ouverture) et à chaque action de fermeture détectée (représentée par le dernier état d'ouverture d'une vraie ouverture) afin d'évaluer la précision du modèle dans la détection des ouvertures de fenêtres. Le score commence à 1 pour une correspondance parfaite entre l'ouverture réelle et l'ouverture détectée, puis diminue linéairement de 0,33 pour chaque différence d'un pas de temps. Cette pénalité de 0,33 est calculée en fonction de la limite de deux pas de temps fixée pour définir les vraies et fausses ouvertures. Les scores de précision associés à chaque ouverture sont ensuite moyennés pour obtenir un score moyen pour toutes les vraies ouvertures détectées par le modèle. Dans Figure 29, le modèle (i) affiche un score moyen de précision d'ouverture de 0,83 et un score moyen de précision de fermeture de 0,66, tandis que le modèle (ii) obtient un score moyen de précision d'ouverture et de fermeture de 0,66.

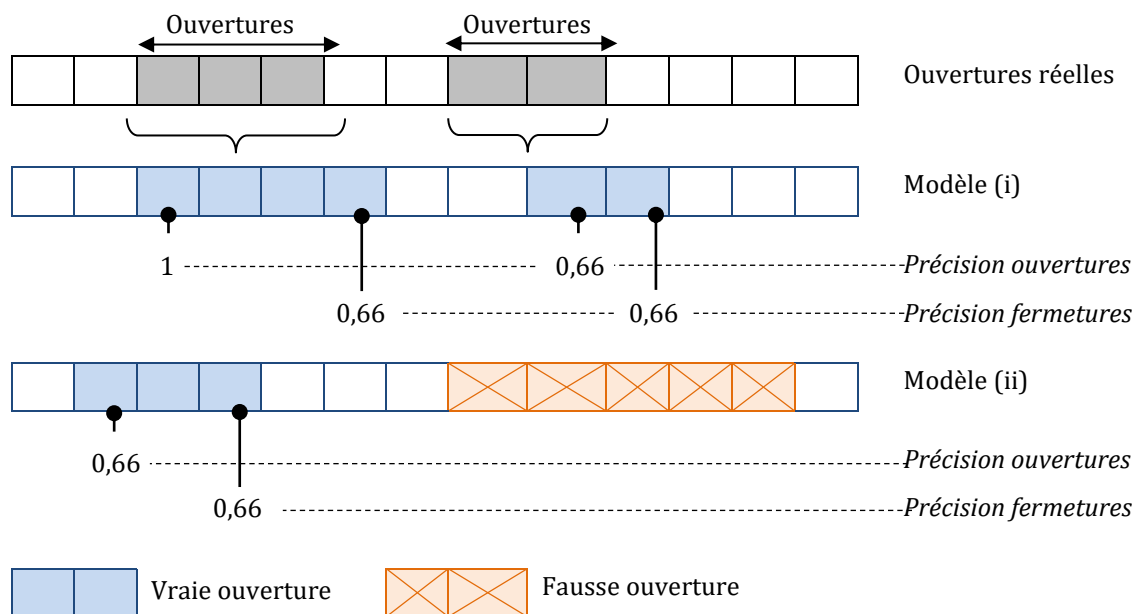


Figure 29 Exemples des scores de précision associés aux actions d'ouvertures et de fermetures de vraies ouvertures

4.3.2. Présentation des modèles

Dans cette partie, les différents modèles utilisés dans ces travaux seront introduits et leur fonctionnement sera expliqué en détail.

4.3.2.1. Les Séparateurs à Vaste Marge (SVM)

Les séparateurs à vaste marge, également appelées machines à vecteurs de support, sont des méthodes d'apprentissage automatique supervisé couramment utilisées pour la classification, la régression et la détection de nouveautés. Ces méthodes, dites « à noyaux », sont adaptées aux données qui ne peuvent pas être séparées de manière linéaire. Elles ont été introduites au milieu des années 1990 par Cortes et Vapnik (1995).

Les modèles SVM sont des classificateurs linéaires qui fonctionnent selon deux principes clés, illustrés dans la Figure 30. Lorsqu'il s'agit d'un problème de classification binaire où il est supposé que les données sont séparables dans l'espace qu'elles forment, il existe de nombreuses frontières possibles pour les séparer (i). Un modèle SVM cherche à déterminer l'hyperplan correspondant à la frontière optimale entre les deux classes en maximisant la distance, aussi appelée marge, entre chaque échantillon des deux classes les plus proches de cette frontière, appelées vecteurs de support (ii) (Bishop, 2006; Duda et al., 2001). Bien que les SVM soient intrinsèquement des classificateurs linéaires, leur capacité à traiter des données non linéaires découle d'une astuce appelée « *kernel trick* » ou « astuce du noyau ». En effet, si les données ne sont pas linéairement séparables dans leur espace d'origine, la fonction noyau des SVM permet de les projeter implicitement

dans un espace de dimension supérieure (potentiellement infini) où elles le deviennent (Theodoridis and Koutroumbas, 2008). Dans ce contexte, l'emploi du terme *implicite* signifie que la fonction de noyau simule le comportement des données dans un espace de dimension supérieure sans les projeter ni calculer leurs nouvelles coordonnées. Cette approche revient ainsi à calculer directement le produit scalaire entre les images des points dans le nouvel espace sans effectuer de transformations, ce qui la rend moins coûteuse en temps de calcul par rapport à une projection. Différents types de noyaux peuvent être utilisés selon la nature des données et des problèmes traités. Parmi les exemples courants de noyaux, il y a le noyau polynomial, le noyau *Radial Basis Function* (RBF) ou encore le noyau tangente hyperbolique (Theodoridis and Koutroumbas, 2008). Un exemple de résultats obtenus en utilisant un noyau RBF dans l'espace d'origine des données est illustré dans la Figure 30 (iii).

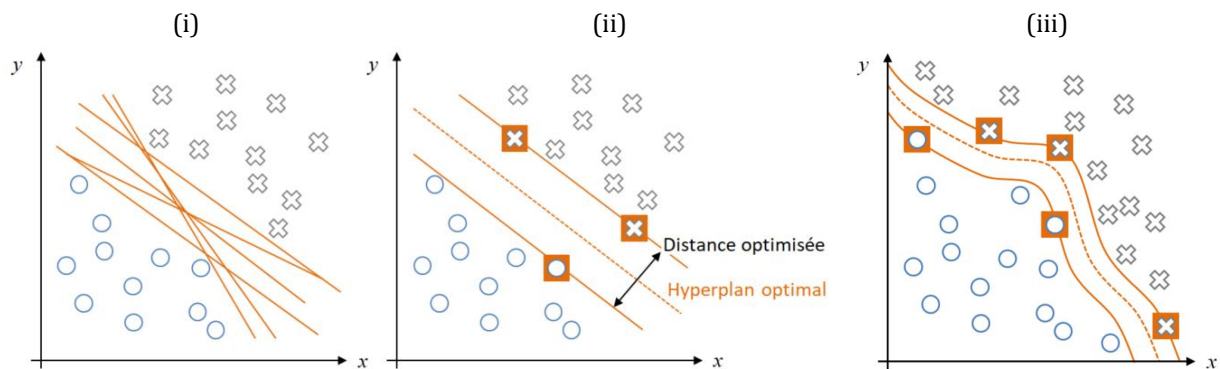


Figure 30 Illustration des principes clés des modèles de séparateurs à vaste marge à travers trois exemples de classification dans un espace en deux dimensions (x,y) . De gauche à droite : (i) montre l'infini de frontières possibles pour séparer deux classes ; (ii), modifiée d'après Cortes et Vapnik (1995), présente les vecteurs de support en orange plein, définissant la distance optimale pour séparer les deux classes ; (iii) modifiée d'après Theodoridis and Koutroumbas (2008), représente l'hyperplan optimal obtenu en utilisant un noyau RBF pour des données non linéairement séparables (ligne orange discontinue).

4.3.2.2. L'Analyse Discriminante Linéaire (ADL)

L'analyse discriminante linéaire (ADL) est une technique supervisée de réduction de dimension couramment utilisée pour la classification. Dans un problème de classification binaire, avec la même hypothèse que celle présentée pour les SVM, un modèle d'ADL construit une projection linéaire visant à maximiser la distance entre les moyennes des différentes classes (variabilité interclasse) tout en minimisant la dispersion au sein de chaque classe projetée (variabilité intraclasse) (Bishop, 2006; Duda et al., 2001). Un exemple de projection, ici aléatoire, est illustré dans la Figure 31. Plus précisément, en utilisant les données d'entraînement et les étiquettes de classe correspondantes, deux matrices de dispersion intra et interclasses sont calculées. L'objectif de cette méthode est de trouver une transformation (ou projection) qui maximise le rapport de ces dispersions. Les vecteurs propres associés à cette transformation, qui forment la base de l'espace

vectorel sur lequel les données sont projetées, sont appelés discriminants et correspondent à une direction optimale dans l'espace des caractéristiques. Au maximum, $C-1$ discriminants sont produits, où C est le nombre de classes. La classification de nouveaux échantillons repose en général sur le théorème de Bayes, qui estime la probabilité *a posteriori* d'appartenance d'un échantillon à une classe (Ghojogh and Crowley, 2019; Gokcen and Peng, 2002).

Bien que l'utilisation de l'ADL ne soit optimale que pour des données présentant une distribution normale et des matrices de covariance égales, sa simplicité et sa robustesse en font une option valable même lorsque ces conditions ne sont pas remplies (Li et al., 2006).

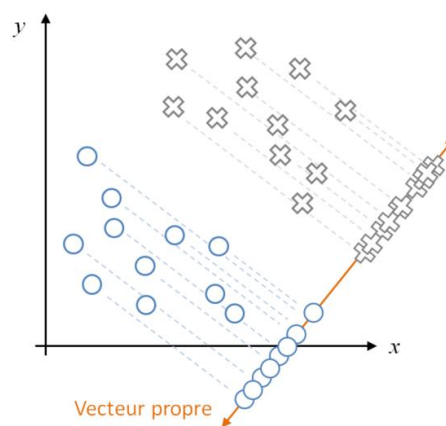


Figure 31 Exemple de réduction à une dimension par analyse discriminante linéaire pour un problème à deux classes en deux dimensions (x,y) . Les données sont projetées sur un vecteur propre, choisi ici aléatoirement, pour maximiser le rapport entre la variabilité interclasse et la variance intraclasse.

4.3.2.3. La forêt aléatoire

La forêt aléatoire est une méthode d'apprentissage automatique supervisée couramment utilisée pour la classification et la régression. Elle repose sur l'utilisation conjointe de nombreux arbres de décisions entraînés suivant une méthode ensembliste et a été introduite au début des années 2000 par Breiman (2001).

Un arbre de décision est une structure de modèle de décision qui utilise une approche hiérarchique pour résoudre un problème. Ils sont structurés selon un ensemble de nœuds, de branches et de feuilles (ou nœuds terminaux). Un test est effectué à chaque nœud, qui induit une subdivision représentée par une branche, laquelle mène à un autre nœud ou à une feuille qui équivaut à une décision finale (Kotsiantis, 2013). Plus précisément, l'arbre est construit en subdivisant récursivement l'ensemble des données à l'aide de tests de caractéristiques généralement basés sur des mesures telles que le gain d'information ou l'indice de Gini (Belgiu and Drăguț, 2016; Deng and Runger, 2013). La construction de l'arbre se poursuit jusqu'à atteindre un critère d'arrêt, tel qu'un nombre minimal

d'échantillons dans un nœud ou une longueur maximale atteinte. Cependant, les arbres ont une tendance au sur-ajustement (*overfitting* en anglais) pendant l'apprentissage (Breiman, 2001), ce qui se traduit souvent par un biais faible et une variance élevée. Pour réduire ce risque, les forêts aléatoires utilisent la méthode du « *bagging* », introduite par Breiman (1996), pour entraîner plusieurs arbres de décision. Le *bagging* est une méthode ensembliste qui implique la sélection aléatoire d'échantillons avec remplacement parmi l'ensemble des données d'entraînement pour former chaque arbre (Duda et al., 2001). De plus, lors de la division de chaque nœud, un sous-ensemble aléatoire est également sélectionné. Cette stratégie assure une diversité parmi les arbres, rendant la forêt aléatoire plus robuste et moins sujette au sur-ajustement (Bishop, 2006; Breiman, 2001; Hastie et al., 2009).

Dans le cas de données non numériques, il n'y a pas de réelle interprétation géométrique quant à la manière dont les nœuds séparent les données. Cependant, pour des données numériques, chaque nœud divise l'espace en un ensemble de rectangles qui répartissent de manière optimale les données en classes. Des exemples de représentation des arbres de décision pour des données non numériques et numériques sont fournis dans la Figure 32.

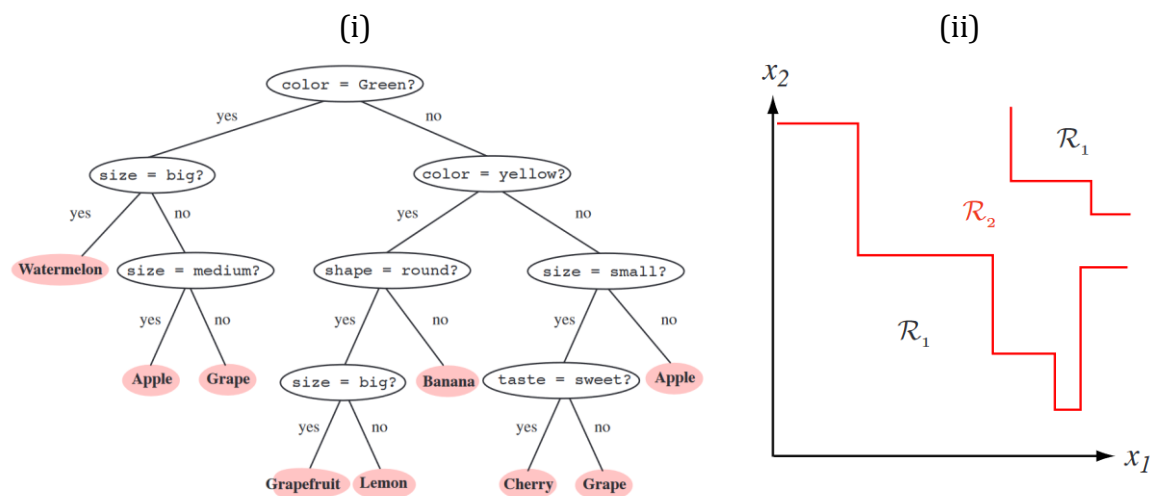


Figure 32 Exemples de représentation d'un arbre de décision pour des données non numériques (i) et des données numériques (ii), issus de Duda et al. (2001). Dans l'arbre (i), les nœuds correspondent aux questions, les branches aux réponses, et les feuilles portent les noms des fruits en fonction de leurs caractéristiques. Dans l'arbre (ii) deux régions de décision, $\mathcal{R}_1$ et $\mathcal{R}_2$, sont représentées.

4.3.2.4. Les Réseaux de Neurones Récurrents (RNN)

Les réseaux de neurone récurrents représentent une sous-famille des réseaux de neurones artificiels qui sont des modèles populaires d'apprentissage automatique.

Inspirés de la structure et du fonctionnement des neurones biologiques, ces réseaux sont organisés en couches successives, comme illustré dans la Figure 33 (i). Ils

comprennent une couche d'entrée, une ou plusieurs couches cachées et une couche de sortie, composées de neurones artificiels. Ces neurones sont interconnectés par des « synapses », chacune ayant un poids et un biais associés. Chaque neurone reçoit des valeurs en entrée des neurones précédents, effectue un calcul en pondérant chaque entrée, sommant les produits obtenus, et ajoutant le biais, puis applique une fonction d'activation (Bishop, 1994). Cette fonction peut prendre différentes formes, comme celle d'une sigmoïde, d'une tangente hyperbolique ou d'une fonction *Rectified Linear Unit* (ReLu), introduisant une non-linéarité pour permettre au réseau d'apprendre des représentations complexes (Bishop, 1994). Le neurone produit alors une valeur en sortie, qui, si elle dépasse un certain seuil d'activation, est dirigée vers d'autres neurones, voire directement vers l'utilisateur, si le neurone est situé dans la dernière couche (ii). L'apprentissage d'un réseau de neurones est guidé par la minimisation d'une fonction de coût (*loss* en anglais), telle que l'écart quadratique moyen (RMSE pour *Root-Mean-Square-Error*), l'erreur absolue moyenne (MAE pour *Mean Absolute Error*) ou l'entropie croisée (CE pour *cross-entropy*), qui évalue la différence entre les valeurs en sortie du réseau et les valeurs réelles attendues. Pendant l'entraînement, les poids et les biais du réseau sont ajustés par rétropropagation pour réduire l'erreur définie par cette fonction de coût (Rumelhart et al., 1986). Ce processus d'entraînement utilise des algorithmes d'optimisation avec essentiellement la descente de gradient (Wythoff, 1993) et ses variantes. Les réseaux de neurones peuvent être utilisés pour différentes applications telles que la classification, la prédiction, la réduction de dimension ou la reconstruction de données (Chalapathy and Chawla, 2019; Himeur et al., 2021b).

Contrairement aux réseaux de neurone traditionnels, où les neurones transmettent leurs sorties à d'autres neurones, les Réseaux de Neurones Récurrents (RNR) possèdent également une sortie dirigée vers eux-mêmes (iii et iv). Chaque neurone reçoit à la fois l'entrée actuelle et la sortie (ou état caché) du pas de temps précédent. Ces connexions récurrentes (ou cycliques) sont essentielles pour traiter des séquences, puisqu'elles permettent au réseau de conserver une sorte de mémoire des entrées précédentes. Ainsi, les RNR sont particulièrement adaptés pour capturer à la fois les relations temporelles et les dépendances de séquences, ce qui les rend régulièrement utilisés pour des tâches impliquant des données séquentielles, comme la traduction automatique, la prévision de séries temporelles ou la classification (Schmidhuber, 2015). Toutefois, en pratique, les RNR ne gèrent pas les dépendances à long terme puisqu'ils sont sujets aux problèmes de dissipation ou d'explosion du gradient (Bengio et al., 1994; Schmidhuber, 2015). Ce terme fait référence à la baisse de performance provoquée par un surplus d'information trop important que le réseau reçoit. Les réseaux de neurones à mémoire à court et long termes (LSTM pour *Long Short-Term Memory*) sont une variante de RNR spécifiquement conçue par Hochreiter et Schmidhuber (1997) pour résoudre ce problème en effectuant un oubli sélectif des informations à long terme. Pour ce faire, ils utilisent des structures de portes,

lesquelles sont au nombre de trois : la porte d'entrée, la porte d'oubli et la porte de sortie. Ces portes régulent le flux d'informations en décidant quelles parties de l'information doivent être respectivement stockées, oubliées, ou utilisées.

Les réseaux de neurones récurrents à portes (GRU pour *Gated Recurrent Unit*) sont une variante des LSTM qui utilisent moins de paramètres d'apprentissage, ce qui les rend moins gourmands en mémoire, plus rapides et potentiellement meilleurs que les LSTM dans certaines tâches (Chung et al., 2014; Greff et al., 2017; Jozefowicz et al., 2015). Les GRU combinent les portes d'oubli et d'entrée en une seule, nommée porte de mise à jour, tandis que la deuxième porte est appelée porte de réinitialisation et contrôle la quantité de données passées à considérer.

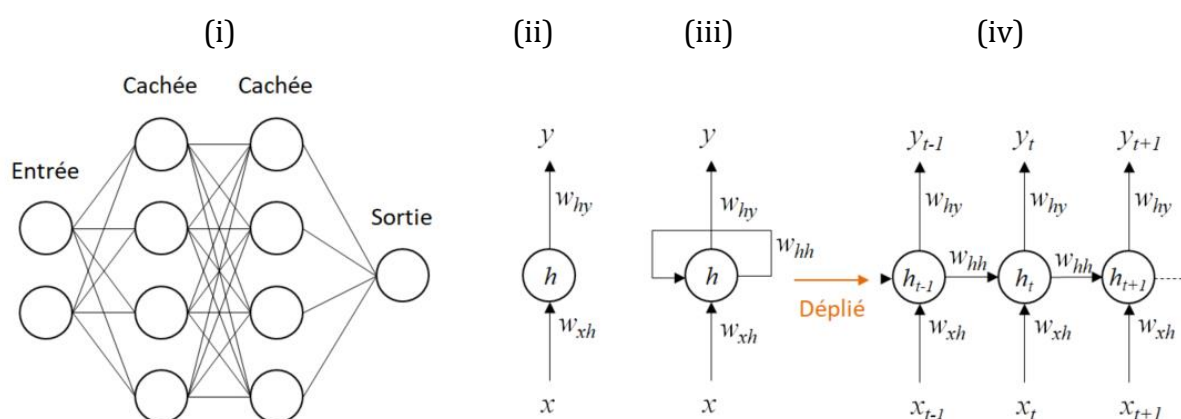


Figure 33 Représentation d'un réseau de neurones (i), d'une cellule neuronale classique (ii), d'une cellule neuronale récurrente (iii), et de la même cellule dépliée dans le temps (iv). Les couches d'entrée, cachée et de sortie à un moment donné t sont respectivement annotées x_t , h_t et y_t . Les poids associés aux connections des couches entrée-cachée, cachée-cachée et cachée-sortie sont respectivement annotés w_{xh} , w_{hh} et w_{hy} .

4.3.3. Configuration et entraînements des modèles

Dans cette partie, les processus d'entraînement réalisés pour chacun des modèles, ainsi que les paramètres retenus, seront présentés. Tous les modèles ont été construits en utilisant Python 3.8 avec les bibliothèques *Scikit-learn* (Pedregosa et al., 2011), *tensorflow*⁸ (Goldsborough, 2016) et plus spécifiquement *Keras*⁹ (Ketkar, 2017). Tous les entraînements ont été réalisés sur un processeur Intel Core i7-8850H CPU (6 cœurs à 2,6 GHz).

⁸ <https://www.tensorflow.org/>

⁹ <https://keras.io/>

4.3.3.1. Modèles SVM, ADL et forêt aléatoire

Pour ces travaux, des modèles de SVM avec noyau RBF, des modèles d'ADL et des modèles de forêt aléatoire ont été entraînés en suivant une procédure de validation croisée en cinq blocs adaptée aux séries temporelles (Berrar, 2019). Un algorithme de recherche adaptatif a également été employé pour optimiser les paramètres de chaque modèle. La validation croisée est une méthode d'échantillonnage issue de la famille des méthodes de Monte Carlo permettant d'évaluer les capacités de généralisation d'un modèle et de prévenir le phénomène de sur-apprentissage (Duda et al., 2001). Pour chaque combinaison et plage de paramètres énumérés dans le Tableau 7, les F_1 -scores calculés à partir de la validation croisée sont moyennés, et les modèles présentant meilleures performances sont sélectionnés afin d'être évalués.

Tableau 7 Liste et plage d'optimisation des paramètres et hyperparamètres de chaque modèles.

Modèles	Paramètres et hyperparamètres de configuration
SVM	Paramètre de régularisation (C) : 0,1 à 100 ; Gamma : 0,01 à 10
ADL	Solveur : <i>Singular value decomposition</i> , <i>Least squares solution</i> ou <i>Eigenvalue decomposition</i> ; Shrinkage : Ledoit-Wolf lemma ou aucune ; Limite absolue : 0,001 à 0,00001
Forêt aléatoire	Nombre d'arbres : 10 à 150 ; Nombre minimum d'échantillons requis pour scinder un nœud : 2 à 10 ; Nombre minimum d'échantillons requis pour être à un nœud final (feuille) : 1 à 4
LSTM et GRU	Nombre de couches cachées : 1 à 2 ; Nombre d'unités (taille) : 4 à 128 ; Valeur de la couche dropout : 0 à 0,5

4.3.3.2. Les modèles LSTM et GRU

Contrairement aux modèles présentés précédemment, les hyperparamètres des LSTM et des GRU, énumérés dans le Tableau 7, ont été préalablement optimisés pour atteindre un équilibre entre les performances, les difficultés d'apprentissage et le temps de calcul (Hasebrook et al., 2023). Les modèles de LSTM et de GRU appliqués dans cette étude se composent d'une première couche de 16, 32 ou 64 unités (selon le nombre de variables utilisées pour l'apprentissage), suivie d'une couche de sortie dense de taille 2 avec une fonction d'activation de type *Softmax*. L'algorithme d'optimisation *Adaptive Moment Estimation (Adam)* a été utilisé avec un taux d'apprentissage de 0,001 (Kingma and Ba, 2017), ainsi qu'une fonction coût adaptée aux problèmes binaire basée sur l'entropie croisée (*binary cross entropy*). L'algorithme *Adam* a été sélectionné pour ses performances et sa réputation dans la littérature comme dans la communauté (Soydaner, 2020). Pour prévenir le sur-ajustement, 25 % des données d'entraînement ont été réservées pour composer un ensemble de validation, les données restantes sont

mélangées avant chaque époque de l'entraînement (*epoch*) et une technique d'arrêt anticipé (*early stopping*) est utilisée pour mettre fin à l'apprentissage lorsque le modèle cesse de s'améliorer (Prechelt, 2012).

4.3.3.3. Spécificités des entraînements

À l'exception des modèles de forêts aléatoires, les données sont normalisées entre 0 et 1, en utilisant uniquement les caractéristiques observées dans l'ensemble d'entraînement. Ces caractéristiques de normalisation sont ensuite appliquées à l'ensemble des données de test. Les modèles de SVM, d'ADL et de forêts aléatoires sont entraînés de sorte à ce que chaque élément individuel en entrée du modèle produise une seule sortie, ce qui est aussi appelé approche '*one to one*' et est présenté dans la Figure 34 (i). En revanche, les modèles LSTM et GRU, conçus pour les séquences temporelles, sont entraînés selon le principe '*many to one*' avec une fenêtre d'observation glissante qui se décale d'un pas dans le temps. Dans ce cas, plusieurs éléments sont renseignés en entrée du modèle pour produire une seule sortie. Les fenêtres glissantes ont été centrées de manière à ce que la cible corresponde à l'état de la fenêtre au centre de la fenêtre d'observation. Cette approche est régulièrement appliquée pour la détection *a posteriori* puisqu'elle apporte des informations à la fois passées et futures (Blázquez-García et al., 2022). La Figure 34 (ii) illustre cet exemple avec une fenêtre de seulement trois pas de temps. La taille de la fenêtre d'observation est fixée à trois heures pour équilibrer le temps d'entraînement et les performances. L'approche '*many to one*' adaptée à la détection en temps réel, en ciblant le dernier élément de la séquence en entrée, est également présentée pour information en (iii). Dans le cas d'une approche basée sur la prédiction, la cible serait décalée d'un ou plusieurs pas supplémentaire vers la droite.

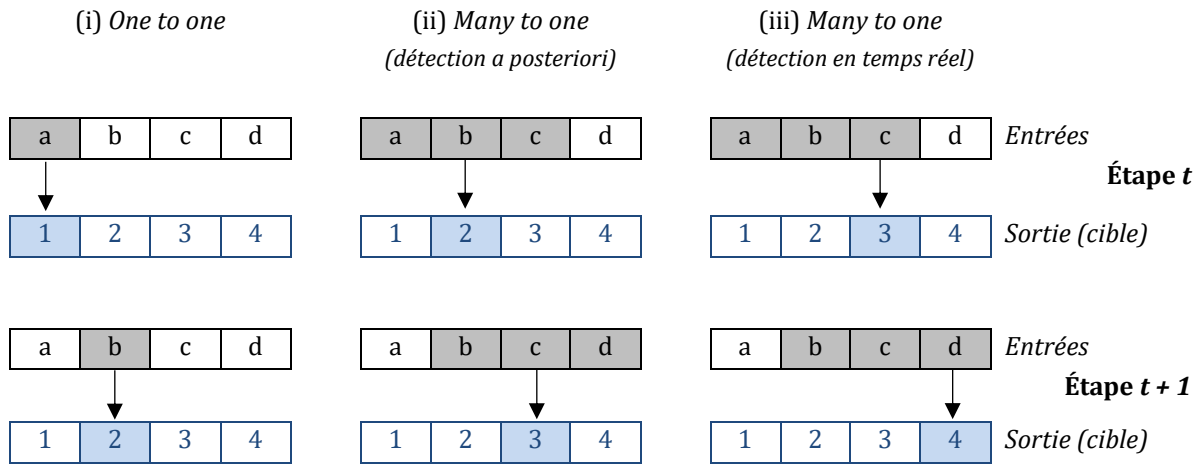


Figure 34 Approche d'entrainement réalisée selon les modèles à deux étapes temporelles successives t et $t+1$.

4.4. Description et prétraitement des données

Cette partie donne un aperçu des variables collectées et des diverses transformations appliquées. Les spécificités et les caractéristiques propres aux données utilisées pour les phases d'entrainement et de test sont ensuite brièvement expliquées. Enfin, le processus mené de transformation et de combinaison des données, provenant d'une expertise métier, est décrit.

4.4.1. Collecte et préparation des données

Les données brutes utilisées dans cette étude sont des mesures au pas de temps de 1 minute, collectées sur une période de deux ans, de juillet 2019 à février 2021. Elles proviennent d'une chambre orientée nord-ouest dans un appartement de type T3, situé dans le centre-ville de Bordeaux (France). Ces données brutes comprennent des mesures de l'ambiance intérieure, telles que la température, l'humidité relative et la concentration en CO_2 , ainsi que des mesures des conditions extérieures telles que la température et l'humidité relative. De plus, des mesures de l'état des fenêtres sont enregistrées au moyen de contacteurs (état ouvert : 1, état fermé : 0) et sont récoltées toutes les heures, sauf en cas de changement d'état. La position des capteurs pour toutes les mesures étudiées est illustrée dans la Figure 35. Parallèlement à ces données brutes, des données complémentaires sont générées et ajoutées à celles existantes. Par exemple, des informations telles que les minutes, les heures ou les jours de la semaine sont extraites des horodatages, tandis que d'autres variables, comme l'humidité absolue, sont calculées à partir des mesures existantes.

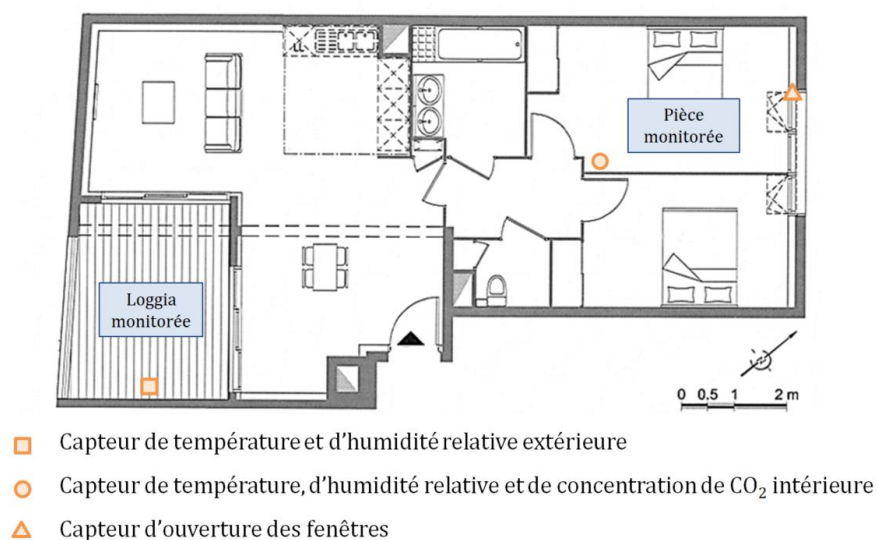


Figure 35 Positionnement des capteurs étudiés.

Sur les deux années de mesures disponibles, nous avons sélectionné quelques mois présentant un faible taux d'anomalies. Ces mois ont été nettoyés, agrégés à un pas de temps de 15 minutes, puis séparés en deux périodes distinctes : une période d'entraînement et une période de test. La période d'entraînement couvre une durée de quatre mois, s'étalant de fin octobre 2019 à début mars 2020. Elle comprend plus de 12 000 observations collectées pendant la saison de chauffage. Les caractéristiques de cette période d'entraînement sont résumées dans le Tableau 8. Il convient de noter que cette période d'étude présente quelques limitations, notamment le fait qu'elle se concentre sur une saison de chauffage, ce qui se traduit généralement par des températures intérieures supérieures à celles de l'air extérieur. La période de test, quant à elle, couvre un mois de données allant de la mi-décembre 2020 à la mi-janvier 2021, totalisant environ 3 100 observations. Cette période est également représentative d'une saison de chauffage et sera étudiée plus en détail dans la partie suivante.

Tableau 8 Caractéristiques des données utilisées pour l'entraînement.

Variable mesurée	Valeur maximale	Valeur minimale	Moyenne	Écart type
Température intérieure (°C)	23,0	19,6	21,7	0,5
Humidité relative intérieure (%)	69,8	29,6	49,7	6,1
Concentration de CO ₂ intérieure (ppm)	2000,0	436,2	727,3	237,4
Température extérieure (°C)	23,5	4,8	12,3	2,9
Humidité relative extérieure (%)	93,6	36,0	71,2	10,2
État des fenêtres (0-1)	1	0	-	-

4.4.2. Analyse des données de test

La Figure 36 donne un aperçu des données de test utilisées pour évaluer chaque modèle, en mettant en évidence deux périodes aux dynamiques distinctes : sans occupants (zone centrale bleue) et avec occupants (zones restantes).

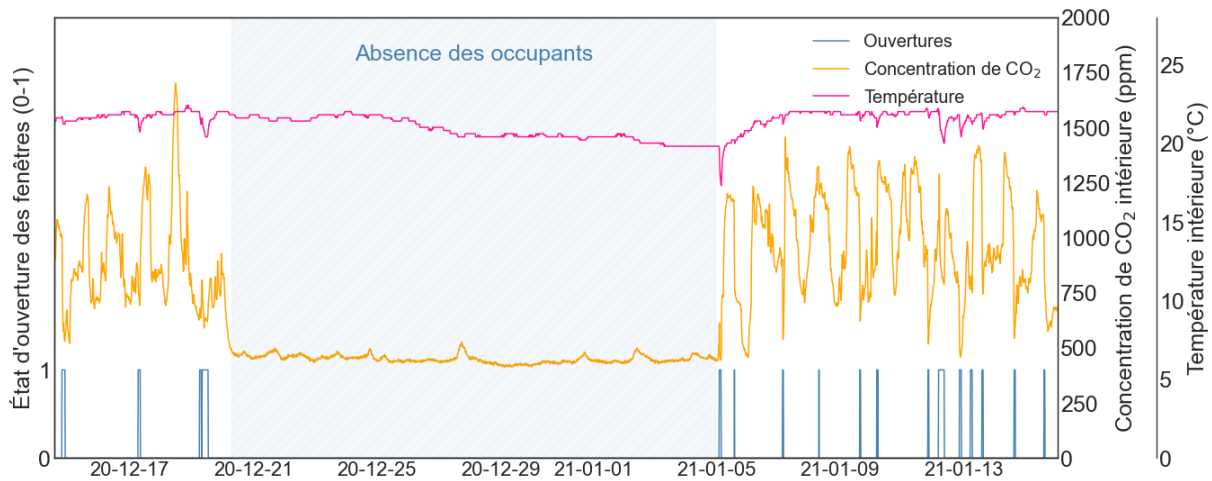


Figure 36 Période de test avec occupation et inoccupation (zone bleue centrale annotée).

Pendant la période d'inoccupation de l'appartement, la température intérieure présente une lente baisse d'environ 2 °C, et la concentration de CO₂ reste basse et stable. Cela est illustré dans le Tableau 9. En revanche, pendant la période occupée, tant la température intérieure que la concentration de CO₂ sont plus élevées en moyenne, se rapprochant des valeurs observées dans les données d'entraînement. Cependant, il y a des différences notables, notamment en ce qui concerne la concentration de CO₂. Ces écarts caractéristiques d'un changement de la variabilité et de la tendance des observations, peuvent influencer l'évaluation des modèles et fournir des informations utiles sur leur comportement vis-à-vis de données atypiques.

Pendant toute la période de test, le temps cumulé pendant lequel les fenêtres étaient ouvertes est de 25,75 heures, pour un total de 18 ouvertures. Ces ouvertures peuvent être divisées en deux catégories en fonction de leur impact sur les conditions intérieures : (1) les **ouvertures à faible impact**, caractérisées par des fluctuations lentes, généralement de petites tailles et de faible intensité, des mesures intérieures, souvent dues à une brève ou légère ouverture des fenêtres, et (2) les **ouvertures à fort impact**, qui comprennent toutes les autres ouvertures ayant un impact significatif tant par leur taille que leur intensité. Les ouvertures à faible impact représentent 5 heures du temps total pendant lequel les fenêtres étaient ouvertes, sur les 25,75 mesurées, et sont considérées comme plus difficiles à détecter pour les modèles par rapport aux ouvertures à fort impact.

Tableau 9 Caractéristiques des données de test, avec (O) pour la période occupée, et (Ø) pour la période non occupée. Les valeurs qui diffèrent le plus des données d'entraînement sont en gras.

Variable mesurée	Valeur maximale	Valeur minimale	Moyenne	Écart type
(O) Température intérieure (°C)	22,4	17,3	21,6	0,6
(Ø) Température intérieure (°C)	22,0	19,8	20,7	0,7
(O) Humidité relative intérieure (%)	58,1	26,5	43,4	6,3
(Ø) Humidité relative intérieure (%)	54,0	34,0	42,8	5,8
(O) Concentration de CO ₂ intérieure (ppm)	1700,7	440,7	932,9	236,9
(Ø) Concentration de CO ₂ intérieure (ppm)	526,2	411,5	449,2	17,8

4.4.3. Prétraitement des données par transformation et association

La sélection et la représentation des données jouent un rôle crucial dans les performances des modèles, comme le soulignent Bengio et al. (2014), Emmott et al. (2016) et Razavi et al. (2019). Cette étape se révèle particulièrement critique pour les problèmes de détection et de classification, car elle simplifie et améliore le processus d'entraînement en fournissant des informations distinctives que le modèle aurait du mal à extraire par lui-même. Dans le contexte de ces travaux, les ouvertures de fenêtres peuvent être considérées comme des anomalies, puisqu'elles représentent un processus qui génère des observations sensiblement différentes des autres (Hawkins, 1980), se distinguant par des motifs spécifiques. Dans cette perspective, l'ingénierie des caractéristiques consiste en partie à **transformer** les données pour mieux caractériser ou extraire ces motifs spécifiques, voire à les mettre en évidence, pour faciliter l'apprentissage de l'algorithme.

Ainsi, plusieurs variables complémentaires sont générées par des transformations de données dont le choix relève autant d'observations que de connaissances métier. Certaines de ces variables découlent des caractéristiques particulières d'une saison de chauffage et visent à refléter l'impact de l'ouverture d'une fenêtre entre deux environnements distincts : l'intérieur et l'extérieur. D'autres variables sont créées pour capturer différentes spécificités relatives aux données, telles que la dynamique. Leur impact sur l'ensemble des données de température et de concentration de CO₂ intérieures utilisées pour ces travaux est illustré dans la

- **Moyenne Mobile Exponentielle (MME).** Cette transformation repose sur une technique de lissage couramment utilisée pour réduire le bruit de mesure et ne capturer que les motifs importants. La MME d'une mesure est calculée selon l'équation (1), où M_t correspond à la valeur de la mesure au temps t , MME_{M_t} est la valeur de la

MME pour cette mesure au temps t , et α est un coefficient de lissage constant compris entre 0 et 1. Dans cette étude, une MME est appliquée aux données avec un coefficient de lissage léger ($\alpha = 0,10$), choisi dans un intervalle de valeurs compris entre 0,10 et 0,25. Les mesures lissées sont désignées comme **MME**_{mesure}.

$$MME(\alpha)_{M_t} = \begin{cases} M_0 & t = 0 \\ \alpha \times M_t + (1 - \alpha) \times MME_{M_{t-1}} & t > 0 \end{cases} \quad (1)$$

- **Dérivée.** Cette transformation est utilisée pour capturer les variations soudaines des mesures en différenciant les tendances saisonnières et cycliques, représentées par des dérivées successives de faible amplitude, des chutes soudaines, telles que celles provoquées par l'ouverture d'une fenêtre, prenant la forme de dérivées successives à forte amplitude. Cette transformation est présentée dans l'équation (2), où M_t est la valeur de la mesure à l'instant t , $Dérivée_{M_t}$ est la valeur de la dérivée de cette mesure à l'instant t et Δt correspond à l'intervalle de temps entre les mesures. Les mesures dérivées sont désignées comme **Dérivée**_{mesure}.

$$Dérivée_{M_t} = \begin{cases} 0 & t = 0 \\ \frac{M_t - M_{t-\Delta t}}{\Delta t} & t > 0 \end{cases} \quad (2)$$

- **Résidu_{STL}.** Cette transformation repose sur la méthode *Seasonal-Trend decomposition using LOESS* (STL), une technique statistique qui permet de décomposer les séries temporelles en trois composantes distinctes (Audibert, 2021) : une composante saisonnière représentant les motifs récurrents dans le temps, une composante de tendance, et un résidu (ou bruit) composé de fluctuations aléatoires ou imprévisibles (Cleveland et al., 1990). Par conséquent, le résidu peut être utilisé pour détecter des motifs inhabituels, tel que ceux résultant de l'ouverture d'une fenêtre. Le résidu_{STL} est évalué à l'aide de l'équation (3), où M_t est la valeur de la mesure à l'instant t et $Saisonnalité_{M_t}$, $Tendance_{M_t}$ et $Résidu_{M_t}$ sont respectivement les composantes de saisonnalité, de tendance et résidu de la mesure à cet instant. Les résidu_{STL} d'une mesure sont désignés comme **Résidu**_{mesure}.

$$Residu_{M_t} = M_t - Tendance_{M_t} - Saisonnalité_{M_t} \quad (3)$$

- **Différence MME-Mesure (DMM).** Cette combinaison de variables, à l'objectif similaire à la transformation basée sur le résidu_{STL}, consiste à soustraire à une mesure brute (M_t) la même mesure préalablement lissée par la méthode de la MME (MME_{M_t}). L'équation (4) décrit cette opération, où M_t est la valeur de la mesure au temps t , MME_{M_t} est la valeur de la MME appliquée à cette mesure au temps t , et **Différence**_{Mt}

correspond à la différence entre les deux au temps t . La MME appliquée est configurée avec un coefficient de lissage (α) adapté, compris entre 0,01 et 0,10, pour extraire uniquement les tendances des mesures. Par conséquent, une valeur élevée (resp. faible) de $Différence_{M_t}$ caractérise un point de mesure qui s'éloigne (resp. se rapproche) de la tendance, ce qui rend plus propice à être considéré comme inhabituel (resp. habituel). Plusieurs coefficients de lissage ont été testés sur la période d'apprentissage, mais une valeur de 0,04 a été retenue pour ces travaux. Cette variable combinée est désignée comme ***Différence***_{mesure}.

$$Différence_{M_t} = M_t - MME_{M_t} \quad (4)$$

Ces transformations sont appliquées aux mesures d'ambiance intérieure, notamment la température, l'humidité relative et absolue et la concentration de CO₂. Elles peuvent être associées entre elles pour fournir différentes informations aux modèles. Pour chaque mesure intérieure, une vingtaine d'associations différentes sont réalisées. En comparant les phases d'occupation (O) et d'inoccupation prolongées (Ø) de la période de test à la période d'entraînement,

permet de montrer que les transformations proposées, notamment le résiduTM (B), la DMM (C), et la dérivée (D), semblent fournir des informations exemptes de variations saisonnières (aussi bien locales que globales), avec des données centrées. Cet effet est d'autant plus notable pour les mesures de température, dont la dispersion initiale semble sensiblement différente entre les périodes d'entraînement et de test (A).

Ces transformations ont pour objectif de rendre les modèles moins sensibles aux changements de dynamique s'étendant sur de longues périodes, en privilégiant les changements brusques à court terme. Cependant, elles présentent quelques inconvénients : les transformations ou combinaisons basées sur la MME peuvent introduire un décalage proportionnel à la constante de lissage, tandis que les transformations basées sur la dérivée peuvent produire des données avec une faible variance, ce qui peut rendre leur exploitation complexe selon les modèles à utiliser. Dans la suite de ces travaux, le terme d'**association** sera utilisé pour désigner les appariements de transformations pour une même variable, tandis que le terme de **combinaison** sera employé lorsque plusieurs variables, ou associations de transformations, sont appariées. Par exemple, une mesure de température non transformée à laquelle s'ajoute une ou plusieurs de ses transformations (comme sa dérivée) constitue une association alors que la même association à laquelle on ajoute une mesure, une transformation ou une association provenant de la mesure d'humidité représente une combinaison.

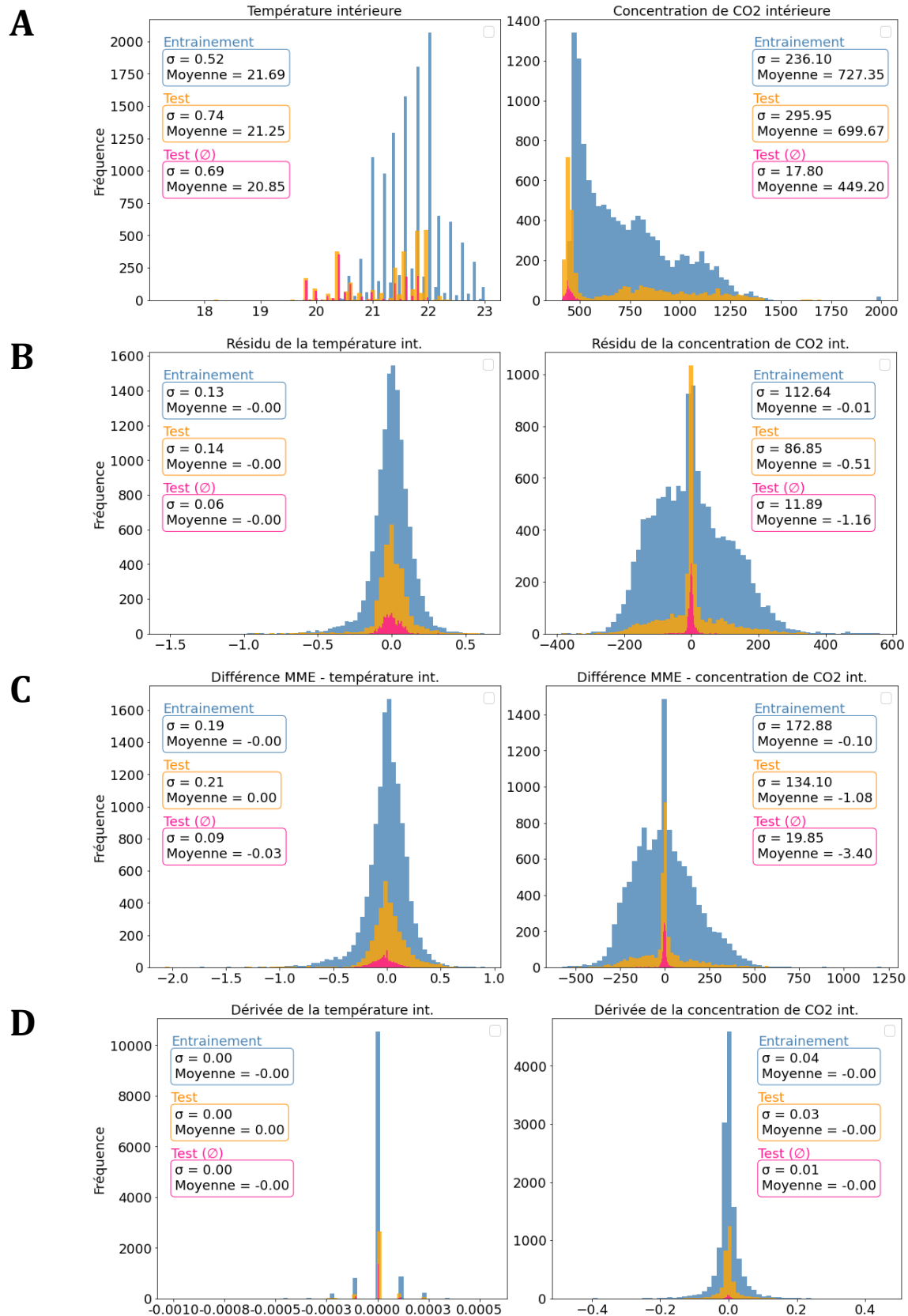


Figure 37 Distribution des données pour les mesures de température et de concentration de CO₂ intérieures (A) ainsi que pour leurs transformations et combinaisons respectives. Les transformations incluent le résidu (B), la DMM (C), et la dérivée (D). Les données sont représentées pour les périodes d'entraînement (en bleu), de test (en orange) et de test sans occupants (en rose).

4.4.4. Combinaisons et associations des données

Au total, environ 800 combinaisons différentes de mesures, de transformations et d'associations de mesures ont été générées pour chaque modèle, comme présenté dans la Figure 38. Initialement, une base d'environ 20 associations différentes a été évaluée pour chaque mesure d'ambiance intérieure (température, humidité relative, humidité absolue et concentration de CO₂) ainsi que pour les données temporelles (A). Au regard des résultats obtenus, environ 175 combinaisons supplémentaires ont été générées en combinant par paire les associations provenant de chacune des mesures intérieures tout en gardant une mesure commune basée sur la température (B). Cela a conduit à la création de deux paires de mesures : température + CO₂ et température + humidité. En suivant un processus similaire, jusqu'à 100 combinaisons additionnelles ont été générées en associant les mesures de température, d'humidité et de concentration de CO₂ intérieures (C), puis jusqu'à 300 autres combinaisons ont été créées en ajoutant les mesures extérieures (température et humidité) à toutes les combinaisons précédentes (D).

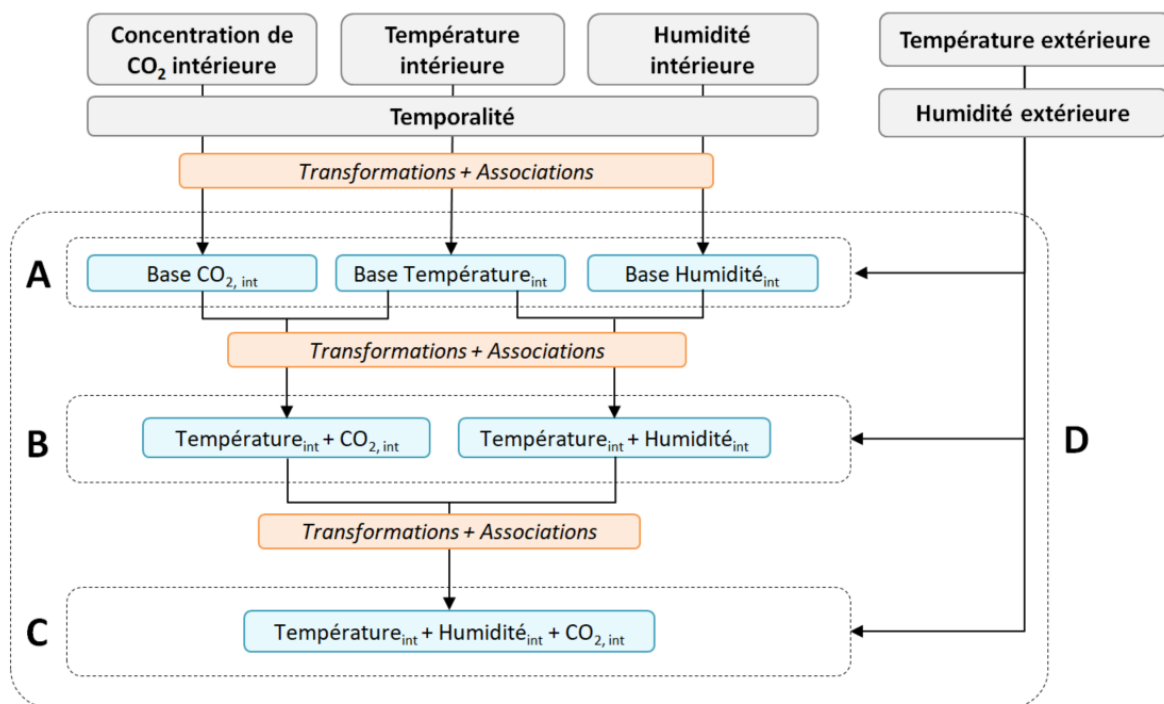


Figure 38 Schéma du protocole d'expérimentation pour la combinaison et l'association des mesures.

4.5. Résultats

Pour les modèles LSTM, GRU et de forêts aléatoires, chacune des 800 combinaisons a été répétée dix fois en raison de leur caractère stochastique. Parmi ces répétitions, seule la meilleure en matière de F₁-score a été retenue pour l'évaluation. Les modèles ADL et

SVM, fournissant des résultats plus stables, ont été évalués une seule fois pour chaque combinaison. Pour simplifier la présentation des résultats, seules les associations basées sur l'humidité absolue sont présentées, en désignant simplement cette variable comme « humidité ». De plus, l'utilisation de données liées à la temporalité telles que les minutes, les heures et les jours de la semaine, a conduit à une diminution des performances des modèles et n'est par conséquent pas présentée.

Les résultats suivants sont évalués en fonction du F₁-score et d'autres métriques alternatives, comme mentionné précédemment. Elles permettront de quantifier et de discuter des transformations, associations et combinaisons de mesures les plus adéquates pour détecter les ouvertures de fenêtres. Il est important de noter que l'objectif de ces travaux ne se limite pas à évaluer les performances brutes des modèles en matière de détection des ouvertures de fenêtres. Bien que ces performances fournissent des informations pertinentes, elles sont étroitement liées à ce cas d'étude spécifique et ne peuvent être généralisée sans précaution.

4.5.1. Évaluation basée sur le F₁-score

4.5.1.1. Influence des transformations de chaque mesure sur les performances de détection

Dans cette section, nous évaluons les performances de chaque modèle en utilisant les vingt transformations et associations de base pour chaque mesure d'ambiance intérieure. Les résultats sont résumés dans le Tableau 10, qui présente le F₁-score moyen et l'écart-type de la vingtaine de résultats pour chaque modèle. Nous constatons que les associations basées uniquement sur l'humidité de l'air ou la concentration de CO₂ intérieures (H_{int} et C_{int}) ne donnent pas de bons résultats, avec un F₁-score moyen généralement inférieur à 0,39. En revanche, les associations basées sur la mesure de la température intérieure (T_{int}) semblent offrir des résultats plus prometteurs, avec un F₁-score moyen supérieur pour tous les modèles, en particulier pour les RNR, qui atteignent une valeur moyenne proche de 0,70.

Tableau 10 F₁-score moyen et écart-type des 20 transformations et associations basées sur une mesure pour chaque modèle. Les meilleurs résultats par modèle sont en gras.

F₁-score: <i>moyenne ± écart-type</i>	GRU	LSTM	ADL	SVM	Forêt aléatoire
Humidité abs. int. (H _{int})	0,35 ± 0,23	0,39 ± 0,20	0,11 ± 0,11	0,18 ± 0,16	0,21 ± 0,13
CO ₂ int. (C _{int})	0,28 ± 0,15	0,19 ± 0,14	0,01 ± 0,01	0,04 ± 0,04	0,18 ± 0,08
Température int. (T _{int})	0,73 ± 0,07	0,68 ± 0,18	0,54 ± 0,24	0,38 ± 0,28	0,46 ± 0,21

Toutefois, ces résultats présentent une forte dispersion qui peut être expliquée en examinant les F_1 -score pour les cinq transformations de base, sans associations, présentés dans le Tableau 11. Pour les mesures d'humidité de l'air ambiant, les transformations basées sur la DMM et le résidu_{STL} augmentent les performances des modèles par rapport aux autres transformations ou aux mesures initiales. Les meilleures performances sont observées avec la DMM pour les modèles LSTM et GRU, atteignant un F_1 -score d'environ 0,70. Concernant la température de l'air, la dispersion des résultats semble être due aux faibles performances des modèles lorsqu'ils sont entraînés avec des mesures non transformées, lissées par MME ou avec d'autres associations pouvant les inclure. Ces effets peuvent s'expliquer par les différences entre l'ensemble d'entraînement et l'ensemble de test. Les mesures de la période de test présentent des tendances et des valeurs non observées pendant la phase d'apprentissage, ce qui permet aux transformations plus robustes aux variations de tendance, telles que la dérivée, le résidu_{STL} ou la DMM, d'être plus performantes et d'augmenter le F_1 -score de 0,60 à 0,77 pour les modèles LSTM et GRU (ce qui représente une amélioration de 25 %) et de 0,00 à environ 0,60 pour les modèles ADL, SVM et de forêts aléatoires. Cependant, ces résultats n'incluent pas les combinaisons entre différentes mesures, lesquelles peuvent grandement influencer sur les performances de détection et sont présentées dans la partie suivante.

Tableau 11 F_1 -score pour les cinq transformations de base appliquées aux mesures de température intérieure (T_{int}) et d'humidité intérieure (H_{int}). Les meilleurs résultats par modèle sont en gras.

F_1-score des transformations	GRU	LSTM	ADL	SVM	Forêt aléatoire
T_{int}	0,60	0,57	0,00	0,04	0,04
MME T_{int}	0,54	0,00	0,00	0,00	0,03
Dérivée T_{int}	0,77	0,75	0,25	0,25	0,50
Différence T_{int}	0,75	0,72	0,65	0,62	0,60
Résidu T_{int}	0,77	0,70	0,55	0,54	0,49
H_{int}	0,00	0,26	0,00	0,00	0,04
MME H_{int}	0,00	0,00	0,00	0,00	0,06
Dérivée H_{int}	0,43	0,40	0,16	0,14	0,07
Différence H_{int}	0,72	0,71	0,13	0,17	0,41
Résidu H_{int}	0,57	0,52	0,41	0,40	0,35

4.5.1.2. L'influence du choix des mesures sur les performances de détection

Pour obtenir une vue d'ensemble des performances optimales pour chaque modèle en fonction des combinaisons de mesures, les vingt meilleurs modèles pour chaque combinaison de mesures sont sélectionnés en utilisant le critère du F1-score. La moyenne et l'écart-type des résultats sont résumés dans le Tableau 12, couvrant ainsi un tiers des 800 combinaisons d'origine (280) pour chaque modèle.

Les combinaisons de paires de mesures intérieures ($T_{int} + H_{int}$ et $T_{int} + C_{int}$) montrent des performances de détection d'ouverture de fenêtre plus élevées, avec une augmentation systématique du F1-score maximum et moyen. Parmi toutes les combinaisons, la paire $T_{int} + C_{int}$ se distingue en offrant des performances significativement supérieures à celles de $T_{int} + H_{int}$. L'ajout de la troisième mesure intérieure ($T_{int} + H_{int} + C_{int}$) semble apporter une légère amélioration, voire aucune, par rapport à la paire $T_{int} + C_{int}$. En ce qui concerne les F1-score, les LSTM, les GRU, les ADL et les forêts aléatoires ne semblent pas bénéficier de manière significative de l'ajout d'informations apportées par l'humidité, contrairement aux SVM. Par conséquent, les résultats des modèles LSTM et GRU pour la détection de l'ouverture des fenêtres semblent plafonner autour d'un F1-score moyen de 0,76 à 0,78, tandis que ceux des ADL, des SVM et des forêts aléatoires semblent être limités autour d'un score moyen plus bas, de 0,72 à 0,73. Enfin, à l'exception des modèles d'ADL, l'ajout des mesures de température extérieure (T_{ext}), d'humidité extérieure (H_{ext}) ou des deux ($T_{ext} + H_{ext}$) à toutes les combinaisons de mesures intérieures tend à généralement détériorer les performances de détection de tous les modèles, entraînant une baisse commune du F1-score maximal et moyen. Cependant, même si l'ajout de la température extérieure aux mesures intérieures semble légèrement détériorer les performances maximales atteignables, avec une baisse du score F1-score pour les RNR, cette approche s'avère plus intéressante pour les ADL, les SVM et les forêts aléatoires que l'utilisation exclusive de la température de l'air intérieur.

Parmi tous les modèles, les LSTM et les GRU semblent être les plus efficaces pour la détection de l'état des fenêtres, affichant les meilleurs F1-score moyens et maximums, même lorsqu'une seule mesure est utilisée. Ces deux modèles, auxquels on peut ajouter la forêt aléatoire, semblent également sensibles à l'ajout d'informations complémentaires, ce qui nécessite une sélection minutieuse des variables ou, à défaut, une transformation appropriée des données. Les modèles ADL et SVM semblent plus fiables, avec une répartition plus stable des résultats et la capacité à maintenir ou améliorer leurs performances malgré l'ajout de mesures qui détériorent les résultats des autres modèles. Enfin, dans l'ensemble des résultats obtenus, l'approche basée sur les forêts aléatoires semblent être la moins stable.

En résumé, même si la combinaison des trois mesures intérieures ($T_{\text{int}} + H_{\text{int}} + C_{\text{int}}$) semble offrir les meilleurs résultats pour la détection des ouvertures, l'utilisation des deux mesures intérieures telles que $T_{\text{int}} + C_{\text{int}}$ ou même $T_{\text{int}} + H_{\text{int}}$, pourrait suffire à obtenir de bons voire d'excellents résultats. Malgré leur fluctuation, qui ne dépendent pas uniquement de l'ouverture des fenêtres (elle est également influencée par l'occupation, la position de l'occupant dans la pièce et la circulation naturelle de l'air), les données de concentration de CO_2 intérieure semblent être préférables à l'humidité intérieure. De plus, et selon les transformations appliquées, la mesure de température intérieure s'avère à elle seule suffisante pour fournir des résultats de détection d'ouverture de fenêtres comparables aux combinaisons à deux ou trois variables. En outre, étant donné que les mesures de la température extérieure ont tendance à diminuer les performances de détection pour la majorité des modèles, la suite de cette étude se concentrera principalement sur les mesures intérieures.

Tableau 12 F-score moyen et écart-type des 20 meilleures combinaisons entre : la température d'air intérieure (T_{int}), l'humidité intérieure (H_{int}), la concentration de CO_2 intérieure (C_{int}), la température d'air extérieure (T_{ext}), et l'humidité extérieure (H_{ext}). Les meilleurs résultats par modèle sont en gras.

F-score: moyenne $\pm$ écart-type	GRU	LSTM	ADL	SVM	Forêt aléatoire
H_{int}	0,35 $\pm$ 0,23	0,39 $\pm$ 0,20	0,11 $\pm$ 0,11	0,18 $\pm$ 0,16	0,21 $\pm$ 0,13
C_{int}	0,28 $\pm$ 0,15	0,19 $\pm$ 0,14	0,01 $\pm$ 0,01	0,04 $\pm$ 0,04	0,18 $\pm$ 0,08
T_{int}	0,73 $\pm$ 0,07	0,68 $\pm$ 0,18	0,54 $\pm$ 0,24	0,38 $\pm$ 0,28	0,46 $\pm$ 0,21
$T_{\text{int}} + T_{\text{ext}}$	0,70 $\pm$ 0,03	0,70 $\pm$ 0,03	0,67 $\pm$ 0,03	0,66 $\pm$ 0,04	0,62 $\pm$ 0,08
$T_{\text{int}} + H_{\text{int}}$	0,76 $\pm$ 0,01	0,75 $\pm$ 0,02	0,69 $\pm$ 0,01	0,68 $\pm$ 0,03	0,65 $\pm$ 0,04
$T_{\text{int}} + H_{\text{int}} + T_{\text{ext}}$	0,71 $\pm$ 0,03	0,69 $\pm$ 0,05	0,68 $\pm$ 0,02	0,68 $\pm$ 0,04	0,63 $\pm$ 0,11
$T_{\text{int}} + H_{\text{int}} + H_{\text{ext}}$	0,71 $\pm$ 0,03	0,71 $\pm$ 0,02	0,69 $\pm$ 0,01	0,66 $\pm$ 0,04	0,53 $\pm$ 0,05
$T_{\text{int}} + H_{\text{int}} + T_{\text{ext}} + H_{\text{ext}}$	0,73 $\pm$ 0,03	0,72 $\pm$ 0,02	0,70 $\pm$ 0,01	0,68 $\pm$ 0,03	0,53 $\pm$ 0,07
$T_{\text{int}} + C_{\text{int}}$	0,78 $\pm$ 0,01	0,76 $\pm$ 0,01	0,72 $\pm$ 0,01	0,70 $\pm$ 0,01	0,73 $\pm$ 0,02
$T_{\text{int}} + C_{\text{int}} + T_{\text{ext}}$	0,70 $\pm$ 0,03	0,65 $\pm$ 0,05	0,71 $\pm$ 0,02	0,67 $\pm$ 0,02	0,71 $\pm$ 0,04
$T_{\text{int}} + H_{\text{int}} + C_{\text{int}}$	0,78 $\pm$ 0,01	0,76 $\pm$ 0,01	0,71 $\pm$ 0,01	0,72 $\pm$ 0,01	0,73 $\pm$ 0,01
$T_{\text{int}} + H_{\text{int}} + C_{\text{int}} + T_{\text{ext}}$	0,70 $\pm$ 0,05	0,64 $\pm$ 0,07	0,71 $\pm$ 0,01	0,69 $\pm$ 0,04	0,71 $\pm$ 0,04
$T_{\text{int}} + H_{\text{int}} + C_{\text{int}} + H_{\text{ext}}$	0,73 $\pm$ 0,02	0,70 $\pm$ 0,02	0,72 $\pm$ 0,01	0,70 $\pm$ 0,02	0,68 $\pm$ 0,02
$T_{\text{int}} + H_{\text{int}} + C_{\text{int}} + T_{\text{ext}} + H_{\text{ext}}$	0,72 $\pm$ 0,01	0,69 $\pm$ 0,03	0,73 $\pm$ 0,01	0,69 $\pm$ 0,02	0,65 $\pm$ 0,02

4.5.1.3. Impact des associations et transformations sur les performances de détection

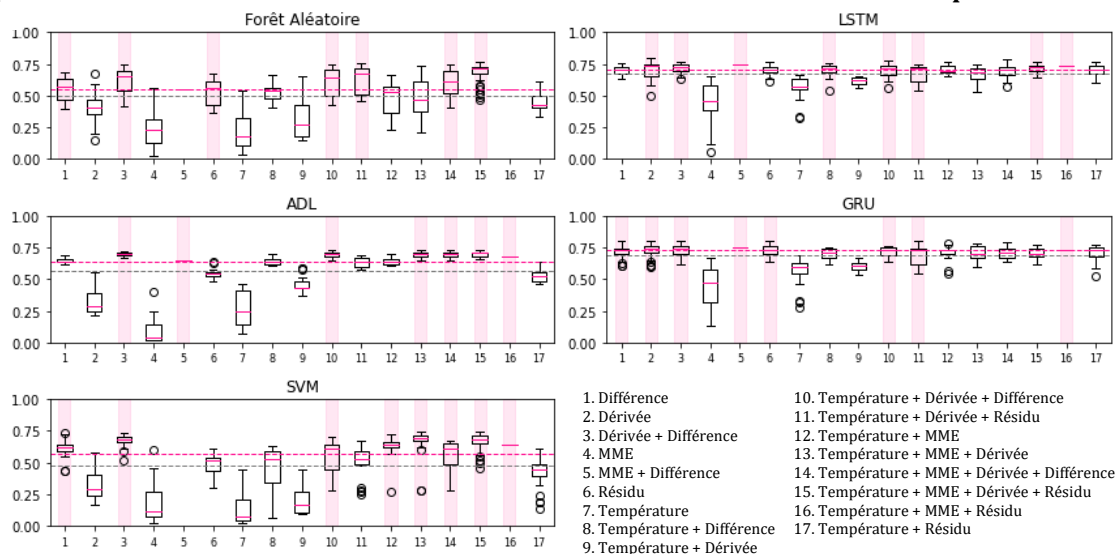
Pour évaluer l'influence des transformations et associations sur les performances de chaque modèle, l'ensemble des résultats provenant des différentes combinaisons des mesures d'ambiance intérieure sont utilisés : T_{int} , C_{int} , H_{int} , $T_{int} + C_{int}$, $T_{int} + H_{int}$ et $T_{int} + H_{int} + C_{int}$, ce qui équivaut à environ 500 combinaisons pour chaque modèle. Les informations relatives à la temporalité ont été exclues, d'où la présence d'un peu moins de 20 associations par variable d'ambiance intérieure. Les F_1 -scores pour les transformations, associations, variables et modèles sont présentés sous la forme de diagrammes en boîtes (appelés aussi boîtes à moustaches) dans la Figure 39. Les transformations et associations relatives à la température intérieure sont illustrées en (A). Les médianes de chaque boîte sont indiquées par un trait plein rose tandis que la médiane et la moyenne de l'ensemble des associations, pour chaque modèle, sont représentées par des traits discontinus rose et gris. Une représentation similaire est appliquée en employant une couleur orange pour les transformations et associations issues de la variable de concentration de CO_2 intérieure en (B), ou une couleur bleue pour celles provenant de l'humidité intérieure en (C). Les transformations et associations mises en évidence par une zone colorée représentent celles dont la médiane, pour un modèle donné, est supérieure à la médiane de l'ensemble des associations pour ce même modèle.

Les résultats montrent que les transformations et associations basées sur l'humidité donnent généralement les moins bons résultats médians, moyens et maximums pour tous les modèles. Cependant, les plus performantes semblent être celles qui utilisent uniquement la dérivée $_{H_{int}}$ (C.1), la différence $_{H_{int}}$ (C.2) ou les deux (C.3). Contrairement à la différence $_{H_{int}}$, qui s'avère être une transformation suffisante pour détecter les ouvertures de fenêtres, comme indiqué précédemment (voir 4.5.1.1), la dérivée $_{H_{int}}$ semble plus performante en combinaison avec d'autres variables.

En ce qui concerne les transformations et associations liées à la température, il convient de faire une distinction entre les modèles. Pour les modèles LSTM et GRU, l'utilisation seule de la dérivée $_{T_{int}}$ (A.2) semble suffisante alors que pour les modèles d'ADL et de SVM et les forêts aléatoires, elle doit être combinée avec d'autres transformations. Cela s'explique par le fonctionnement des techniques utilisées. La dérivée d'une mesure caractérise une ouverture de fenêtre par une succession de valeurs négatives (ouverture) suivie de valeurs positives (fermeture). Cependant, les valeurs de ces séquences ne sont pas suffisamment différentes des autres variations pour être discriminatoires pour des modèles qui se basent sur la séparation des données dans l'espace, comme les SVM et les ADL. En revanche, pour les modèles LSTM et GRU, qui traitent des séquences, la succession de ces valeurs agit comme un élément discriminant.

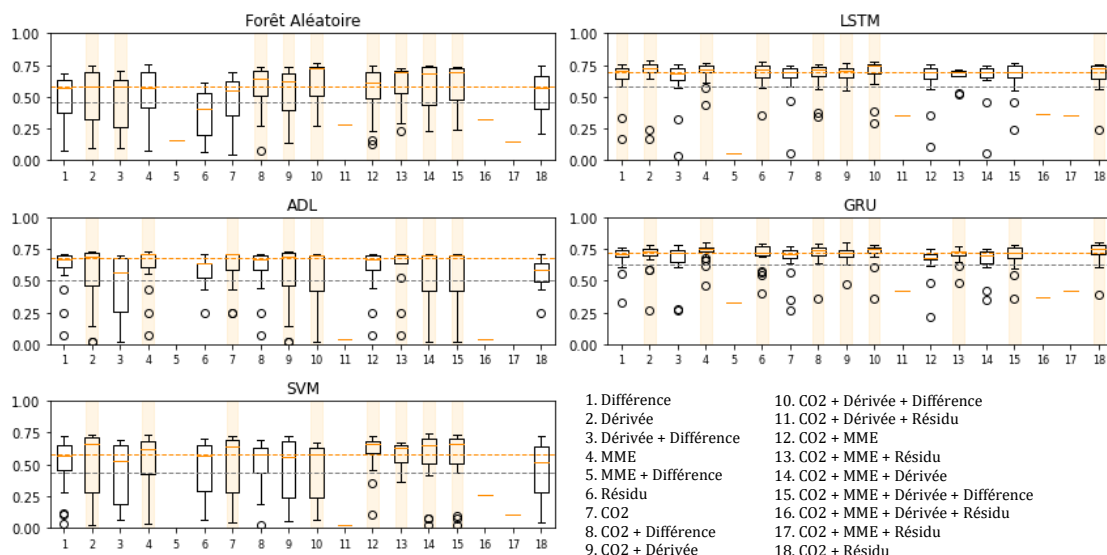
A

F1-score de toutes les transformations et associations basées sur la température



B

F1-score de toutes les transformations et associations basées sur la concentration de CO₂



C

F1-score de toutes les transformations et associations basées sur l'Humidité

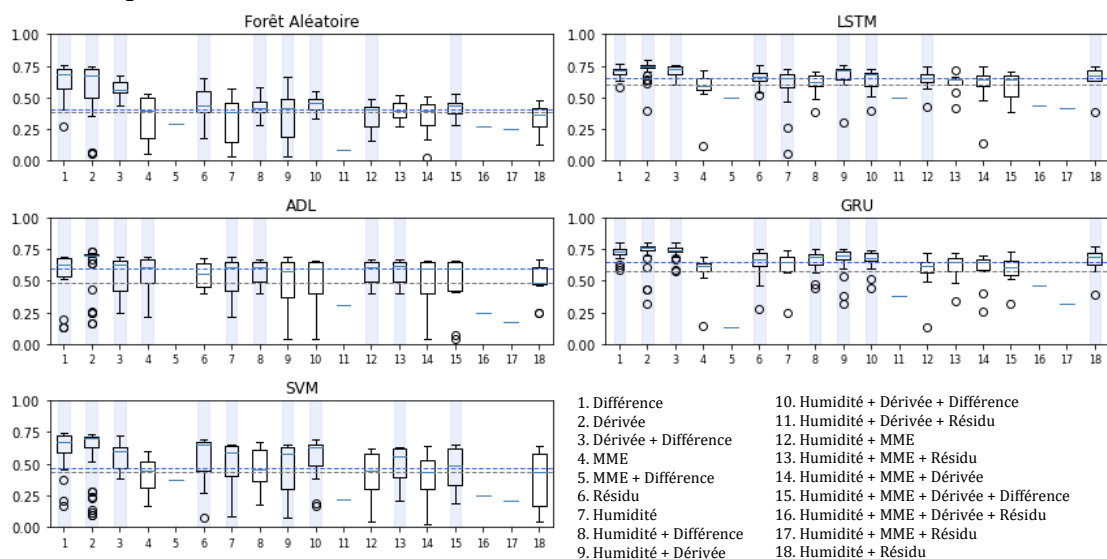


Figure 39 F1-scores pour chaque transformation, association, et combinaison de mesure d'ambiance pour un modèle donné

Pour les modèles SVM et ADL, un grand nombre d'associations de transformation de la température semblent être nécessaires pour obtenir de bons résultats, contrairement aux modèles de RNR, qui peuvent obtenir de bons résultats avec une seule association de transformation ou même aucune, telle que la dérivée_{Tint}, la différence_{Tint} ou le résidu_{Tint} (A.1, 2 et 6). Ces dernières transformations semblent particulièrement adaptées pour les SVM, les ADL, et les forêts aléatoires. Les performances supérieures des modèles SVM et ADL avec des transformations basées sur la DMM ou le résidu_{STL} renforcent nos conclusions précédentes sur la transformation de dérivée. Ces deux transformations, qui mesurent la distance par rapport à la tendance générale des observations avoisinantes, permettent de classer les données en fonction de leurs valeurs. Les valeurs les plus éloignées sont souvent identifiées comme des ouvertures de fenêtres.

Il est également intéressant de noter que les associations ou combinaisons de mesures qui se limitent à la température sans transformations ou avec lissage par MME, montrent les résultats les plus faibles, avec des médianes particulièrement basses, pour tous les modèles. Contrairement à nos observations précédentes, il n'y pas de consensus clair en ce qui concerne les associations de transformation basées sur la concentration de CO₂. Les associations comprenant la dérivée_{Cint}, le résidu_{Cint} ou la différence_{Cint} apparaissent comme efficaces. Cependant, contrairement aux autres variables, les mesures non modifiées ou lissées par MME semblent aider les modèles dans leur prise de décision. Ce résultat est cohérent compte tenu de la nature versatile des mesures de CO₂, qui reflètent mieux les ouvertures de fenêtres tout en étant plus sujettes au bruit que les autres données d'ambiance. Bien que la concentration de CO₂ seule ne permette pas de détecter les ouvertures, elle constitue un complément précieux. De plus, les associations comprenant la dérivée sont souvent mises en avant.

En résumé, il est préférable d'éviter d'utiliser uniquement les mesures de base sans associations ni transformations, ou seulement lissées, car elles donnent généralement de faibles résultats. Cependant, il existe une exception notable pour la mesure de concentration de CO₂, mais uniquement lorsqu'elle est combinée avec d'autres mesures. En revanche, la température intérieure est une base solide, même si elle nécessite une transformation. En effet, l'utilisation seule de la dérivée, du résidu_{STL} ou de la DMM semble suffisante pour produire de bons résultats dans la détection des ouvertures de fenêtre pour les modèles de RNR. Pour les modèles d'ADL, de SVM et de forêt aléatoire, il est préférable de privilégier plusieurs associations ou combinaisons de mesures.

4.5.2. Évaluation avec métriques additionnelles

Des métriques additionnelles, introduites dans la partie 4.3.1, apportent une interprétation plus approfondie sur les différences observées. Ces métriques incluent le

nombre de vraies et de fausses ouvertures (A), la durée totale (en heures) des ouvertures (B), et les scores moyens de précision pour les ouvertures et les fermetures (C). Ils sont présentés sous forme de diagrammes en boîte dans la Figure 40. Cette figure est construite en exploitant les mêmes vingt meilleurs modèles par combinaison de mesures utilisés dans le Tableau 12. Toutefois, elle se concentre sur les combinaisons de mesures qui ont donné les meilleurs résultats, à savoir T_{int} , $T_{int} + C_{int}$, $T_{int} + H_{int}$ et $T_{int} + H_{int} + C_{int}$, soit un total de 80 combinaisons au lieu des 470 initiales pour chaque modèle.

Comme observé précédemment avec le F_1 -score, nous remarquons à nouveau que les combinaisons de température et de CO_2 intérieures sont légèrement plus performantes que celles de température et d'humidité intérieures dans l'ensemble. Les combinaisons incluant la variable de CO_2 semblent globalement plus performantes, avec des scores médians ou maximaux supérieurs pour toutes les métriques considérées, ainsi que des résultats maximaux plus élevés pour la détection du nombre de vraies ouvertures (A.i) et les score d'ouvertures et de fermetures (C). Cette observation pourrait s'expliquer par la propension de la concentration de CO_2 à fluctuer de façon bien plus marquée que l'humidité, ce qui permet, avec une transformation adéquate, de mieux détecter ou définir les ouvertures. En outre, pour tous les modèles, à l'exception des GRU et des LSTM, l'utilisation d'une combinaison d'au moins deux mesures intérieures conduit à une nette amélioration des résultats par rapport à l'utilisation de la seule température intérieure, même si le score le plus élevé pour chaque métrique tend rarement à s'éloigner significativement quelle que soit la combinaison utilisée.

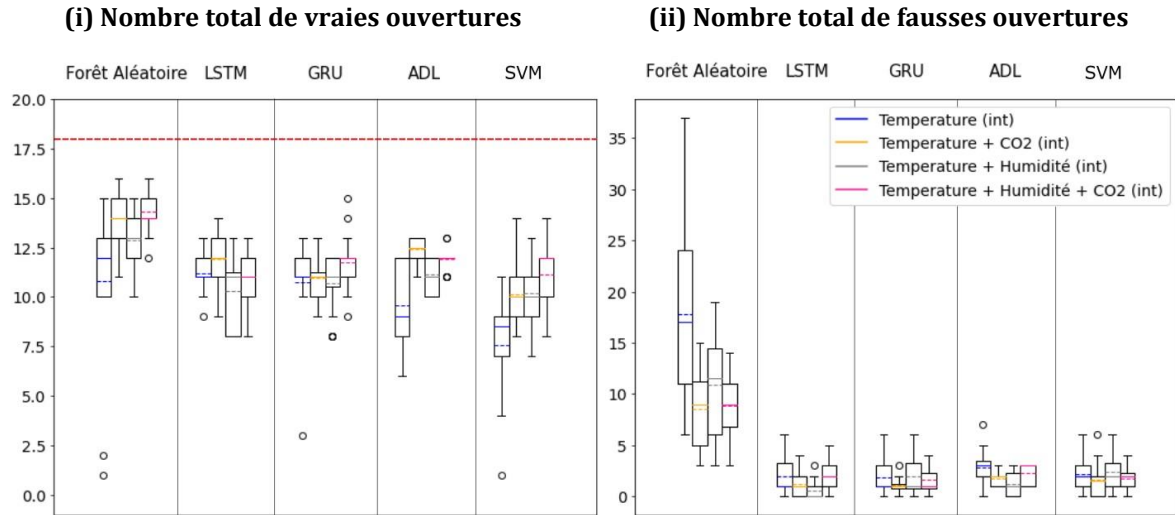
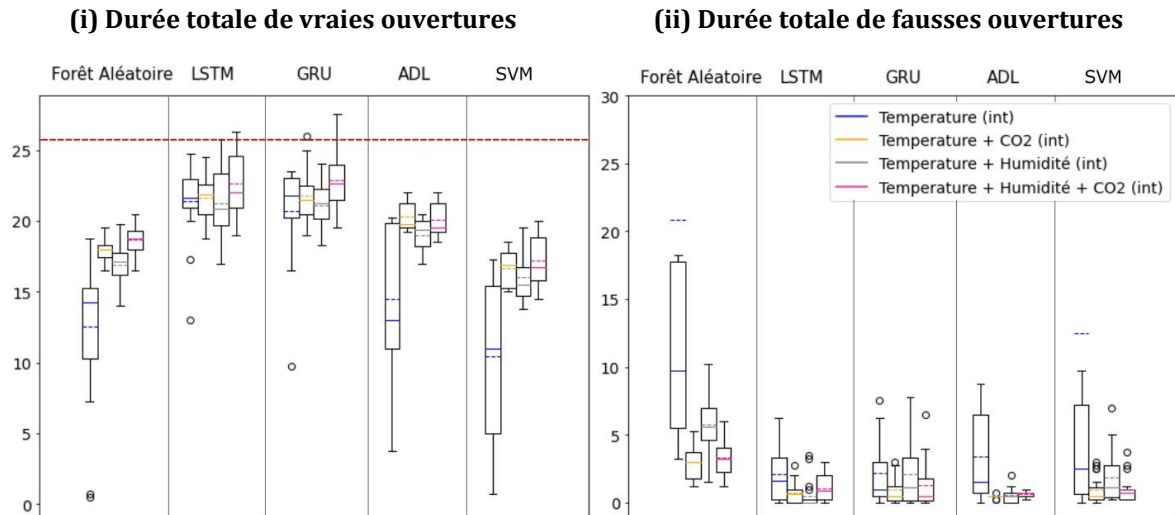
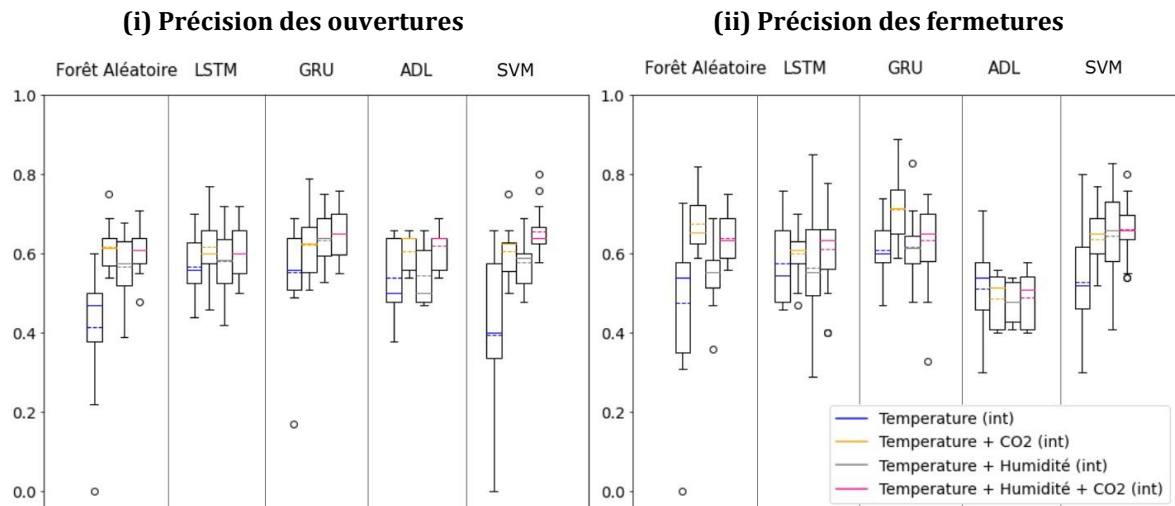
A**B****C**

Figure 40 Métriques additionnelles pour les 20 meilleurs modèles des combinaisons de température (T_{int}), d'humidité (H_{int}) et de concentration de CO₂ (CO_{2int}) intérieures.

Sur la base des F1-scores du Tableau 12, les GRU et les LSTM semblent produire des résultats similaires et suivent des tendances identiques. Cependant, la Figure 40 montre que, quelle que soit la combinaison, les modèles LSTM détectent légèrement plus d'ouvertures (A.i), bien que cela reste relativement modeste, alors que les modèles GRU semblent être plus précis dans la détection des états d'ouverture et de fermeture (C). Les modèles SVM et ADL semblent également fournir des résultats assez proches en matière de détection d'ouvertures, mais les SVM apparaissent comme moins aptes à détecter les vraies ouvertures (A.i), tandis que les ADL sont nettement moins performants pour évaluer les fermetures, puisqu'ils présentent les scores de précision les plus faibles (C.ii). Le modèle de forêt aléatoire semble être le plus inconstants mais aussi les plus sensibles aux variations dans les mesures, avec le plus grand nombre de vraies ouvertures et de fausses ouvertures détectées (A) parmi tous les modèles, ainsi qu'une durée totale de fausses ouvertures élevée (B.ii). La variabilité des résultats au sein d'une même boîte à moustaches semble révélatrice de l'impact des transformations et des associations réalisées pour chaque modèle ou combinaison de mesures. Par exemple, en examinant les 20 transformations et associations de température pour chaque modèle (en bleu) dans la Figure 40, nous observons des évolutions significatives dans la détection du nombre de vraies ouvertures (A.i) ou de la durée totale de vraies ouvertures (B.i). Cela confirme le constat précédent selon lequel les modèles dépendent largement des informations permettant de caractériser les comportements qu'ils cherchent à classer.

En conclusion, bien que tous les modèles de la Figure 40 détectent un nombre médian de vraies ouvertures de 10 à 13 sur les 18 existantes (représentées par une ligne pointillée rouge en A.i), ce résultat doit être nuancé. Comme expliqué dans la partie 4.4.2, plusieurs ouvertures ont peu ou pas d'impact sur l'environnement intérieur et sont donc plus difficiles, voire impossibles, à détecter. La Figure 41, qui recense toutes les ouvertures détectées par les 20 meilleurs modèles, illustre ce constat. Certaines ouvertures telles que l'O₆, l'O₁₁, l'O₁₅, l'O₁₇ et l'O₁₈ sont rarement, voire jamais, détectées par la majorité des modèles, car leur impact sur les mesures de température, d'humidité et de concentration de CO₂ intérieures est très faible et peu différenciable des observations sans ouverture de fenêtre. Ces impacts sont systématiquement caractérisés par des ouvertures de petites tailles entraînant des variations dans les données de (très) faible intensité. La forêt aléatoire semble être le seul modèle capable de les détecter, mais le nombre de fausses ouvertures présentées dans Figure 40 (A.ii) remet en question la fiabilité de cette détection.

Cependant, tous les modèles tendent à détecter en moyenne de 17 à 22 heures d'ouverture sur les 25,75 heures existantes (représentées par une ligne rouge en pointillé dans la Figure 40 B.i), avec une moyenne de 30 minutes à 2 heures de fausses ouvertures. De plus, les modèles LSTM et GRU montrent des résultats satisfaisants en utilisant

uniquement la température intérieure, en détectant en moyenne 84 à 88 % du temps d'ouverture (soit 21 à 22 heures sur 25,75) pour une moyenne de 1,5 à 2 heures de fausses ouvertures. Ces résultats tendent à montrer que la plupart des ouvertures ayant un impact sont détectées au cours de cette période de test d'un mois. Cependant, le principal domaine à améliorer semble résider dans la précision des détectations d'ouvertures et de fermetures, qui semblent toujours être légèrement en avance ou en retard d'un pas de temps, avec un score moyen de 0,50 ou 0,60.

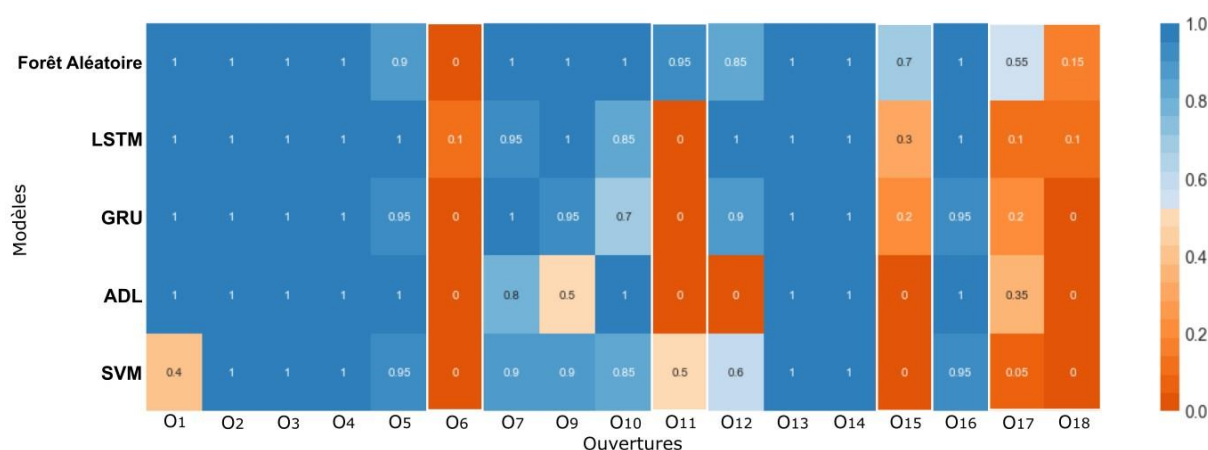


Figure 41 Carte thermique des ratios d'ouvertures détectées par les 20 meilleurs modèles, toutes combinaisons de mesures confondues. Les ouvertures sont numérotées de 1 à 18 et sont mentionnées en abscisse sous l'abréviation « O ». Par exemple, la cinquième ouverture est notée « O₅ ». L'O₈ n'a jamais été détectée et n'est donc pas affichée.

4.5.3. Synthèse appliquée à la détection d'ouvertures des fenêtres

Plusieurs combinaisons, associations et transformations des mesures ont été réalisées pour entraîner différents modèles d'apprentissage automatique et évaluer leur efficacité dans la détection de l'état d'ouverture des fenêtres. Cependant, malgré ces avancées dans un contexte d'application spécifique, plusieurs limitations subsistent. La plupart des transformations appliquées aux mesures (résid_{STL}, dérivée ou DMM) sont conçues pour refléter l'impact de l'ouverture d'une fenêtre entre deux environnements aux caractéristiques différentes, en l'occurrence l'intérieur et l'extérieur. Bien qu'elles aient permis d'améliorer et de stabiliser les performances des modèles, elles peuvent ne pas être adaptées à d'autres climats ou saisons et nécessitent une évaluation approfondie dans d'autres contextes. Cependant, ce cas spécifique a d'illustré la pertinence d'appliquer des transformations permettant d'extraire certaines informations pertinentes des mesures, en exploitant une connaissance approfondie des données, de leurs caractéristiques et spécificités, démontrant ainsi l'importance de l'expertise métier dans la résolution de problèmes de détection.

Il est apparu que, pour les données non transformées, les résultats obtenus dans ces travaux étaient corrélés avec ceux observés dans d'autres études, notamment en ce qui concerne l'importance des températures intérieures et extérieures (Dai et al., 2020). Cependant, cette tendance change lorsque des transformations sont appliquées aux mesures intérieures, et l'ajout de variables extérieures peut détériorer les résultats. En outre, il convient de noter, comme l'ont montré Markovic et al. (2019), que l'humidité de l'air intérieur (et extérieur, dans notre cas) semble avoir un impact faible sur les performances des modèles.

Concernant la sélection de mesures, les modèles LSTM et GRU obtiennent des résultats satisfaisants avec la seule utilisation des mesures de la température intérieure, bien que l'ajout de la concentration de CO₂ intérieure semble stabiliser et améliorer légèrement leurs résultats. Pour les modèles de SVM, d'ADL et de forêts aléatoires, l'utilisation minimale de la température intérieure et de la concentration de CO₂ tend à être recommandée même si une légère amélioration des résultats peut être observée en ajoutant la mesure de l'humidité intérieure.

Les métriques complémentaires introduites dans cette étude ont offert une perspective différente sur les performances des modèles en ce qui concerne la détection des états d'ouverture. Ces métriques permettent d'apporter une expertise métier aux résultats des modèles et pourraient faciliter la sélection du modèle le plus adapté aux besoins d'un projet, par exemple, en privilégiant le nombre d'ouvertures détectées par rapport à leur précision, ou en comparant de manière plus pertinente les résultats entre différentes études différentes. Cependant, contrairement aux métriques couramment utilisées et adaptées aux tâches de classification binaires, telles que le F₁-score, leur mise en œuvre est complexe et nécessite l'examen simultané de six métriques différentes.

4.6. Conclusion

Ce chapitre a permis d'appliquer différentes techniques supervisées de classification, à savoir les *Long Short-Term Memory* (LSTM), les *Gated Recurrent Unit* (GRU), les Séparateurs à Vaste Marge (SVM), les Analyses Discriminantes Linéaires (ADL) et les forêts aléatoires, pour les comparer dans une tâche similaire à la détection d'anomalies. Les ouvertures de fenêtre peuvent être considérées comme des anomalies collectives, dont la nature semble varier selon les données considérées. Ces travaux ont évalué un processus de détection basé sur trois mesures différentes : la température intérieure, l'humidité relative intérieure et la concentration de CO₂ intérieure. Le contexte de l'étude a été modulé en variant les combinaisons de mesures parmi un ensemble de cinq mesures

et en appliquant des transformations visant à mieux caractériser les dynamiques (ou motifs) spécifiques à détecter.

Dans l'ensemble, quelle que soit la technique de classification utilisée et les mesures considérées, il semble préférable de travailler avec un contexte composé d'au moins deux variables, incluant *a minima* la température intérieure, ce qui caractérise ce problème de détection comme principalement multivarié. Cependant, il a été démontré que l'application de transformations adéquates pouvait permettre, selon les modèles et les transformations utilisées, d'effectuer une détection dans un contexte univarié.

En ce qui concerne les techniques de classification, bien que les résultats ne soient pas diamétralement opposés, plusieurs distinctions ont pu être observées. Tout d'abord, le choix des techniques influence le choix des transformations. Selon les techniques utilisées, certaines transformations se révèlent plus adaptées pour caractériser les données que d'autres. L'exemple le plus parlant est apparu dans les performances liées à l'utilisation de la dérivée.

En général, les modèles de Réseaux de Neurones Récurrents (RNR) semblent être moins sensibles au choix des transformations ou des associations, mais plus sensibles aux choix des variables par rapport aux SVM, aux ADL et aux forêts aléatoires. Cependant, l'utilisation des RNR nécessite une préparation plus poussée ainsi que la prise en compte de davantage de paramètres, ce qui se traduit par un investissement en temps et en matériel plus conséquent. En revanche, l'emploi des RNR permet d'améliorer les performances de détection d'environ 10 % par rapport aux autres modèles, puisque les meilleures performances obtenues par les modèles LSTM et GRU avoisinent un F1-score de 0,80 contrairement aux autres modèles qui plafonnent autour d'un score de 0,73. Ils se démarquent notamment lorsqu'ils sont appliqués à des mesures de température intérieure transformées, atteignant leurs performances presque maximales, avec un F1-score de 0,77 à 0,75, sans nécessiter l'ajout de variables complémentaires. D'une manière générale, et pour l'ensemble des modèles, l'application de transformation adaptées permet d'améliorer sensiblement les performances de détection par rapport à l'utilisation des mesures sans transformation.

De plus, deux éléments ont été identifiés comme ayant un impact considérable sur les performances des techniques employées. Le premier concerne la dynamique des données. Les changements dans les profils d'occupation sur de longues périodes, constatés lors de la phase de test, ont eu un impact significatif sur les performances des modèles. Cependant, l'application de transformations et d'associations a permis d'atténuer cet impact. Il est important de noter que de tels événements restent rares, et garantir leur présence dans l'ensemble d'entraînement n'est pas une solution simple, voire même nécessaire. Ces événements rares soulignent l'importance de prendre en compte, *a priori*,

certaines dynamiques occasionnelles dans le but de développer des modèles robustes, même si cela pose des questions de pertinence en raison de leur rareté. Il serait envisageable de ne pas appliquer le modèle dans certaines conditions spécifiques, d'entraîner plusieurs modèles en fonction des cas, ou même en fonction des mesures disponibles. Enfin, le deuxième élément ayant un effet considérable sur les performances des techniques est en partie lié à ce qui a été qualifié de dynamique des anomalies dans le chapitre précédent (*cf.* partie 3.3.4). Selon l'intensité de leur impact sur les données, certaines ouvertures de fenêtres peuvent échapper à la détection. L'évaluation de cet impact semble ainsi être un axe intéressant à explorer pour la suite de ces travaux.

Ce chapitre a examiné un cas spécifique, dans une optique de preuve de concept, d'un problème caractérisé par une anomalie présentant une dynamique unique. Le chapitre 5 se concentrera sur la détection de diverses anomalies, chacune ayant des dynamiques et des impacts variés sur différentes données du bâtiment. Bien que nous ayons pu présenter un exemple d'utilisation de techniques de détection d'anomalies supervisées, la rareté des données étiquetées, notamment celles représentant des anomalies, limite leur applicabilité dans des scénarios réels. Par conséquent, l'application de techniques de détection semi-supervisées et non supervisées sera privilégiée dans la suite.

Chapitre 5

Détection semi-supervisée et non supervisée d'anomalies dans des mesures du bâtiment

Ce chapitre vise à appliquer différentes techniques d'apprentissage semi-supervisé et non supervisé à plusieurs profils d'anomalies, générées artificiellement pour évaluer les performances des modèles de détection. Ces anomalies varient en dynamique, taille et intensité pour tester la capacité des modèles à détecter des situations difficiles. Nous explorerons également différentes combinaisons, associations et transformations de données, ainsi que plusieurs tailles des fenêtres et approches de détection pour évaluer leur impact sur les performances et en dégager des tendances. Le choix des techniques de détection est d'abord présenté dans la section 5.2. Puis le processus de génération des anomalies de même que les modèles, leurs constructions et leur processus d'évaluation sont décrits dans la section 5.3. Les données collectées et les transformations utilisées ainsi que le protocole de génération de jeux de test sont détaillés dans la section 5.4. Enfin, plusieurs résultats sont présentés dans la section 5.5 et 5.6, lesquelles sont suivies d'une synthèse en 5.7.

« Je ne peux le porter pour vous... Mais je peux vous porter vous. »

Extrait traduit du, Seigneur des Anneaux : le retour du roi, de J.R.R. Tolkien

Sommaire du chapitre 5

5.1. Contexte et démarche	125
5.2. État de l'art pour le choix des techniques	126
5.3. Méthodologie de l'expérimentation	128
5.3.1. La génération d'anomalies.....	128
5.3.2. Modèles d'apprentissage automatique	133
5.3.3. Configuration et entraînement des modèles	142
5.3.4. Processus d'évaluation des modèles.....	145
5.4. Description, prétraitement des données et génération des ensembles d'évaluation	151
5.4.1. Collecte et préparation des données.....	151
5.4.2. Génération des ensembles d'évaluation anormaux	153
5.4.3. Transformation et association des données.....	155
5.5. Résultats	156
5.5.1. Contexte sur l'entraînement et l'exploitation des modèles	157
5.5.2. Le cas des mesures de température d'air intérieur.....	160
5.5.3. Le cas des mesures d'humidité relative d'air intérieur	165
5.5.4. Les cas des mesures de concentration de CO ₂ intérieure.....	166
5.5.5. Le cas des mesures de consommation électrique	172
5.5.6. Observations communes à l'ensemble des variables	175
5.5.7. Synthèse des résultats	183
5.6. Exemple d'application sur des anomalies issues de mesures	186
5.6.1. Contexte	186
5.6.2. Visualisation des résultats.....	187
5.7. Conclusion.....	191

5.1. Contexte et démarche

Le chapitre précédent a examiné l'influence des combinaisons, associations et transformations de données sur la détection d'anomalies. Il a également soulevé des interrogations sur la quantification de l'impact des anomalies. Ce chapitre élargit le champ d'investigation pour mettre en lumière des observations et des tendances liées aux données, aux anomalies et aux modèles, susceptibles d'être utiles pour de futures études ou applications.

Pour ce faire, diverses techniques de détection d'anomalies sont évaluées sur un ensemble de problématiques relatives aux données et aux anomalies. Ces problématiques peuvent être formulées en associant différentes données à des anomalies de tailles et de natures variées (conformément au chapitre 3). Les données disponibles pour ces travaux proviennent de mesures de bâtiments connectés, notamment la consommation électrique cumulée, la température d'air, l'humidité relative et la concentration de CO₂ intérieures.

La littérature révèle que la plupart des anomalies étudiées sont de type « pic » (aussi appelés *extremum*), que nous avons définies précédemment comme des anomalies ponctuelles. Les anomalies collectives, notamment de tailles variables, et les anomalies globales, sont moins fréquemment abordées (Blázquez-García et al., 2022). Bien que certains travaux se penchent sur les anomalies collectives, à l'instar de Sharma et al. (2010) ou Liu et al. (2020), elles tendent souvent à être tellement éloignées des observations normales qu'elles ressemblent davantage à des anomalies collectives indépendantes, qui sont généralement plus faciles à détecter (Emmott et al., 2016).

Dans notre typologie des anomalies présentée au chapitre 3 (3.3.4), nous avons proposé plusieurs niveaux pour définir les défis liés à la détection d'anomalies. Dans ce cas, les anomalies collectives de nature contextuelle, qu'elles soient univariées ou multivariées, apparaissent comme un défi de détection plus complexe. Comme ces travaux portent sur les données du bâtiment et visent à évaluer l'influence présumée de la dynamique des types d'anomalies formulée dans le chapitre 3, nous allons examiner l'ensemble des anomalies rencontrées dans des mesures réelles, telles que recensées dans la partie 3.3.2. Nous allons ainsi appliquer six types d'anomalies, dont une ponctuelle avec les « pics » et cinq collectives avec les « bruits », les « écarts », les « écarts plats », les « plateaux » et les « changements de niveaux ». Pour évaluer un éventail aussi large de types d'anomalies, nous allons les générer artificiellement, créant ainsi un cadre contrôlé permettant de quantifier les performances des modèles sur des anomalies de nature contextuelle.

À l'image du chapitre 4, nous allons appliquer uniquement les modèles semi-supervisés et non supervisés aux jeux de données contaminées. Le but est d'évaluer leurs performances et, si possible, observer des tendances. Ces méthodes d'entraînement sont pertinentes car elles ne nécessitent peu ou pas d'étiquetage des données et figurent parmi celles les plus couramment utilisées (Himeur et al., 2021b). De plus, certains travaux, comme ceux de Schmidl et al. (2022), constatent peu de différences de performances entre ces deux approches et l'approche supervisée. Cependant, il est nécessaire de considérer des techniques ayant des fonctionnements variés pour pouvoir comparer et évaluer l'impact de différentes méthodes.

Dans un premier temps ; nous détaillerons le choix des modèles sélectionnés pour cette étude. Dans un second temps, nous présenterons la méthodologie de l'expérimentation, en décrivant le modèle de génération d'anomalies mis en place et en introduisant les différents modèles sélectionnés. Nous expliquerons aussi leur processus d'entraînement et les différentes approches de détection d'anomalies suivies. Les données utilisées pour l'entraînement et les tests des modèles seront présentées, suivies du protocole de contamination appliqué pour générer les jeux de données anormaux. Enfin, nous exposerons les résultats obtenus.

5.2. État de l'art pour le choix des techniques

Le choix des techniques pour la détection d'anomalies dans les séries temporelles peut s'appuyer sur plusieurs approches et méthodes semi-supervisées ainsi que non supervisées. Dans une étude similaire, mais plus étendue, réalisée par Schmidl et al. (2022) sur les séries temporelles, diverses techniques ont été explorées et regroupées en plusieurs familles de méthodes de détection avec principalement celles dites : de prédiction, de reconstruction, de distance et de forêt d'isolation.

1. Les méthodes de prédiction, référencées sous l'appellation de « méthodes de régression » dans la partie 2.4.3.3, sont généralement employées suivant une approche semi-supervisée pour anticiper les observations futures à partir des observations passées. Elles se divisent en deux catégories : univariées et multivariées. Parmi les méthodes univariées, l'*Auto-Regressive Moving Average* (ARIMA) est largement utilisée, car elle permet de détecter plusieurs types d'anomalies avec succès, comme les « pics », les « écarts » ou les « changements de niveau » (Kraiem, 2021). D'autres techniques de régression linéaire plus simples peuvent également être incluses dans cette catégorie (Foulon, 2020; Liu et al., 2020). Dans l'approche multivariée, les modèles les plus employés sont les réseaux de neurones denses classiques (Blázquez-García et al., 2022) ou les

réseaux récurrents basés sur les *Long Short-Term Memory* (LSTM) (Chalapathy and Chawla, 2019).

2. Les méthodes de reconstruction sont majoritairement semi-supervisées et réduisent la dimension des observations pour les reconstruire par la suite (cf. 2.4.3.7). Comme ces modèles sont formés pour généraliser des données normales, les données anormales, étant rares, elles ne peuvent supposément être reconstruites. Cependant, il est important de noter que ces modèles généralisent si bien les données qu'ils peuvent parfois tout de même reproduire les anomalies (Angiulli et al., 2022). Les modèles d'*Encoder-Decoder* (ED) et les *Autoencoder* (AE) sont parmi les plus populaires (Blázquez-García et al., 2022; Chalapathy and Chawla, 2019). Ils sont construits à l'aide de réseaux de neurones et présentent de nombreuses variantes (Chalapathy and Chawla, 2019). Par exemple, Liu et al. (2020) emploient des AE, des ED, des réseaux denses et des LSTM prédictifs pour détecter des anomalies dans des mesures de concentration de CO₂ provenant d'un mur végétal.
3. Les méthodes de distances et (4) les méthodes de forêt d'isolation comprennent celles classées dans les familles basée sur le voisinage (cf. 2.4.3.5) et sur le *clustering* (cf. 2.4.3.2). Les méthodes les plus représentatives de ces familles incluent les *K plus proches voisins* (KNN), le *Local Outlier Factor* (LOF), l'*Isolation Forest* (IF) et le *K-means*. Habituellement, ces méthodes s'emploient pour évaluer l'anormalité des observations individuellement, bien qu'elles puissent également servir à évaluer des similarités ou les différences entre les séquences (Blázquez-García et al., 2022; Schmidl et al., 2022). Cette approche peut être mise en œuvre en comparant, à l'aide de mesures de distance, une séquence d'observations jugée normale à un ensemble d'autres séquences, ou en comparant toutes les séquences entre elles (Benkabou, 2018).

Plus récemment, nous avons assisté à la montée en puissance des *Generative Adversarial Network* (GAN) (Goodfellow et al., 2014). Ces modèles semi-supervisés, initialement conçus pour générer des données, ont également trouvé leur application dans la détection d'anomalies. Ils sont constitués d'un modèle générateur et d'un modèle discriminateur qui sont en « compétition » l'un contre l'autre. Le générateur tente de créer des données si réalistes que le discriminateur ne peut pas les distinguer des données réelles ou artificielles. Dans le contexte de la détection d'anomalies, le discriminateur est employé pour estimer la probabilité qu'une séquence d'observations soit anormale. Cependant, il convient de noter que ces modèles sont coûteux à entraîner et semblent mieux adaptés aux données complexes de grande dimension (Chalapathy and Chawla, 2019). Dans certaines applications, des techniques plus simples à mettre en œuvre, comme les KNN, peuvent même offrir de meilleures performances (Škvára et al., 2018).

Dans notre cas, compte tenu de la tâche de détection envisagée, qui concerne à la fois des anomalies collectives représentées par des séquences d'observations anormales et des anomalies ponctuelles, une approche diversifiée semble préférable. Par conséquent, nous évaluerons plusieurs méthodes, y compris des méthodes de prédiction, de reconstruction et de voisinage. Pour ce faire, nous utiliserons trois variantes de modèles prédictifs : la méthode **ARIMA**, des réseaux de neurones denses prédictifs, notés **Dense-p**, ainsi que des réseaux de neurones récurrents LSTM prédictifs, notés **LSTMP-p**. Pour les méthodes de reconstruction, des **Autoencoder** et des modèles **Encoder-Decoder**. Enfin, pour les méthodes de voisinage, nous utiliserons les modèles d'**IF**, de **KNN** et de **LOF**. Les variantes spécifiques de chaque modèle seront introduites plus en détails dans la section 5.3.2.

5.3. Méthodologie de l'expérimentation

Cette section présente le modèle de génération d'anomalies que nous avons développé et décrit en détail le processus de construction des différents types d'anomalies que nous avons choisis pour ces travaux. Nous introduirons également les modèles utilisés et fourniront des informations supplémentaire sur leur construction et leur entraînement. Enfin, nous aborderons les méthodes que nous avons sélectionnées pour évaluer et comparer les performances des modèles.

5.3.1. La génération d'anomalies

Pour évaluer la détection d'anomalies sur différents profils d'anomalies et des mesures variées, nous avons créé un modèle de génération d'anomalies. Ce modèle permet la contamination artificielle de jeux de données basés sur des mesures réelles avec cinq types différents d'anomalies, comme expliqué dans la partie 3.3.2 : les pics, les valeurs constantes, les bruits, les écarts et les changements de niveau. Ce modèle est conçu pour adapter ces anomalies aux caractéristiques des données sur lesquelles elles sont appliquées. De plus, il permet de générer ces anomalies à deux niveaux d'intensité différents, à savoir *faible* et *fort*. Les anomalies de faible intensité ont un impact minime sur les données, contrairement aux anomalies forte intensité. Cette distinction en intensité vise à évaluer les performances des modèles en tenant compte de la capacité des anomalies à se confondre ou à se distinguer des données normales. Cette approche s'inspire des réflexions amorcées dans le chapitre 4 et est en corrélation avec des travaux antérieurs, tels que ceux d'Emmott et al. (2016) ou de Sharma et al. (2010). Les anomalies de faible intensité sont considérées comme étant plus proches des observations normales que celles de forte intensité, ce qui les rend potentiellement plus difficiles à détecter. Ainsi,

l'ajout de cette échelle d'intensité vise à mettre en lumière la sensibilité des modèles utilisés à cette caractéristique.

Les anomalies artificiellement produites par notre modèle sont décrites ci-après. Leur formulation résulte de l'expertise métier ainsi que de l'examen de plusieurs travaux, dont ceux de Sharma et al. (2010), Noshad et al. (2019) et Kullaa (2013). De plus, elles sont toutes conçues pour respecter les définitions des anomalies contextuelles telles qu'établies dans la partie 3.3.4 du chapitre 3. Pour la suite de cette partie, notons $X = \{x_1, x_2, \dots, x_M\}$, une variable saine composée de M ($\in \mathbb{N}$) observations normales notées x_i (avec $1 < i < M$), et X^* , une variable contaminée comprenant à la fois des observations normales et anormales (x_i^*).

- **Les pics.** Ils correspondent à une unique observation dont la valeur s'écarte significativement de celle des observations avoisinantes, pouvant être soit positifs ou négatifs. Pour les introduire, nous choisissons aléatoirement une observation x_i parmi l'ensemble des observations de X , selon une distribution uniforme. Ensuite, nous la remplaçons par une nouvelle observation anormale notée x_i^* . La valeur de cette observation est choisie aléatoirement entre deux bornes différentes, selon l'intensité renseignée.

Si l'intensité est faible, la valeur est sélectionnée conformément aux bornes présentées dans l'équation (5), où σ_x représente l'écart type de la variable originale X . Si l'intensité est forte, elle est choisie de manière aléatoire entre les bornes de l'équation (6), utilisant respectivement les valeurs maximales et minimales, notées avec max_x et min_x . Des exemples d'anomalies artificielles de type pic, de différentes intensités, sont illustrés dans la Figure 42.

$$\text{Pic faible :} \quad x_i^* \in \left[x_i + \frac{1}{2} \sigma_x ; x_i + \frac{3}{2} \sigma_x \right] \cup \left[x_i - \frac{3}{2} \sigma_x ; x_i - \frac{1}{2} \sigma_x \right] \quad (5)$$

$$\text{Pic fort :} \quad x_i^* \in \left[x_i + \frac{3}{2} \sigma_x ; max_x \right] \cup \left[min_x ; x_i - \frac{3}{2} \sigma_x \right] \quad (6)$$

- **Les bruits.** Ils correspondent à des anomalies impliquant plusieurs observations qui présentent une dispersion plus importante que celle attendue ou constatée dans l'ensemble des observations de la variable. Pour les introduire, nous choisissons aléatoirement une séquence d'observations de taille $N > 1$ parmi l'ensemble des observations de X , en utilisant une loi uniforme. Cette séquence d'observations est notée $S_{X,i} = \{x_i, x_{i+1}, \dots, x_{i+N-1}\}$. En parallèle, nous générons un segment de bruit de même taille, $S_{B,i} = \{b_i, b_{i+1}, \dots, b_{i+N-1}\}$, en sélectionnant des valeurs b_i de manière aléatoire et uniforme entre deux bornes différentes, en fonction de l'intensité renseignée. Ce segment est créé de sorte à ce que sa moyenne soit nulle. Ensuite, nous l'ajoutons à $S_{X,i}$, et la séquence résultante, qui est anormale, est notée $S_{X,i}^*$.

Si l'intensité est faible, le segment de bruit est généré conformément aux bornes présentées dans l'équation (7), où σ_X représente l'écart type de X . Si l'intensité est forte, le bruit est alors choisi aléatoirement entre les bornes de l'équation (8). Les bornes sont maintenues faibles à dessein, puisque cette anomalie se caractérise avant tout par une séquence erratique d'observations, ce qui entraîne une variance élevée (Sharma et al., 2010). L'emploi de bornes plus importantes pourrait introduire un biais en ajoutant davantage d'observations anormales ayant plus de profils de pics successifs que de bruit. Un exemple d'anomalie artificielle de type bruit, de faible intensité, est illustré dans la Figure 42.

$$\text{Bruit faible :} \quad S_{X,i}^* = \underbrace{\{x_i + b_i, \dots, x_{i+N-1} + b_{i+N-1}\}}_{S_{X,i} + S_{B,i}} \text{ avec } b_i \in \left[-\frac{1}{4} \sigma_X; \frac{1}{4} \sigma_X\right] \quad (7)$$

$$\text{Bruit fort :} \quad S_{X,i}^* = \underbrace{\{x_i + b_i, \dots, x_{i+N-1} + b_{i+N-1}\}}_{S_{X,i} + S_{B,i}} \text{ avec } b_i \in \left[-\frac{1}{3} \sigma_X; \frac{1}{3} \sigma_X\right] \quad (8)$$

- **Les écarts.** Ils correspondent à des anomalies impliquant plusieurs observations dont les valeurs sont brusquement modifiées d'une quantité fixe, soit positivement, soit négativement. Ces modifications peuvent être réversibles, c'est-à-dire que les valeurs peuvent revenir à la normale rapidement ou progressivement. Dans ce cas d'application spécifique, nous ne considérons que le premier scénario. Pour introduire ce type d'anomalie, nous choisissons aléatoirement une séquence d'observations de taille $N > 1$ parmi l'ensemble des observations de X , en utilisant une distribution uniforme. Cette succession d'observations est notée $S_{X,i} = \{x_i, x_{i+1}, \dots, x_{i+N-1}\}$. Parallèlement, nous générons un segment de gain de même taille, $S_{G,i}$, en sélectionnant également de manière aléatoire une unique valeur de gain g entre deux bornes différentes, selon l'intensité renseignée. Ensuite, nous ajoutons ce segment à $S_{X,i}$, et la séquence résultante, qui est anormale est notée $S_{X,i}^*$.

Si l'intensité est faible, le segment de gain est généré conformément aux bornes présentées dans l'équation (9), où σ_X représente l'écart type de X . Si l'intensité est forte, le bruit est alors choisi aléatoirement entre les bornes de l'équation (10). Il est important de noter que les valeurs des observations de $S_{X,i}^*$ ne peuvent dépasser la valeur maximale ni la valeur minimale de X . Si cela se produit, une nouvelle valeur de gain est sélectionnée récursivement jusqu'à ce que la condition soit respectée, ou jusqu'à ce qu'un critère d'arrêt soit atteint. Si le critère d'arrêt est atteint, une nouvelle séquence d'observations de même taille est choisie. Un exemple d'anomalie artificielle de type écart, de faible intensité, est présenté dans la Figure 42.

$$\text{Écart faible :} \quad S_{X,i}^* = \underbrace{\{x_i + g, \dots, x_{i+N-1} + g\}}_{S_{X,i} + S_{G,i}} \text{ avec } g \in \left[\frac{1}{2} \sigma_X; \sigma_X \right] \cup \left[-\frac{1}{2} \sigma_X; -\sigma_X\right] \quad (9)$$

$$\text{Écart fort : } S_{X,i}^* = \underbrace{\{x_i + g, \dots, x_{i+N-1} + g\}}_{S_{X,i} + S_{G,i}} \text{ avec } g \in [\sigma_x; 3\sigma_x] \cup]-\sigma_x; -3\sigma_x] \quad (10)$$

- **Les valeurs constantes.** Elles correspondent à une séquence d'observations ayant toutes la même valeur. Il est possible de les aborder de deux manières différentes : soit comme un écart où toutes les observations sont identiques, soit comme une séquence où toutes les observations sont égales à la première observation. Dans le premier cas, cette anomalie est appelée, **écart plat**, tandis que dans le second cas, elle est appelée, **plateau**. Pour créer l'anomalie de type plateau, nous choisissons aléatoirement une séquence d'observations de taille $N > 1$ parmi l'ensemble des observations de X , en utilisant une distribution uniforme. Cette séquence d'observations est notée $S_{X,i} = \{x_i, x_{i+1}, \dots, x_{i+N-1}\}$. Ensuite, toutes les observations de cette séquence sont modifiées pour qu'elles aient la même valeur que la première observation x_i . La nouvelle séquence, notée $S_{X,i}^*$, est ainsi composée de N observations identiques à x_i et est utilisée pour remplacer $S_{X,i}$. Dans ce cas, l'intensité n'est pas prise en compte.

Pour l'anomalie de type écart plat, l'approche est similaire à celle de l'anomalie de type écart. La seule différence réside dans le remplacement de toutes les valeurs de la séquence $S_{X,i}$ par la première valeur x_i de cette séquence. À titre d'exemple, plusieurs représentations d'anomalies artificielles écart plat de faible intensité, et plateau, sont illustrées dans la Figure 42.

- **Les changements de niveaux.** Ils correspondent à des variations brusques et permanentes des observations, pouvant être soit additives soit soustractives. Ces variations peuvent affecter une ou plusieurs observations, mais se caractérisent principalement par une variation brutale (Foorthuis, 2021). Pour introduire cette anomalie, une séquence d'observations de taille $N \geq 1$ est choisie aléatoirement parmi l'ensemble des observations de X , en utilisant une distribution uniforme. Cette séquence d'observations est notée $S_{X,i} = \{x_i, x_{i+1}, \dots, x_{i+N-1}\}$. Cette séquence est considérée comme une coupure qui sépare X en deux séquences différentes. La première séquence est constituée de toutes les observations antérieures à la coupure et est notée $X' = \{x_0, x_1, \dots, x_{i-1}\}$. La deuxième séquence est composée de toutes les observations postérieures à la coupure et est notée $X'' = \{x_{i+N}, x_{i+N+1}, \dots, x_M\}$. Ensuite, un gain g est sélectionné de manière aléatoire, en utilisant une loi uniforme, entre deux bornes différentes selon l'intensité renseignée. Ce gain est ensuite ajouté à l'ensemble des observations de X'' , ce qui donne naissance à la séquence anormale X''^* .

Si l'intensité est faible, le gain est généré conformément aux bornes présentées dans l'équation (11), où σ_X représente l'écart type de X . Si l'intensité est forte, le gain est choisi aléatoirement entre les bornes de l'équation (12). Le segment $S_{X,i}$ est également modifié pour que chacune de ses observations s'aligne linéairement sur la dernière

observation de X' et à la première de X''^* . Ce segment est alors noté $S_{X,i}^*$. Enfin, la variable est reconstruite en associant l'ensemble des séquences de manière à obtenir $X^* = X' \cup S_{X,i}^* \cup X''^*$.

Changement de niveau *faible* :

$$X^* = \underbrace{\{x_0, \dots, x_{i-1}\}}_{X'} \cup S_{X,i}^* \cup \underbrace{\{x_{i+N} + g, \dots, x_m + g\}}_{X''^*} \text{ avec } g \in \left[\frac{1}{2} \sigma_x ; \sigma_x \left[\cup \right] -\sigma_x ; -\frac{1}{2} \sigma_x \right] \quad (11)$$

Changement de niveau *fort* :

$$X^* = \underbrace{\{x_0, \dots, x_{i-1}\}}_{X'} \cup S_{X,i}^* \cup \underbrace{\{x_{i+N} + g, \dots, x_m + g\}}_{X''^*} \text{ avec } g \in \left[\frac{3}{2} \sigma_x ; 2 \sigma_x \left[\cup \right] -2 \sigma_x ; -\frac{3}{2} \sigma_x \right] \quad (12)$$

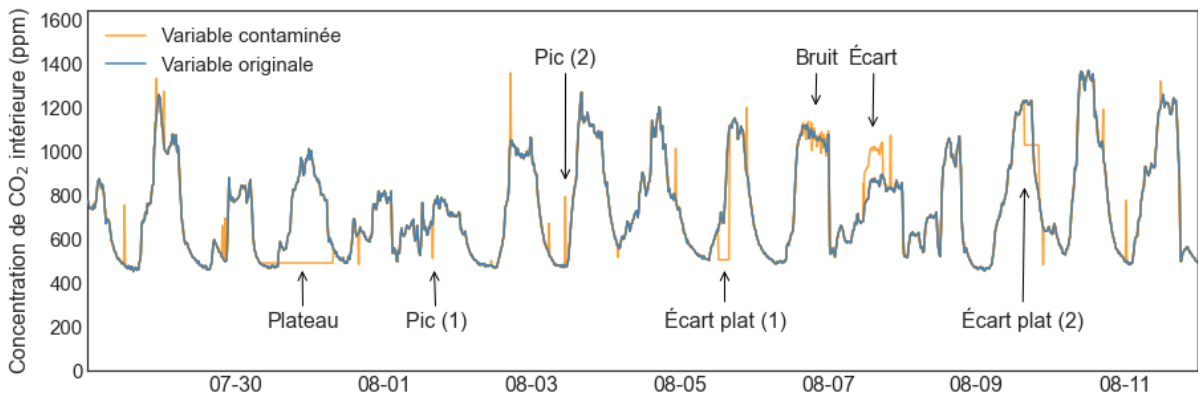


Figure 42 Exemples d'anomalies générées artificiellement sur une mesure de concentration de CO₂. La variable originale est incluse pour évaluer les différences induites par la contamination. Toutes les anomalies non annotées correspondent à des pics d'intensité variable.

Il est important de noter que cette approche présente à la fois des avantages et des limitations. Elle permet d'évaluer les performances des modèles sur divers types d'anomalies, avec des dynamiques, des tailles et des intensités différentes, appliqués à des données variées. Ainsi, elle offre un aperçu de la précision et de la robustesse des méthodes employées (Emmott et al., 2016; Sharma et al., 2010), ce qui serait difficile à réaliser autrement. Cependant, ces anomalies ne sont pas le produit d'un processus naturel, ce qui, comme l'indiquent Emmott et al. (2016), est l'approche à privilégier lorsque cela est possible. L'accès aux données, qu'elles soient étiquetées ou non, reste un facteur limitant significatif dans le domaine de l'apprentissage automatique (Himeur et al., 2021b).

Pour pallier cette limitation, le modèle proposé a été conçu pour s'adapter à plusieurs variables différentes et générer des anomalies, bien qu'artificielles, qui restent plausibles. Néanmoins, ces travaux veilleront à fournir quelques aperçus visuels des résultats de détection des méthodes employées sur des jeux d'observations contenant des anomalies réelles, même si elles ne sont pas non étiquetées.

5.3.2. Modèles d'apprentissage automatique

Dans cette partie, nous introduirons les différents modèles utilisés dans ces travaux et détaillerons leurs fonctionnements. Ces modèles se divisent en trois familles de méthodes, chacune étant associée à un symbole et une couleur spécifiques pour faciliter la lecture : voisinage (+), prédiction (▲) et reconstruction (■). Le fonctionnement des modèles basés sur les réseaux de neurones denses classiques et sur les *Long Short-Term Memory* (LSTM) ne sera pas réexpliqué ici (pour de plus amples informations, se référer à la partie 4.3.2.4 du chapitre 4). Cependant, des explications seront fournies sur le fonctionnement des modèles de prédiction ou de reconstruction qui les utilisent.

5.3.2.1. Les K plus proches voisins (+)

La méthode des K plus proches voisins, aussi appelée *K-Nearest Neighbour* (KNN), peut être utilisée à la fois de manière supervisée (p. ex., classification, régression) et de manière non supervisée pour la détection d'anomalies (Ramaswamy et al., 2000). Cette méthode, appartenant à la famille des techniques basées sur le voisinage, évalue la distance entre chaque observation ou séquence d'observations et ses K plus proches voisins. Plus précisément, elle calcule, pour une observation donnée, les distances la séparant des autres observations. Ces distances peuvent être calculées selon différentes approches (p. ex., euclidienne, de Manhattan). Ensuite, elles sont triées par ordre croissant, et la moyenne ou médiane des K distances les plus faibles, considérées comme correspondant aux K voisins les plus proches, est utilisée pour caractériser l'observation. L'observation peut aussi être directement classifiée comme normale ou anormale selon Ramaswamy et al. (2000). Dans ce cas, une observation est classée comme anormale s'il y a trop peu d'observations ayant une distance au $K^{\text{ème}}$ voisin supérieure à un nombre N prédéfini. Le choix de K est primordial, car s'il est trop faible la comparaison peut se faire entre observations anormales et être biaisée, tandis qu'un K trop important peut conduire à une généralisation excessive (Foulon, 2020).

Il existe quelques variantes notables de cette méthode, notamment les *Ensemble K-Nearest Neighbour*, qui combinent les KNN avec d'autres techniques (p. ex., Autoencoder-KNN) (Song et al., 2017) ou la méthode *Subsequence K-Nearest Neighbour* (sub-KNN), qui utilise des séquences d'observations au lieu d'observations indépendantes (Keogh et al., 2006). Cette approche, introduite pour la première fois par Keogh et al. (2006), maintenant régulièrement appelée '*time series discord*', consiste à évaluer une sous-séquence anormale d'une série temporelle par rapport au reste de la série. Cette sous-série temporelle est ensuite considérée comme discordante (Blázquez-García et al., 2022). Les méthodes des **KNN** (+) ainsi que celle des **sub-KNN** (+) seront utilisées conjointement dans ces travaux.

5.3.2.2. Le Local Outlier Factor (+)

La méthode du *Local Outlier Factor* (LOF) est une technique non supervisée basée sur le voisinage, utilisée pour la détection d'anomalies. Cette méthode, introduite par Breunig et al. (2000), permet d'identifier les anomalies en se basant sur des mesures de densités locales. Pour chaque observation, une mesure de densité est calculée. En comparant ces densités aux densités des observations voisines, il est possible d'évaluer un degré d'anormalité. Plus précisément, la méthode sélectionne les K plus proches voisins d'une observation donnée et calcule la distance qui la sépare de la $K^{\text{ème}}$ observation. Cette distance, appelée distance_K , peut être calculée selon différentes approches (p. ex., euclidienne, de Manhattan). Ensuite, le degré d'isolement de l'observation est estimé à l'aide d'une distance appelée « distance d'accessibilité ». Pour deux observations A et B , la distance d'accessibilité de A correspond à la valeur maximale entre la distance séparant A de B et la distance_K du point B (13).

$$\begin{aligned} \text{Distance d'accessibilité}(A, B) \\ = \max \{ \text{distance}(A, B), \text{distance}_K(B) \} \end{aligned} \quad (13)$$

L'utilisation de cette distance d'accessibilité permet de normaliser la valeur de la distance par rapport à la densité de l'environnement local, limitant ainsi l'influence de zones denses sur des zones moins denses. Un exemple est illustré dans la Figure 43 (i), inspirée de Breunig et al. (2000), pour une observation notée O et une valeur de K de 5. Le degré d'isolement d'une observation, aussi appelé densité d'accessibilité, correspond alors à l'inverse de la moyenne des distances d'accessibilité de l'observation vers ses K plus proches voisins. Un degré d'anormalité, aussi appelé score LOF, est calculé pour l'observation en effectuant la moyenne des ratios de l'ensemble des densités locales de l'observation sur celles de ses K voisins. Plus ce score est inférieur à 1 (resp. supérieur), plus la densité locale de l'observation est forte (resp. faible), ce qui signifie qu'elle peut être considérée comme normale (resp. anormale). Un score proche de 1 indique également que l'observation a une densité similaire à celle de ses voisins (Foulon, 2020). Un exemple de résultat obtenu pour des données en 2 dimensions, tiré de Kriegel et al. (2009) est présenté dans la Figure 43 (ii).

Il existe quelques variantes notables de cette méthode, notamment la *Connectivity-based Outlier Factor* (COF), qui évalue la densité en fonction du nombre de connexions nécessaires, aussi appelées connectivité, pour relier deux observations en utilisant leurs plus proches voisins (Tang et al., 2002). Il y a aussi la *Clustering-Based Local Outlier Factor* (CBLOF), qui se concentre sur le partitionnement de l'espace suivant une approche similaire au K -Means (cf. partie 2.4.3.2) (He et al., 2003) ainsi que la *Subsequence Local Outlier Factor* (sub-LOF), qui consiste à renseigner des séquences d'observations au lieu d'observations indépendantes (Schmidl et al., 2022). Les méthodes

du **LOF** (+) ainsi que celles du **sub-LOF** (+) seront toutes deux employées dans ces travaux.

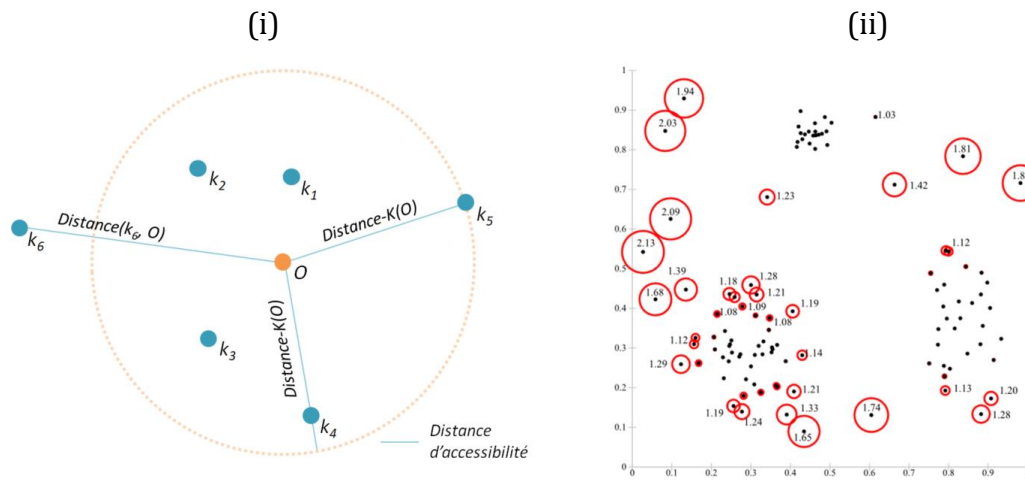


Figure 43 Exemples illustrant le fonctionnement de la méthode LOF. (i) Illustration du calcul de la distance d'accessibilité pour trois points différents et avec $K = 5$. Inspirée de Breunig et al. (2000). (ii) Présentation du score LOF pour des données en 2 dimensions. Les périmètres des cercles rouges représentent le score LOF, plus ils sont étendus, plus le score s'éloigne de 1. Tirée de Kriegel et al. (2009).

5.3.2.3. L'Isolation Forest (+)

L'*Isolation Forest* (IF), aussi appelée forêt d'isolement, est une méthode non supervisée et classée, à des fins de simplification, dans la famille des techniques à base de voisinage. Elle est spécifiquement employée pour la détection d'anomalies (Hariri et al., 2021). Cette méthode, introduite par Liu et al. (2008), est basée sur des arbres comme les forêts aléatoires (cf. partie 4.3.2.3) mais utilise une approche unique, dite « d'isolement ». Cette méthode construit un ensemble d'arbres aléatoires qui partitionnent les observations (ou séquences d'observations) dans l'espace, comme présenté dans la Figure 44.

Le modèle de l'*Isolation Forest* (IF) opère de la manière suivante : il sélectionne aléatoirement une caractéristique de l'espace qui constituera la base de l'arbre ainsi qu'une valeur comprise entre le minimum et le maximum de cette variable. Cette valeur appelée seuil de division correspond à un nœud et permet de séparer les observations. Des nœuds sont récursivement ajoutés suivant ce même principe jusqu'à ce qu'une observation soit isolée dans une feuille ou jusqu'à ce qu'un critère d'arrêt soit atteint (i et ii). Ce dernier correspond généralement à une profondeur maximale d'arbre qui représente le nombre maximum de nœud pouvant être ajoutés. L'ensemble de ce processus est alors répété jusqu'à former une forêt. Pour évaluer si une observation est anormale, elle est passée au travers de tous les arbres de la forêt, et la moyenne de son chemin, c'est-à-dire le nombre de nœuds nécessaires pour l'isoler, est calculé. Les observations anormales, étant rares et différentes des observations normales, elles sont

en général plus simple à isoler (i et ii) (Xu et al., 2023). Elles sont proches des racines des arbres et sont donc caractérisées par des chemins plus courts (iii) (Schmidl et al., 2022). Cette longueur moyenne du chemin sert ensuite de score pour déterminer la normalité ou l'anormalité.

Diverses variantes de l'*Isolation Forest* méritent d'être mentionnées, comme l'*Extended Isolation Forest* (EIF), qui est capable de diviser l'espace en diagonale (Hariri et al., 2021). Il existe aussi des approches hybrides telles que l'*Hybrid Isolation Forest* ou l'*Ensemble Isolation Forest*, qui combinent l'IF avec d'autres méthodes, par exemple le LOF-IF (Cheng et al., 2019). En outre, la méthode *Subsequence Isolation Forest* (sub-IF) est une variante qui s'applique à des séquences d'observations au lieu d'observations indépendantes (Schmidl et al., 2022). Dans le cadre de ces travaux, à la fois l'IF (+) et le sub-IF (+) seront utilisés conjointement.

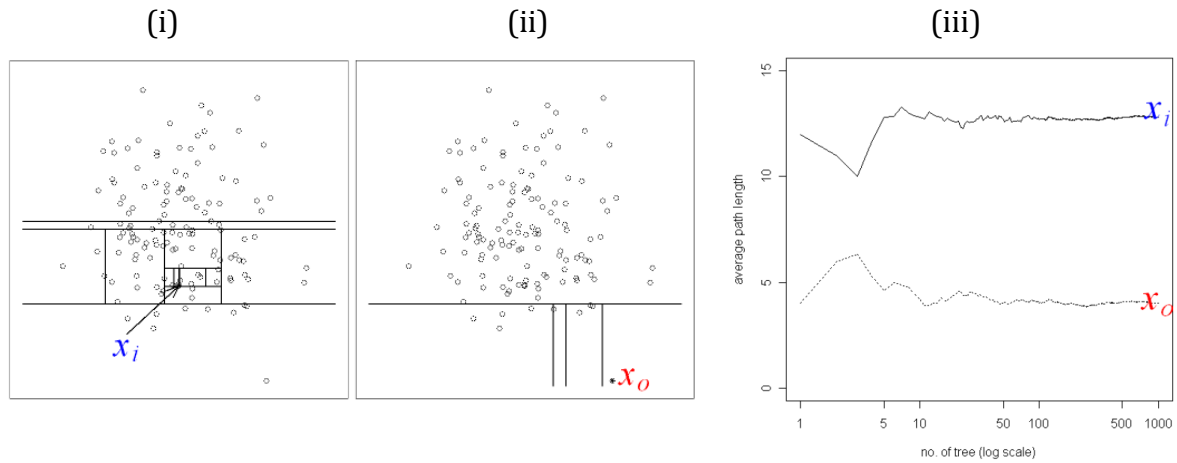


Figure 44 Exemple de fonctionnement de la méthode *Isolation Forest*. (i) Illustration d'un arbre composé de 12 nœuds qui isole une observation normale (x_1). Présentation d'un autre arbre composé de seulement trois nœuds permettant d'isoler une observation anormale (x_0). (iii) Montre la moyenne des distances pour chaque observation selon le nombre d'arbres de la forêt. Peu importe le nombre d'arbres, la distance moyenne de l'observation anormale (x_0) reste inférieure à celle de l'observation normale (x_1). Tirée de Liu et al. (2008).

5.3.2.4. L'Auto-Regressive Integrated Moving Average et autres méthodes de prédiction (▲)

L'*Auto-Regressive Integrated Moving Average* (ARIMA) est une adaptation pour les données non stationnaires du modèle de régression *Auto-Regressive Moving Average* (ARMA) (Foulon, 2020). Ces modèles sont couramment utilisés pour la détection d'anomalies en apprentissage semi-supervisé dans les séries temporelles en utilisant une approche prédictive (Bianco et al., 2001). Les modèles ARIMA se composent de trois éléments (Siarni-Namini et al., 2018) :

- (i) **Auto-Régression (AR).** Ce modèle de régression utilise les observations passées pour prédire les observations futures avec un décalage.

- (ii) **Intégré (I).** Il s'agit d'un processus de transformation, souvent une différenciation, qui peut être appliqué pour éliminer les tendances des observations et les rendre stationnaires.
- (iii) **Moyenne Mobile (Moving Average ou MA).** Ce modèle évalue l'erreur ou le bruit qui se produit à intervalles réguliers en appliquant une moyenne mobile (Vandewalle et al., 1999) aux observations passées.

Dans le cas d'un modèle ARMA, seules les composantes « AR » et « MA » sont utilisées. Il existe des variantes significatives de cette méthode, notamment la *Vector Auto-Regression Integrated Moving Average* (VARIMA ou *Vector ARIMA*), adaptée aux séries temporelles multivariées non stationnaires (Öller, 1985). Une autre variante est la *Seasonal Auto-Regressive Integrated Moving Average* (SARIMA), qui intègre une composante liée à la saisonnalité (Greis et al., 2018).

Dans le contexte de la détection d'anomalies, l'approche prédictive commune à tous les modèles consiste en ce qui peut être appelé une *rétro-prédiction*. Cela implique de prédire une ou plusieurs observations passées au moyen d'observations antérieures, comme illustré dans la Figure 45 (i), tirée des travaux de Blázquez-García et al. (2022). Pour ce faire, le modèle est d'abord entraîné sur des données considérées comme normales, puis appliqué à de nouvelles données. Une erreur de prédiction est ensuite calculée en comparant les mesures réelles aux prédictions réalisées (ii) (Malhotra et al., 2015). Une observation ou une séquence d'observations prédite qui s'éloigne fortement des mesures réelles présente une erreur de reconstruction plus importante et a davantage de chances d'être considérée comme anormale.

Il existe plusieurs méthodes pour calculer l'erreur de prédiction, parmi lesquelles les plus couramment utilisées sont l'erreur absolue moyenne (*Mean Absolute Error* ou MAE), l'erreur quadratique moyenne (*Mean Squared Error* ou MSE) ou la racine de l'erreur quadratique moyenne (*Root Mean Squared Error* ou RMSE) (Shcherbakov et al., 2013). Cette approche revient donc à détecter des événements nouveaux qui n'ont pas été appris par le modèle, et donc à faire de la détection de nouveauté.

Dans le cadre de ces travaux, la méthode **ARIMA** (▲) sera appliquée conjointement avec des modèles de réseaux de neurones **Dense-p** (▲) et **LSTM-p** (▲). Ces deux derniers modèles fonctionnent conformément à ce qui a pu être détaillé dans la partie 4.3.2.4, bien qu'ils soient ici utilisés pour une tâche de prédiction plutôt que de classification ou détection.

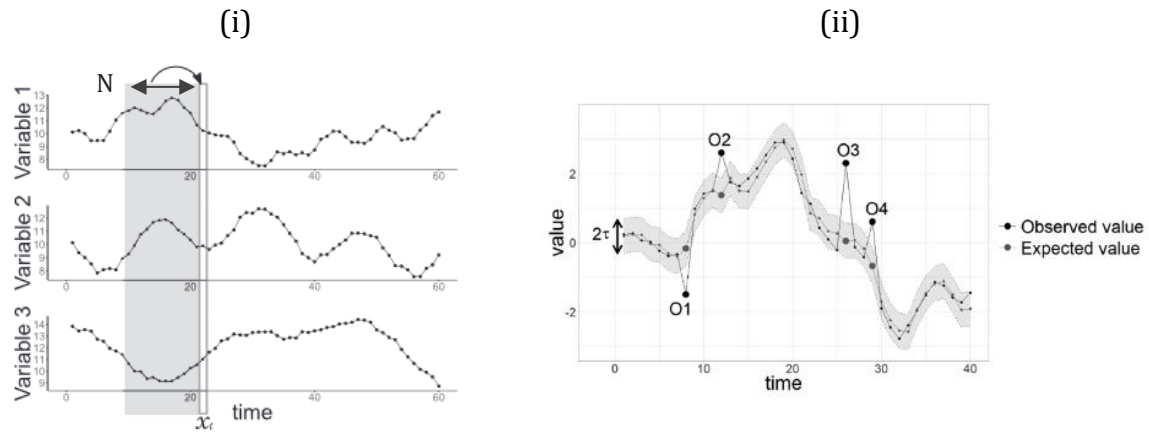


Figure 45 Processus de détection d'anomalies avec des modèles de prédiction. (i) Exemple de rétro-prédiction : une observation, notée x_t , est prédite en utilisant une séquence (zone grise) de N observations antérieures. (ii) Exemple de comparaison entre valeurs prédites et valeurs réelles. Les figures (i) et (ii) sont tirées de Blázquez-García et al. (2022).

5.3.2.5. L'Auto-Encodeur et l'Encodeur-Décodeur (■)

Les modèles d'Encodeur-Décodeur (ED) et les modèles d'Auto-Encodeur (AE), également appelés *Encoder-Decoder* et *Autoencoder*, sont des méthodes d'apprentissage automatique de reconstruction (Schmidl et al., 2022), appartenant à la famille des réseaux de neurones. Leur architecture comprend deux composantes principales : un encodeur (pour l'encodage) et un décodeur (pour le décodage), comme illustré dans la Figure 46 (ii), en contraste avec un réseau de neurones « classique » (i). Cependant, bien que ces deux architectures se ressemblent, leurs objectifs et applications diffèrent :

- **L'Autoencodeur.** L'encodeur d'un AE est composé de couches cachées successives, avec un nombre décroissant de neurones artificiels. Il prend en entrée des séquences d'observations et les transforme en une représentation plus compacte, souvent de dimension réduite, appelée *espace latent* (Bank et al., 2021; Girin et al., 2021). Bien que compressée, cette représentation est conçue pour capturer les caractéristiques essentielles des observations d'origine. Le décodeur, constitué de couches de taille croissante, utilise cette représentation pour tenter de reproduire fidèlement les séquences d'entrée d'origine.
- **L'Encoder-Decoder.** Le processus initial est proche de celui de l'AE, car l'encodeur transforme également les séquences d'observations en entrée en une représentation latente. Cependant, le décodeur utilise cette représentation pour générer des séquences en sortie qui sont généralement différentes de celles en entrée. L'Encodeur-Décodeur est donc une généralisation de l'Auto-Encodeur, où les entrées et les sorties peuvent différer (Costa-jussà et al., 2018).

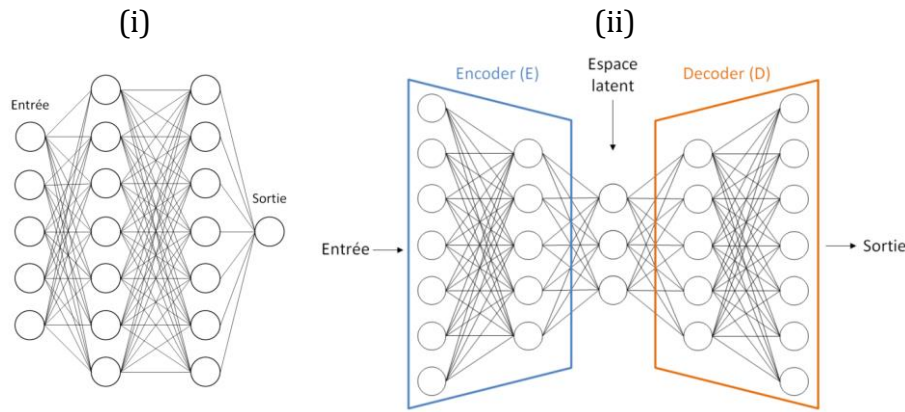


Figure 46 Comparaison de la structure d'un réseau de neurones « classique » (i) avec celle commune à l'Auto-Encodeur et à l'Encodeur-Décodeur (ii).

De façon équivalente aux réseaux de neurones, l'ensemble du processus d'apprentissage est effectué selon la minimisation d'une fonction de coût (p.ex. RMSE) qui permet d'évaluer la qualité des constructions ou reconstructions. Plus l'encodeur et le décodeur présentent de couches plus ils sont considérés profonds. Il existe plusieurs variations des AE et des ED. On pourra citer de manière non exhaustive (Bank et al., 2021) :

- **Le Variational Autoencoder (VAE).** Ce modèle, également appelé Auto-Encodeur variationnel, est un modèle probabiliste dit « génératif ». Appelé *Autoencoder* car il présente une structure d'encodeur-décodeur (Doersch, 2021), ce modèle a été introduit relativement récemment par Kingma et Welling (2013) et Rezende et al. (2014). Contrairement à l'AE classique, qui apprend une représentation latente fixe des séquences d'entrées, le VAE adopte une approche probabiliste. Son encodeur produit des vecteurs de moyennes et d'écart-types pour chaque séquence d'entrée, définissant ainsi une distribution normale. Le décodeur utilise ensuite un échantillon de cette distribution pour reconstruire la séquence d'entrée d'origine. En outre, en plus de la fonction coût traditionnelle de l'AE, le VAE effectue également une régularisation à l'aide de la divergence de Kullback-Leibler pour maintenir la distribution latente proche d'une distribution normale (Girin et al., 2021). Cela en fait un modèle particulièrement adapté à la génération de données.
- **Le Recurrent Autoencoder (RAE).** Également appelé Auto-Encodeur récurrent, ce modèle est une variante de l'Auto-Encodeur et de l'Encodeur-Décodeur (RED) dans laquelle les couches traditionnelles sont remplacées par des couches de neurones récurrents, comme présenté dans la partie 4.3.2.4 (Yang et al., 2020). L'utilisation de couches récurrentes permet de capturer des dépendances à plus ou moins long terme dans les données (Kieu et al., 2019).

- Le Denoising Autoencoder (DAE).** Cette variante de l'Auto-Encodeur, ou de l'Encodeur-Décodeur (DED), parfois appelée Auto-Encodeur débruiteur, est conçue pour éliminer le bruit des données. Elle consiste à fournir des données « corrompues » à un AE lors de l'entraînement, de sorte qu'il puisse reconstruire une version « réparée » des données (Vincent et al., 2008). Pour cela un bruit est ajouté pendant la phase d'entraînement à chaque séquence d'observations renseignée en entrée. Ce bruit peut être généré de différentes manières, généralement sous la forme d'un bruit gaussien ou en masquant certaines observations (Vincent et al., 2008). Le décodeur utilise ensuite la représentation issue des entrées bruitées pour tenter de reconstruire la version originale, non bruitée, des données. L'ajout de bruit en entrée permet d'entraîner un AE plus robuste, moins sensible au bruit, ce qui le rend particulièrement adapté à la détection d'anomalies (Chen et al., 2020). L'ensemble de ce processus est illustré dans la Figure 47.

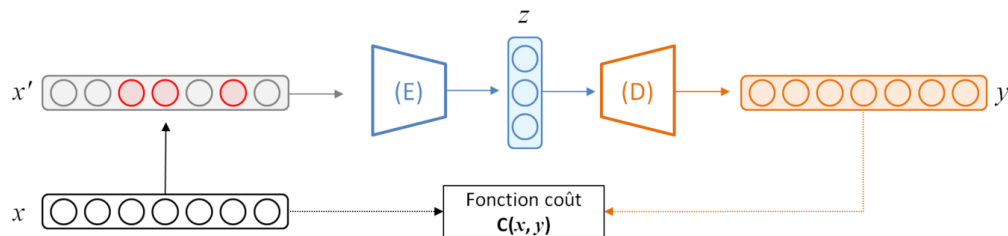


Figure 47 Représentation du fonctionnement d'un DAE. x représente la séquence d'observations d'origine des entrées, x' la séquence bruitée utilisée par le modèle (rouge pour les observations bruitées), z l'espace latent correspondant et y la sortie reconstruite par le modèle.

- Les Sparse Autoencoder (SAE).** Cette variante d'Autoencoder ou d'Encoder-Decoder (SED), parfois appelée Auto-Encodeur épars, est conçue pour éviter le sur-apprentissage en contraignant le modèle à apprendre des caractéristiques plus distinctives des données (Ng, 2010). L'idée est de n'avoir qu'une fraction des neurones actifs à la fois, les forçant à capturer des caractéristiques uniques, spécifiques et représentatives des données. Pour ce faire, une pénalité de *parcimonie* est imposée, généralement mise en œuvre à l'aide de la divergence de Kullback-Leibler ou d'une régularisation d'activité L_1 (Li et al., 2021). Cette parcimonie induite permet de former un modèle plus robuste, particulièrement adapté à la réduction de dimension ou à la détection d'anomalies (Kieu et al., 2019). La Figure 48, inspirée de Koike-Akino et Wang (2020), présente une architecture classique d'un AE (i) comparée à celle d'un SAE (ii).

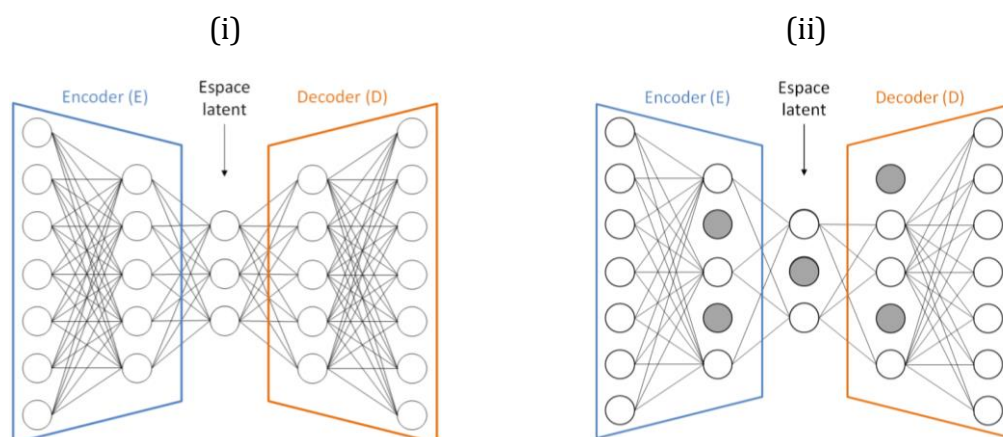


Figure 48 Architecture d'un AE classique (i) et d'un Sparse AE (ii). Les neurones désactivés sont représentés en gris et ne sont plus reliés aux neurones actifs. Inspirée de Kokie-Akino et Wang (2020).

Les modèles d'Auto-Encodeur et d'Encodeur-Décodeur peuvent être utilisés pour la détection d'anomalies selon deux approches différentes.

La première consiste à évaluer l'erreur de reconstruction entre chaque séquence d'entrée et de sortie du modèle au moyen de différentes méthodes de calcul, telles que la RMSE ou la MAE. Cette méthode repose sur l'hypothèse que le modèle ne peut pas reproduire efficacement des événements rares tels que les anomalies. Elle est similaire à l'approche des modèles prédictifs tels que l'ARIMA ou le LSTM-p, où une erreur de reconstruction est élevée indique une plus grande probabilité que la séquence soit anormale.

La deuxième approche, plus récente, consiste à examiner la distribution ou représentation de l'espace latent. Elle repose sur l'idée que les anomalies ont tendance à se regrouper dans des régions de faible densité de l'espace latent (Angiulli et al., 2022). Elles peuvent alors être détectées en utilisant d'autres techniques de détection d'anomalies complémentaires, telles que le *clustering* ou le voisinage. Cette approche permet également d'effectuer des tâches de classification de manière non supervisée (Norlander and Sopasakis, 2019).

Dans ces travaux, plusieurs variantes d'Autoencodeur et d'Encodeur-Décodeur seront utilisées, notamment des **AE** classiques (■), des **RAE** (■), des **DAE** (■) et des **SAE** (■) ainsi que leurs **variantes d'ED** classiques (■), de **RED** (■), de **DED** (■) et de **SED** (■).

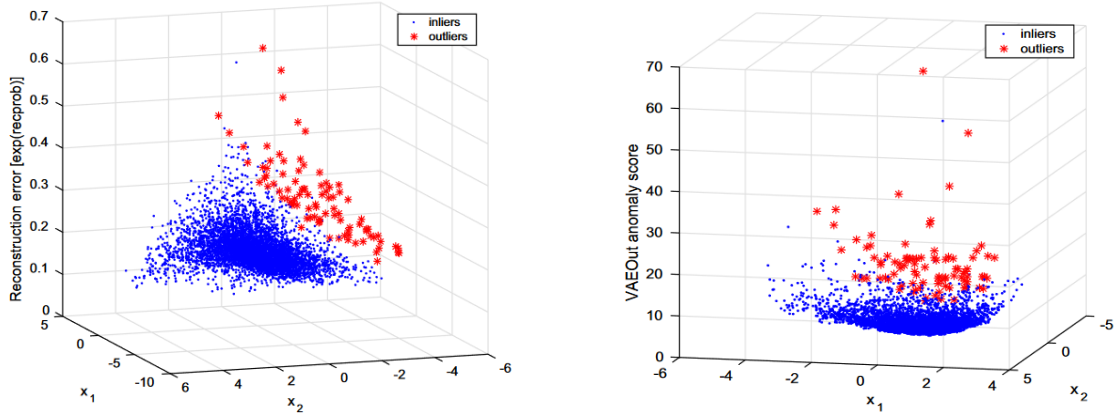


Figure 49 Représentation d'observations normales (bleues) et anormales (rouges) dans l'espace latent d'un VAE. Issue de Agiulli et al. (2022).

5.3.2.6. La Moyenne Mobile Exponentielle, une méthode « naïve »

En plus des modèles précédemment mentionnés, nous utilisons également un modèle de moyenne mobile exponentielle pour proposer une méthode « naïve » de détection d'anomalies. L'objectif de cette approche est de pouvoir comparer les performances de nos modèles à une méthode intuitive et simple d'application. Cette démarche nous permet d'évaluer si nos modèles offrent réellement un avantage par rapport à une méthode de référence.

L'utilisation de cette approche constitue une « *bonne pratique* » importante à considérer, bien que ce ne soit pas toujours le cas (Wu and Keogh, 2021). En fonction des cas et des applications, cette méthode peut prendre différentes formes, par exemple, un modèle de régression classique (Liu et al., 2020). Le lissage introduit par la moyenne mobile exponentielle semble être un apport pertinent pour la détection d'anomalies. Selon le coefficient de lissage (α) choisi, un poids différent est attribué aux observations passées, ce qui pourrait potentiellement être avantageux pour détecteur des anomalies ponctuelles et collectives.

Le modèle de moyenne mobile exponentielle que nous utilisons est nommé **MME-naïve** et utilise différents coefficients de lissage, notamment $\alpha = \{0,2 ; 0,125 ; 0,05 ; 0,02\}$, comme décrit dans la partie 4.4.3.

5.3.3. Configuration et entraînement des modèles

Dans cette partie les spécificités de construction et d'entraînement de chacun des modèles seront présentés, ainsi que leurs paramétrages. Tous les modèles utilisés dans ces travaux ont été construits en utilisant Python 3.8 avec les bibliothèques *Scikit-learn*, *tensorflow* et plus particulièrement *keras*. Toutes les phases d'entraînement ont été

réalisées sur un processeur Intel Core i7-8850H CPU, équipé de 6 cœurs fonctionnant à 2,6 GHz.

5.3.3.1. Conditionnement et construction des modèles de reconstruction (■) et de prédiction (▲)

Les modèles de reconstruction et de prédiction ont été entraînés en utilisant une fenêtre glissante de taille N_e , se décalant d'un pas dans le temps suivant, une méthode similaire à celle décrite dans le chapitre 4, partie 4.3.3.3. Afin d'évaluer l'impact de la taille de la fenêtre glissante sur la tâche de détection (Li et al., 2023; Wei et al., 2022), plusieurs tailles d'entrée N_e et de sortie N_s ont été testées : 12, 24, 48 et 96, correspondant à des fenêtres de 3, 6, 12 et 24 heures. Pour les modèles de reconstruction, la taille de fenêtre correspond également à la taille des séquences en sortie (N_s). En revanche, pour les modèles de prédiction, différentes tailles de sortie ont été évaluées, correspondant à 4 et 12 observations, soit à des prédictions sur 1 et 3 heures. Toutes ces spécifications sont résumées dans le Tableau 13.

Étant donné les performances obtenues par les modèles prédictifs sur un seul pas de temps ($N_s = 1$) dans la détection de séquences anormales, comme observé dans les travaux de Liu et al. (2020), et compte tenu du nombre d'anomalies collectives étudiées, aucune prédiction à un seul pas de temps n'a été réalisée afin d'optimiser les temps de calculs. L'approche présentée dans cette partie, où plusieurs éléments sont renseignés en entrée du modèle, induisant également une sortie composée de plusieurs éléments, est appelée '*many to many*' dans la littérature (Kim et al., 2017).

Concernant la structure des modèles, deux approches ont été adoptées. D'une part, les modèles basés sur les réseaux de neurones, tels que les variantes d'AE, d'ED, les Dense-p et les LSTM-p, et d'autre part, l'ARIMA :

- Pour les modèles basés sur les réseaux de neurones, tous les hyperparamètres ont été optimisés au préalable pour trouver un équilibre entre les performances, les difficultés d'apprentissage et le temps de calcul (Hasebrook et al., 2023). Pour les modèles de reconstruction, l'encodeur est composé d'une première couche de 128 ou 256 unités, d'une deuxième couche de 64 ou 128 unités et d'une dernière couche pour l'espace latent de 16 ou 32 unités. Le décodeur est construit de manière symétrique. Pour les DAE, un bruit gaussien est ajouté en entrée de l'encodeur pendant l'entraînement. Dans le cadre de l'évaluation, un encodeur identique est reconstruit sans bruit, et les poids des neurones résultant de l'entraînement sont transférés. Pour les SAE, une régularisation L_1 est appliquée à l'espace latent pour désactiver aléatoirement des neurones (Li et al., 2021). Les sorties des modèles passent par une couche dense utilisant une activation *sigmoïde*, et le nombre d'unités est adapté à la taille des entrées. L'algorithme d'optimisation *Adam* est utilisé avec un

taux d'apprentissage de 0,001, ainsi qu'une fonction de coût MAE. Pour prévenir le sur-apprentissage, 20 % des données d'entraînement sont réservées comme ensemble de validation, les données restantes sont mélangées avant chaque époque de l'entraînement (*epoch*), et une technique d'arrêt précoce (*early stopping*) est appliquée. Les modèles de prédiction, Dense-p et LSTM-p, sont construits selon une approche similaire aux encodeurs mais possèdent également une couche de *dropout* à 0,2.

- De manière similaire aux modèles basés sur les réseaux de neurones, les hyperparamètres de la méthode ARIMA ont été optimisés et sélectionnés en amont. Cependant, cette méthode nécessite d'être préalablement ajustée aux données pour estimer certains coefficients associés aux composantes « AR » et « MA ». Afin d'éviter que le modèle ne devienne obsolète en s'appuyant sur des données trop anciennes, une actualisation régulière a été mise en place. Cette procédure de mise à jour est coûteuse en matière de temps et de ressources, et constitue l'une des principales faiblesses de cette technique (Kraiem, 2021). La mise à jour s'effectue toutes les 12 à 24 prédictions.

Tableau 13 Résumé des spécificités relatives aux entrées et sorties des modèles de voisinage, de prédiction et de reconstruction utilisés.

Familles	Modèles	Apprentissage	Entrées (Taille N_e)	Sorties (Taille N_s)
Voisinage (✚)	IF KNN LOF	Non supervisé	Observation $N_e = 1$	Score d'anormalité d'observation $N_s = 1$
	Sub-IF Sub-KNN Sub-LOF	Non supervisé	Séquence d'observations $N_e = \{12 ; 24 ; 48 ; 96\}$	Score d'anormalité de séquence $N_s = 1$
Prédiction (▲)	ARIMA	Non supervisé	Séquence d'observations $N_e = \{12 ; 24 ; 48 ; 96\}$	Séquence prédite $N_s = \{4 ; 12\}$
	Dense-p* LSTM-p*	Semi-supervisé	Séquence d'observations $N_e = \{12 ; 24 ; 48 ; 96\}$	Séquence prédite $N_s = \{4 ; 12\}$
Reconstruction (■)	AE*, ED* RAE*, RED* DAE*, DED* SAE*, SED*	Semi-supervisé	Séquence d'observations $N_e = \{12 ; 24 ; 48 ; 96\}$	Séquence reconstruite $N_s = \{12 ; 24 ; 48 ; 96\}$

Enfin, en ce qui concerne l'ensemble des modèles de reconstruction et de prédiction en apprentissage semi-supervisé, les données sont normalisées dans la plage de 0 à 1, en se basant uniquement sur les caractéristiques observées dans l'ensemble d'entraînement.

Les caractéristiques de cette normalisation seront ensuite appliquées à l'ensemble des données de test.

5.3.3.2. Conditionnement et construction des modèles de voisinage (✚)

Les modèles de voisinage sont appliqués de deux façons différentes, comme indiqué dans le Tableau 13. Les modèles classiques d'IF, de KNN et de LOF sont utilisés selon une approche '*one to one*', de sorte à ce que chaque observation individuelle en entrée génère une sortie de même taille représentant son score d'anormalité (cf. partie 4.3.3.3). Les modèles adaptés aux séquences, à savoir sub-IF, sub-KNN et sub-LOF, sont appliqués aux données séquencées à l'aide d'une fenêtre glissante de taille variable qui se déplace d'un pas dans le temps. La procédure et la taille des fenêtres employées sont identiques à celles décrites pour les familles de reconstruction et de prédiction. Cependant, les modèles de voisinage adaptés aux séquences se distinguent par leur sortie constituée d'un seul élément ($N_s = 1$), qui représente le score d'anormalité de la séquence en entrée. Dans ce cas, l'approche est appelée '*many to one*' (cf. partie 4.3.3.3).

Les hyperparamètres de tous les modèles KNN, y compris les variantes séquentielles sub-KNN, ont été préalablement optimisés. Pour chaque modèle, une évaluation récursive a été réalisée en utilisant trois valeurs différentes pour le nombre de voisins proches, noté K_{KNN} , qui dépend de la taille du jeu de données. Ces valeurs peuvent être approximées par $K_{KNN} = \{10 ; 60 ; 120\}$. Une approche similaire a été adoptée pour les méthodes d'IF et sub-IF, où le nombre d'arbres, N_{arbre} , est également déterminé en fonction de la taille de l'ensemble des données. Il est possible d'approximer ce nombre par $N_{arbre} = \{50 ; 100 ; 200\}$. En ce qui concerne les méthodes LOF et sub-LOF, le nombre K_{LOF} de voisins utilisé est sélectionné de la même manière et peut être approximé par $K_{LOF} = \{20 ; 60 ; 120\}$.

Enfin, pour toutes les méthodes de voisinage, les données sont également normalisées entre 0 et 1 uniquement par rapport aux données d'évaluation.

5.3.4. Processus d'évaluation des modèles

5.3.4.1. Estimation des erreurs de prédiction (▲) et de reconstruction (■)

Contrairement aux modèles basés sur les méthodes de voisinage, les modèles de reconstruction et de prédiction ne produisent pas de score directement lié à l'anormalité. Cependant, puisque ces modèles sont entraînés uniquement sur des séquences d'observations normales, ils ont du mal à prédire ou reconstruire efficacement des séquences anormales lors de l'évaluation (Chalapathy and Chawla, 2019; Foulon, 2020; Schmidl et al., 2022). Par conséquent un score d'anormalité peut être calculé en évaluant

les différences entre les séquences d'origine (ou mesurées) et les séquences prédites ou reconstruites. Pour ce faire, plusieurs approches mesurant la *similarité entre séquences* (Benkabou, 2018), sont utilisées :

- **La Mean Absolute Error (MAE).** Cette mesure évalue la distance moyenne absolue entre chaque élément des séquences mesurées et des séquences modélisées. La MAE, ou erreur absolue moyenne, est calculée selon l'équation (14) avec N représentant le nombre total d'observations composant une séquence, x_i désignant une observation de la séquence mesurée $S_x = \{x_1, x_2, \dots, x_N\}$, et y_i représentant une observation de la séquence reconstruite ou prédite $S_y = \{y_1, y_2, \dots, y_N\}$. Cette approche revient à évaluer la distance moyenne de Manhattan entre deux séquences.

$$MAE_{(S_x, S_y)} = \frac{1}{N} \sum_{i=1}^N |x_i - y_i| \quad (14)$$

- **La Mean Squared Error (MSE).** Cette mesure évalue la distance au carré entre chaque élément des séquences mesurées et des séquences modélisées. La MSE, ou erreur quadratique moyenne, se calcule selon l'équation (15) avec N représentant le nombre total d'observations composant une séquence, x_i désignant une observation de la séquence mesurée $S_x = \{x_1, x_2, \dots, x_N\}$, et y_i représentant une observation de la séquence reconstruite ou prédite $S_y = \{y_1, y_2, \dots, y_N\}$. La différence majeure entre la MAE et la MSE réside dans le fait que la MSE pénalise davantage les erreurs conséquentes en raison de l'élévation au carré, ce qui peut avoir une influence significative sur les résultats.

$$MSE_{(S_x, S_y)} = \frac{1}{N} \sum_{i=1}^N (x_i - y_i)^2 \quad (15)$$

- **La Dynamic Time Warping (DTW).** Elle correspond à une mesure de similarité qui a initialement été utilisée pour de la reconnaissance vocale par Sakoe et Chiba (1978). Elle est couramment employée pour évaluer les similarités entre séries temporelles (Bemdt and Clifford, 1994; Csillik et al., 2019) et est souvent utilisée dans des tâches de classification de séquences (Senin, 2008) ainsi que pour évaluer la précision des prédictions (Laperre et al., 2020). Pour deux séquences, S_x et S_y , de longueurs respectives M et N , notées $S_x = \{x_1, x_2, \dots, x_M\}$ et $S_y = \{y_1, y_2, \dots, y_N\}$, la DTW mesure le chemin minimum dans une matrice de distance permettant d'obtenir leur alignement optimal. Un exemple de matrice de distance, basé sur celui de Mori et al. (2016) et adapté, est illustré dans la Figure 50. Chaque case de la matrice, notée $d(x_i, y_j)$, correspond à la distance euclidienne entre une paire de points x_i et y_j , reflétant le coût de l'alignement de l'observation x_i avec y_j . Plusieurs chemins sont calculés de manière récursive tout en respectant trois conditions :

1. un chemin commence en position (1, 1) et termine en position (M, N) ;
2. chaque déplacement ne peut se faire qu'en direction des cases adjacentes à la position actuelle ;
3. et il n'est pas possible de revenir en arrière dans la matrice.

Le chemin minimum est celui pour lequel la somme de tous les coûts d'alignement le composant est la plus basse (Luczak, 2017). La Figure 50 illustre les chemins entre les observations utilisées pour calculer la similarité entre deux séquences. L'avantage de la DTW (ii) réside dans sa flexibilité par rapport à d'autres approches comme la MSE (iii) (Csillik et al., 2019). Cependant, il est important de noter que la DTW n'est pas une mesure de distance car elle ne respecte pas toujours l'inégalité triangulaire (Luczak, 2017).

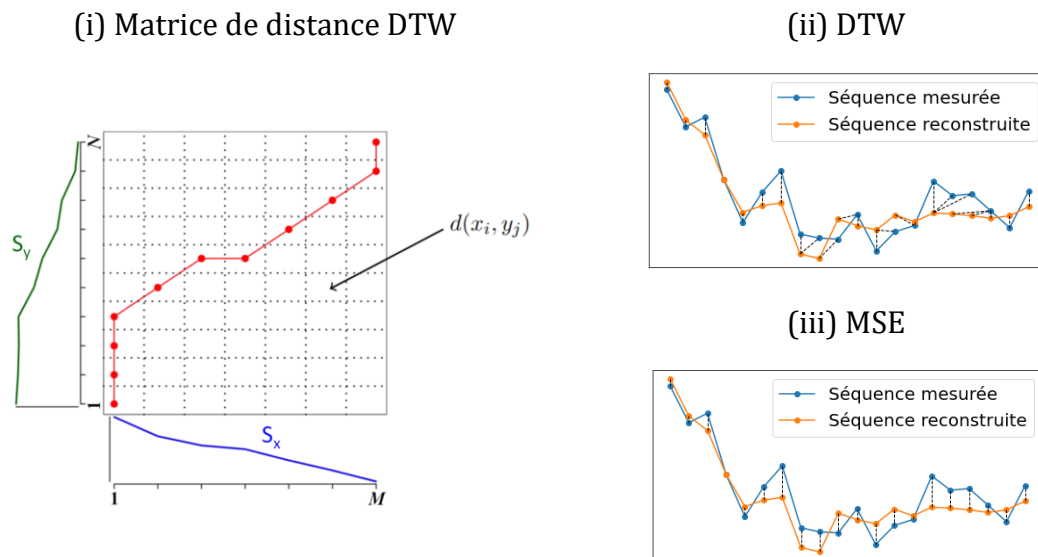


Figure 50 Exemple d'application de la DTW entre deux séquences de tailles M et N différentes (i). Adaptée de Mori et al. (2016). Exemple des chemins suivis pour calculer les similarités entre deux séquences à l'aide de la DTW (ii) et de la MSE (iii).

Ces trois mesures de similarité servent à calculer un score d'anormalité pour chaque séquence prédite ou reconstruite par les modèles. Ce score permet de quantifier la précision d'une prédiction ou d'une reconstruction par un scalaire (ou valeur unique). Par exemple, dans le cas d'un modèle de reconstruction, un score bas indique une séquence reconstruite similaire à l'originale, tandis qu'un score élevé traduit une mauvaise reconstruction, souvent causée par une anomalie (Chalapathy and Chawla, 2019). Ce score est appelé *score d'anormalité de séquences*, et il est similaire à celui produit par les modèles de voisinage sub-IF, sub-KNN et sub-LOF, par opposition au *score d'anormalité des observations* produit par les modèles IF, KNN et LOF.

En complément du score d'anormalité de séquences obtenu grâce aux trois mesures de similarité précédemment présentées (MAE, MSE et DTW), une deuxième mesure est proposée pour évaluer les différences de prédiction ou de reconstruction non pas entre séquences, mais entre observations individuelles. Pour ce faire, il est nécessaire de transformer l'ensemble des séquences prédites ou reconstruites par les modèles en une variable unique en moyennant les observations similaires.

Pour illustrer cette démarche, notons $X = \{x_1, x_2, \dots, x_M\}$, une variable composée de M observations. En utilisant une approche de fenêtrage de taille N se déplaçant d'un pas à la fois, chaque observation apparaît N fois, à l'exception des $N-1$ premières et dernières. Elles sont donc reconstruites ou prédites plusieurs fois, une pour chaque séquence qui les inclut, dans l'ensemble des sorties du modèle, noté $Y = \{S_{y,1}, S_{y,2}, \dots, S_{y,M-N+1}\}$. Il est alors possible de créer une variable de sortie moyenne $\bar{Y}$ de même taille que X en appliquant l'équation (16). Cette opération revient à calculer la moyenne de toutes les observations de sortie ayant le même indice que les observations d'entrées x_i , qui composent Y .

Dans cette équation, $S_{y,i-j+1}\{y_j\}$ représente l'observation de sortie y_j à l'indice j , compris entre 1 et N , pour chacune des séquences qui les incluent, notées $S_{y,i-j+1}$, avec un indice i compris entre 1 et M , et appartenant à l'ensemble Y . Cette approche est appelée **Distance Moyenne par Assemblage (DMA)**. Elle est utilisée pour considérer le score d'anormalité, que ce soit pour les modèles de reconstruction et de prédiction, comme un problème entre observations plutôt qu'entre séquences. Cela peut apporter un éclairage différent sur les performances observées. Ce score est donc désigné comme un *score d'anormalité d'observations*.

$$DMA_{(x_i, \bar{y}_i)} = |x_i - \bar{y}_i| \quad \text{avec} \quad \bar{y}_i = \frac{1}{N} \sum_{j=1}^N S_{y,i-j+1}\{y_j\} \quad (16)$$

5.3.4.2. Processus de détection d'anomalies

Le processus de détection d'anomalies sera évalué en utilisant la courbe ROC, et plus précisément l'indicateur *AUC-ROC* (cf. partie 2.4.2.2), pour l'ensemble des modèles. Cette métrique présente deux avantages lorsqu'elle est appliquée à un problème de détection d'anomalies. Premièrement, comme vu précédemment, elle permet d'évaluer la capacité d'un modèle à distinguer les observations normales des observations anormales sans être liée à un seuil de décision d'anormalité spécifique. Dans notre contexte, cela nous permet d'évaluer les performances de tous les modèles de manière similaire. Deuxièmement, l'*AUC-ROC* est peu sensible au taux de contamination de l'ensemble de test, contrairement à d'autres métriques comme le *F1-score* ou la précision (Fourure et al., 2021). En effet, comme l'ont noté Emmott et al. (2016) ou Wu et Keogh (2021), le taux de contamination

peut avoir un impact significatif sur les performances des modèles. Il est donc important de prendre en compte cet impact pour comparer les performances des modèles sur différentes données, présentant différentes anomalies et, par conséquent, différents taux de contamination.

Deux approches permettant de distinguer une observation normale d'une observation anormale seront utilisées pour construire les courbes ROC et calculer l'indicateur AUC-ROC. Ces approches varient selon l'utilisation d'un score d'anormalité d'observations ou de séquences.

- **Score d'anormalité d'observations.** Il est utilisé pour distinguer les observations anormales des observations normales, en fonction d'un seuil de décision. Cette approche est appliquée aux modèles de voisinage KNN, IF et LOF, qui génèrent un score d'anormalité pour chaque observation. De plus, elle est utilisée pour les modèles de prédiction et de reconstruction lorsque le score d'anormalité est calculé en employant la DMA. Pour détailler cette approche, prenons un exemple illustré dans la Figure 51 avec un ensemble de 40 observations mesurant la concentration de CO₂. Les observations anormales sont marquées d'un « A », tandis que les observations normales sont notées « O ». Avec le seuil (a), seule l'anomalie A₁ est détectée, tandis qu'avec le seuil (b), les deux anomalies A₁ et A₂ sont détectées, ainsi qu'une observation normale O₁₈.

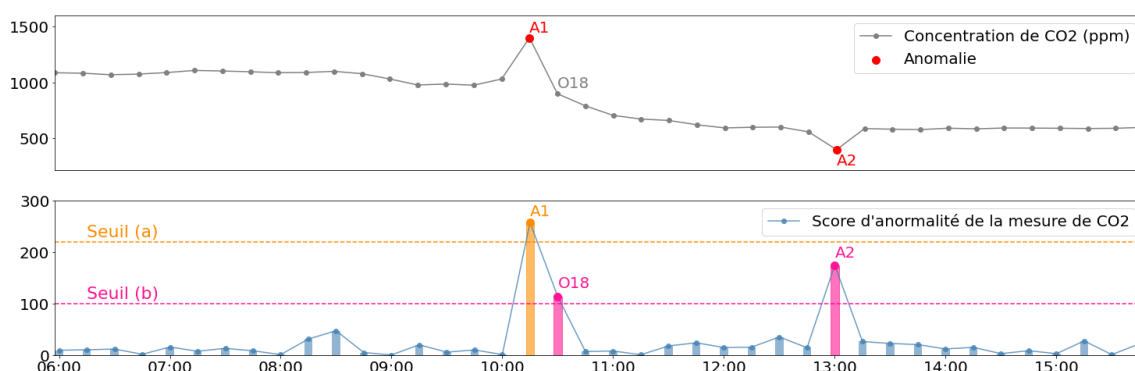


Figure 51 Exemple de détection d'anomalie en utilisant un score d'anormalité des observations. La ligne orange représente un seuil de décision (a) de 220 et la ligne rose un seuil (b) de 100. Les anomalies sont notées « A » et les observations sont notées « O ».

- **Score d'anormalité de séquences.** Il est utilisé pour déterminer si une observation est anormale en fonction du score d'anormalité de toutes les séquences qui l'incluent. Pour qu'une observation x_i soit considérée anormale, il faut que tous les scores d'anormalité associés aux séquences de taille N , allant de l'indice $i - N - 1$ à i , dépassent le seuil de décision. Cette méthode est utilisée pour les modèles séquentiels de voisinage sub-IF, sub-KNN et sub-LOF, qui génèrent un score d'anormalité pour des séquences d'observations. De plus, elle est également appliquée aux méthodes de

reconstruction et de prédiction lorsque le score d'anormalité est obtenu en évaluant la similarité entre séquences à l'aide de la MSE, de la MAE ou de la DTW. La Figure 52 permet d'illustrer cette approche avec, en exemple, un ensemble de 40 observations mesurant la concentration de CO₂. Les observations anormales sont marquées d'un « A », tandis que les observations normales sont notées « O ». L'anomalie A₁ est en 17^{ème} position, tandis que celle notée A₂ est en 28^{ème} position. Les séquences d'observations sont notées $S_i = \{O_i, \dots, O_{i+N-1}\}$ où N est la taille de fenêtre considérée. Par exemple, pour une taille de fenêtre $N = 4$ appliquée dans cet exemple, la 16^{ème} séquence serait ainsi notée : $S_{16} = \{O_{16}, A_1, O_{18}, O_{19}\}$. Avec le seuil (a), aucune anomalie n'est détectée car il aurait fallu que l'ensemble des scores d'anormalité des séquences S_{14}, S_{15}, S_{16} et S_{17} dépasse le seuil fixé, ce qui n'est pas le cas pour S_{14} . En revanche, avec le seuil (b), deux anomalies sont détectées : A₁ mais aussi O₁₈ puisque les scores des quatre séquences dans lesquelles elles apparaissent sont tous au-dessus du seuil fixé. L'observation anormale A₂ n'est quant à elle pas considérée anormale, car les séquences S_{25} et S_{27} ne dépassent pas le seuil (b).

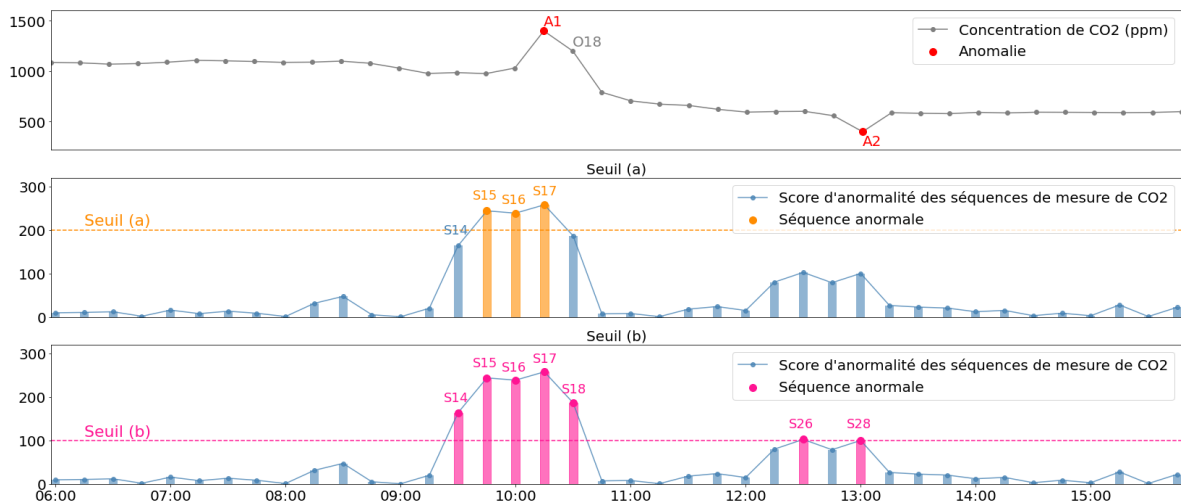


Figure 52 Exemple de détection d'anomalie en utilisant un score d'anormalité de séquences. La ligne orange représente un seuil de décision (a) de 200, tandis que la ligne rose correspond à un seuil (b) de 100. Les anomalies sont notées « A » et les observations sont notées « O ».

5.4. Description, prétraitement des données et génération des ensembles d'évaluation

Cette partie offre un aperçu des variables collectées, des ensembles d'évaluations générés et des diverses transformations appliquées aux données. Elle explique brièvement les particularités et les caractéristiques des données utilisées pour les phases d'entraînement et de test. Enfin, elle présente les transformations appliquées aux données.

5.4.1. Collecte et préparation des données

Les données brutes utilisées dans cette étude consistent en des mesures collectées à un pas de temps de 1 minute à partir de deux emplacements :

- (i) Dans une chambre nord-ouest d'un appartement de type T3 situé dans le centre-ville de Bordeaux (France) sur une période de deux ans (de juillet 2019 à février 2021) (*cf.* chapitre 4, partie 4.4.1). Ces mesures, issues de cet emplacement, seront désignées dans la suite de ces travaux comme provenant de *l'appartement*.
- (ii) Dans une maison située à quelques kilomètres de Bordeaux sur une période légèrement supérieure à un an (de décembre 2018 à février 2020). Les mesures provenant de cet emplacement seront désignées comme provenant de la *maison* pour la suite de ces travaux.

Pour les deux cas d'étude, les données brutes comprennent des mesures de l'ambiance intérieure, telles que la température, l'humidité relative et la concentration en CO₂. De plus, des mesures de consommation électrique sont collectées individuellement sur 16 prises dans la maison (p. ex. six dans la cuisine). La disposition des capteurs pour toutes les mesures relatives à la maison est illustrée dans la Figure 53. Pour l'appartement, le plan est identique à celui proposé dans la Figure 35 de la partie 4.4.1. Des données complémentaires, telles que l'horodatage (minutes, heures et jour de la semaine), sont également générées et ajoutées à l'ensemble des données brutes.

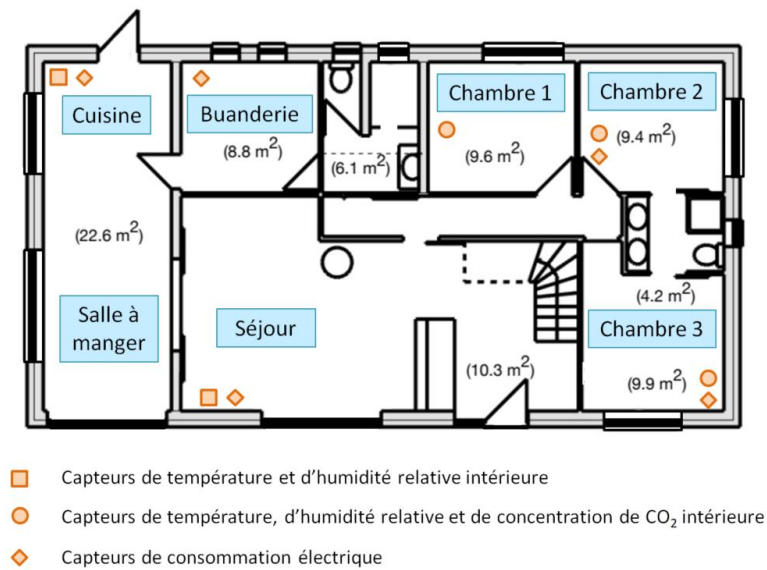


Figure 53 Position des capteurs de la maison.

Sur les trois années de données disponibles, nous avons sélectionné des données avec un faible taux d'anomalies, les avons nettoyées, agrégées à un pas de temps de 15 minutes, puis séparées en une période d'entraînement et une période de test. La période d'entraînement couvre plus d'un an de mesures pour l'appartement, allant de fin juillet 2019 à début octobre 2020, et un an de mesures pour la maison, de décembre 2018 à décembre 2019. Cette période comprend environ 270 000 observations cumulées par variable, à l'exception de la concentration de CO₂, qui en comptabilise environ 160 000.

Nous utilisons deux ensembles d'évaluation distincts. Le premier ensemble est composé de données saines, c'est-à-dire exemptes d'observations anormales, mais ces données seront artificiellement altérées pour évaluer les performances des méthodes sur différents scénarios d'anomalies. Le deuxième ensemble comporte des anomalies réelles (non étiquetées) et est utilisé pour évaluer visuellement les résultats de certaines méthodes sur des anomalies réelles. La composition de ces ensembles varie selon les mesures étudiées.

- Les mesures d'ambiance.** Le premier ensemble est composé de données saines provenant de l'appartement, couvrant environ un mois et demi, de début décembre 2020 à mi-janvier 2021, avec un total d'environ 4 200 points de mesure. Ces données sont en grande partie celles utilisées dans chapitre 4. Nous utilisons également les mesures issues d'un capteur positionné dans le séjour de la maison, bien qu'il présente un grand nombre d'anomalies, pour évaluer visuellement les performances des modèles sur des anomalies réelles. Cet ensemble contient environ deux mois et demi de données, allant de fin décembre 2018 à mi-février 2019, avec un total d'environ 4 800 observations.

- **Les mesures de consommation électrique.** L'ensemble utilisé pour l'évaluation comprend un mois de données saines provenant d'une des prises situées dans la buanderie du bâtiment M, dédiée à cet usage. Ces données ont été collectées de mi-juin à mi-juillet 2019 et totalisent environ 3 000 observations.

5.4.2. Génération des ensembles d'évaluation anormaux

Pour générer les ensembles de données de test avec des anomalies, nous avons varié deux paramètres : l'*intensité de l'anomalie* (cf. partie 5.3.1) et le *taux de contamination* (C). Le taux de contamination fait référence au pourcentage d'observations normales dans un jeu de données qui sont remplacées par des observations anormales. Par exemple, une contamination 2 % signifie qu'un total de 2 % des observations de l'ensemble des données sera contaminé et remplacé par des anomalies. Il s'agit d'un paramètre important à prendre en compte, car il peut avoir une influence sur l'évaluation des performances de détection (Emmott et al., 2016; Fourure et al., 2021; Sharma et al., 2010; Wu and Keogh, 2021).

Nous avons considéré deux taux de contamination : un *faible* et un *fort*. Le premier correspond à un taux de contamination τ_c compris entre 0,5 % et 1,5 %. Le second est inclus dans l'intervalle de 2,5 % à 5 %. Le choix du taux de contamination maximum de 5 % vise à refléter l'une des principales caractéristiques des anomalies, à savoir leur rareté. Ce taux maximal varie généralement entre 1 % et 10 % selon les études (Emmott et al., 2016; Liu et al., 2020; Schmidl et al., 2022), et il affecte également la taille N des anomalies collectives telles que les bruits, les écarts, les écarts plats et les plateaux. Pour illustrer ce point, le Tableau 14 présente en détail les paramètres utilisés pour la contamination de chaque type d'anomalie.

Tableau 14 Paramètres utilisés pour la contamination des jeux de test. M représente le nombre d'observations composant la variable contaminée.

Anomalies	Taux de contamination (C)		Taille (N) $N \in \mathbb{N}$	Nombre (n) $n \in \mathbb{N}$	Intensité
Pic	Faible	→ (0,5 %, 1,5 %)	1	$\tau_c \times M$	[Faible, Forte]
	Fort	→ (2,5 %, 5 %)	1	$\tau_c \times M$	[Faible, Forte]
Bruit	Faible	→ [0,5 %, 1,5 %]	(8, 48)	(1, 3)	[Faible, Forte]
	Fort	→ [2,5 %, 5 %]	(16, 96)	(1, 3)	[Faible, Forte]
Écart	Faible	→ [0,5 %, 1,5 %]	(8, 48)	(1, 3)	[Faible, Forte]
	Fort	→ [2,5 %, 5 %]	(16, 96)	(1, 3)	[Faible, Forte]
Écart plat	Faible	→ [0,5 %, 1,5 %]	(8, 48)	(1, 3)	[Faible, Forte]
	Fort	→ [2,5 %, 5 %]	(16, 96)	(1, 3)	[Faible, Forte]
Plateau	Fort	→ (2,5 %, 5 %)	$\tau_c \times M$	1	-
Changement de niveau	-	-	(1, 6)	1	[Faible, Forte]

Pour un jeu de M observations, le nombre de pics générés dépend du taux de contamination choisi. Le taux de contamination est sélectionné de manière aléatoire selon une loi uniforme, dont la plage dépend du taux fixé (faible ou fort). Par exemple, si le taux de contamination fort est fixé à 3 %, alors 3 % des observations normales seront transformées en pics.

En ce qui concerne les anomalies collectives telles que les bruits, les écarts, les écarts plats et les plateaux, le processus est inversé. Un nombre aléatoire de séquences est généré, variant de 1 à 3, de sorte à ce que la somme de leurs tailles corresponde au taux de contamination spécifié. Cette procédure est répétée jusqu'à ce que cette condition soit satisfaite. La taille maximale des anomalies (N) est fixée pour correspondre à 24 heures d'observations, tandis que la taille minimale varie de 2 à 4 heures. Bien que certains travaux comme ceux de Sharma et al. (2010) puissent proposer des tailles d'anomalies de plus de 1000 observations, notre choix est motivé par le fait qu'une anomalie de 24 heures représente environ 2,5 % des données dans notre cas d'application.

Contrairement aux autres types d'anomalies, les plateaux ne dépendent que d'un seul taux de contamination. En effet, certaines variables, telles que la température, l'humidité absolue intérieure et la consommation électrique, présentent une variabilité moindre par rapport à d'autres variables, comme la concentration de CO₂. Par conséquent, l'anomalie de plateau se distingue des observations normales principalement par sa taille

particulièrement importante. Ainsi, nous avons choisi de ne considérer qu'une seule anomalie de taille unique, déterminée en fonction d'un taux aléatoire sélectionné dans la plage de 2,5 % à 5 %.

Enfin, les anomalies de changements de niveau, caractérisées par une variation soudaine, ne dépendent pas du taux de contamination.

La combinaison de l'intensité et de la contamination conduit à la génération des ensembles d'évaluation suivants pour chaque type d'anomalie :

- **Quatre** ensembles d'évaluation anormaux pour chaque anomalie de pic, de bruit, d'écart et d'écart plat, car elles sont toutes caractérisées par deux niveaux d'intensité et de contamination différents.
- **Deux** ensembles d'évaluation anormaux pour l'anomalie de changement de niveau, car elle est caractérisée uniquement par deux niveaux d'intensité différents.
- **Un** ensemble d'évaluation anormal unique pour les anomalies de plateau, car elles ne présentent ni variation d'intensité ni de contamination.

5.4.3. Transformation et association des données

Pour toutes les méthodes de détection d'anomalies utilisées, plusieurs transformations et associations de données ont été appliquées pour faciliter l'apprentissage ou la détection des anomalies (Bengio et al., 2014). Cette approche est similaire à celle proposée pour le chapitre 4, mais elle est moins complexe en raison du nombre de modèles et de jeux de données évalués. Pour simplifier, les associations de transformations ont été limitées à deux composantes par variable. Voici un résumé de toutes les transformations appliquées, telles qu'elles sont répertoriées dans le Tableau 15.

Les transformations « *moyenne sur fenêtre* » et « *écart type sur fenêtre* », notées respectivement $\text{moyenne}_{\text{Fenêtre}}$ et $\text{écart type}_{\text{Fenêtre}}$, sont inspirées des travaux d'Araya et al. (2016). Elles servent à caractériser une séquence d'observations de taille N en calculant la moyenne ou l'écart type de l'ensemble des observations qui la composent. Ces transformations permettent de résumer les informations contenues dans une séquence en une seule observation. Elles sont appliquées pour agréger des fenêtres de 12, 24, 48 et 96 observations (dans le cas des modèles de voisinage IF, KNN et LOF), ou adaptées à la taille de fenêtre existante (dans le cas des modèles de prédiction).

Les autres transformations, telles que la dérivée, la MME, le résidu ou la tendance de la STL, notés respectivement $\text{résidu}_{\text{STL}}$ et $\text{tendance}_{\text{STL}}$, sont identiques à celles décrites dans la partie 4.4.3.

Tableau 15 Présentation des transformations appliquées.

Transformations	Explications
Moyenne sur fenêtre	Moyenne de toutes les observations d'une fenêtre de taille N ($\in \mathbb{N}$)
Écart type sur fenêtre	Écart type de toutes les observations d'une fenêtre de taille N ($\in \mathbb{N}$)
Dérivée	Dérivée de l'ensemble des observations d'une variable
Résidu de la STL	Résidu obtenu par application de la STL pour l'ensemble des observations d'une variable
Tendance de la STL	Tendance obtenue par application de la STL pour l'ensemble des observations d'une variable
MME	Lissage par MME de l'ensemble des observations d'une variable

5.5. Résultats

Cette section présente tous les résultats obtenus dans cette étude. Pour chaque variable, à savoir la température de l'air, l'humidité relative, la concentration de CO₂ intérieure et la consommation électrique, différentes anomalies de diverses intensités ont été introduites, avec des taux de contamination variables. Pour évaluer les performances des modèles sur un type d'anomalie spécifique, chaque jeu de test a été contaminé par un seul type d'anomalie à la fois, avec une intensité et un taux de contamination prédéfinis, pouvant être faibles ou forts. Au total, en tenant compte des différentes combinaisons et transformations employées, entre 400 et 600 modèles différents ont été utilisées pour chaque ensemble de données contaminées. Par conséquent, les performances sont présentées de manière indépendante pour chaque variable mesurée : température d'air, humidité relative, concentration de CO₂ et consommation électrique. Cependant, une partie de l'analyse abordera les composantes observables et communes à l'ensemble de ces variables. Enfin, une synthèse des résultats sera présentée pour résumer les principales conclusions de cette étude.

Tous les résultats présentés ci-après sont évalués en utilisant l'indice AUC-ROC. Cela nous permettra de quantifier et de discuter de l'impact des transformations et combinaisons de mesures sur les performances la détection d'anomalies. De plus, cela nous permettra de discuter des différentes méthodes d'évaluation du score d'anormalité pour les modèles concernés, notamment MAE, MSE, DTW et DMA. Nous examinerons aussi l'impact du choix de la taille des fenêtres glissantes employées.

Il est important de noter que l'objectif de ces travaux ne se limite pas à évaluer uniquement les performances brutes des modèles en matière de détection d'anomalies. Ces performances sont pertinentes mais spécifiques à ce cas d'étude et ne peuvent pas

être généralisées sans prudence. Par conséquent, nous adoptons une approche statistique pour rendre compte des tendances observées. Cette approche comparative prend en considération l'ensemble des spécificités relatives aux variables, aux anomalies et aux modèles, comme cela a été présenté jusqu'à présent.

5.5.1. Contexte sur l'entraînement et l'exploitation des modèles

Dans le contexte de l'entraînement et de l'exploitation des modèles semi-supervisés basés sur des réseaux de neurones, nous avons suivi une approche similaire à celle présentée dans l'introduction de la partie 4.5. Pour les réseaux prédictifs, à savoir Dense-p et LSTM-p, nous avons entraîné deux modèles en alternance. Le modèle présentant la meilleure performance moyenne de prédiction au cours des dernières itérations sur le jeu de données de validation a été retenu pour l'évaluation.

Initialement, une approche de sélection similaire a été employée pour l'ensemble des modèles de reconstruction. Cependant, les performances des modèles se sont révélées très stables, ce qui a conduit au choix de n'entraîner qu'un seul modèle. Dans les deux cas, le nombre d'itérations était beaucoup plus faible que celui appliqué dans le chapitre 4 (10 itérations). Cela s'explique par deux facteurs : la stabilité des prédictions et des reconstructions réalisées, ainsi que le temps de calcul nécessaire plus important. La stabilité des résultats peut être attribuée à plusieurs facteurs, notamment la quantité bien plus importante de données employées, avec environ 160 000 à 270 000 observations pour une seule variable par rapport à 12 000 précédemment. De plus, les modèles eux-mêmes et la tâche effectuée jouent un rôle essentiel dans cette stabilité.

Concernant les temps de calculs, ils sont illustrés dans la Figure 54, qui présente le temps d'exécution moyen par observation pour l'ensemble des modèles semi-supervisés et non supervisés. Pour les modèles semi-supervisés (en bleu), ce temps comprend à la fois le temps d'entraînement et le temps d'application par modèle, ramené à une observation pour une combinaison ou une transformation de variable unique du jeu d'entraînement. Dans ce cas, le temps d'application est négligeable par rapport au temps d'entraînement ($< 0,1$ % en moyenne). Le temps d'évaluation n'est pas comptabilisé dans cette partie.

Pour les modèles non supervisés (en orange), le temps d'exécution représente le temps d'application par modèle, ramené à une observation pour une variable unique du jeu d'évaluation. Nous observons qu'à l'exception de l'ARIMA, tous les modèles semi-supervisés sont parmi les plus lents, avec un temps d'exécution moyen d'environ 6 ms par observation, soit environ 100 fois plus long que celui des modèles non supervisés, qui est d'environ 50 μ s.

Ce temps d'exécution a eu un impact conséquent sur le choix du nombre d'itérations des modèles, car un seul modèle d'AE nécessite environ 20 à 180 minutes d'entraînement. Ce temps varie en fonction de la convergence des modèles et de l'arrêt induit par la fonction d'arrêt précoce. À titre d'exemple, l'entraînement des seuls modèles de reconstruction pour toutes les variables a pris environ 450 heures de calcul cumulées, soit 19 jours. De plus, l'utilisation de la DTW pour le calcul du score d'anormalité a eu un impact non négligeable sur la durée de la phase d'évaluation. Dans ce cas, selon la taille de la fenêtre considérée, son utilisation peut multiplier le temps d'application total par 5 pour des fenêtres de 12 observations, allant jusqu'à le multiplier par 50 pour des fenêtres de 96 observations. En comparaison, l'emploi de la MSE, de la MAE et de la DMA nécessite en moyenne 0,36 ms par observation, tandis que l'utilisation de la DTW en nécessite 6,6 ms. De plus, l'ajout de combinaisons ou de transformations n'a pas semblé avoir d'impact significatif sur le temps d'entraînement des modèles, à l'exception de potentiellement accélérer la convergence dans certains cas.

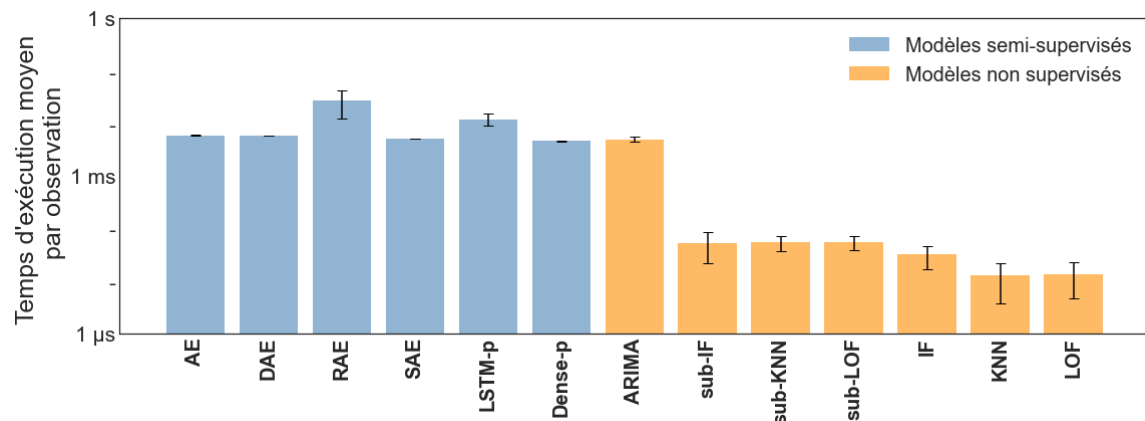


Figure 54 Temps d'exécution moyen par observation pour chaque modèle appliqué sur une combinaison ou transformation de variable unique (échelle logarithmique). Les modèles semi-supervisés sont représentés par une couleur bleue et les modèles supervisés par de l'orange. L'écart type est présenté par une barre noire. Inspirée de Schmidl et al. (2022).

De plus, parmi les modèles semi-supervisés, ceux construits au moyen de neurones récurrents LSTM, les RAE et les LSTM-p, sont jusqu'à 4 fois plus lents, avec un temps d'exécution moyen de 26 ms et 12 ms par observation. Ces modèles sont, par ailleurs, les seuls qui présentent une influence marquée par la taille N des fenêtres employées, laquelle peut aller jusqu'à tripler les temps d'exécution pour les plus grandes tailles ($N = 96$). Pour les autres modèles, quelle que soit la méthode d'apprentissage, la taille de fenêtre glissante employée semble présenter un impact à peine marqué, voire inexistant. Pour les modèles non supervisés, l'ARIMA est particulièrement lente, avec un temps d'exécution moyen de 6 ms. Cet effet peut s'expliquer en grande partie par le processus d'actualisation appliqué. Globalement, sans *a priori* sur les performances obtenues et en restant dans le cadre défini par ces travaux, les méthodes non supervisées de voisinage présentent l'avantage d'être les plus rapides à exécuter et, par définition, moins

demandeuses en données. Elles doivent cependant être appliquées régulièrement sur les données, contrairement aux modèles d'apprentissage semi-supervisés qui peuvent être stockés.

Concernant le stockage, le Tableau 16 résume les poids minimum, moyen et maximum de l'ensemble des modèles semi-supervisés qui ont été formés et enregistrés pour ces travaux. Cette information revêt de l'intérêt dans des cas d'application réels, car elle définit les exigences de stockage et d'utilisation potentielles des modèles. Les poids minimum correspondent généralement aux modèles formés avec une seule variable ou d'une seule transformation, utilisant également les plus petites tailles de fenêtre ($N = 12$). En revanche, les poids maximum correspondent aux modèles combinant le plus grand nombre de variables ou de transformations, jusqu'à six dans ce cas, ainsi que les plus grandes tailles de fenêtre ($N = 96$).

Il est notable que, de manière générale, les modèles constitués de réseaux de neurones récurrents LSTM, les RAE, les RED et les LSTM-p, figurent parmi les plus lourds au sein de leurs familles de méthodes respectives. Cependant, il est important de noter que le nombre de combinaisons ou de transformations n'a pas d'impact significatif sur leur poids, qui reste relativement stable. Cette différence est surtout prononcée pour les poids moyens et minimums, étant jusqu'à 10 fois plus élevés pour ces modèles par rapport aux autres variantes.

Tableau 16 Poids minimum, moyen et maximum de l'ensemble des modèles semi-supervisés entraînés, pour les familles de reconstruction (■) et de prédiction (▲). Les poids sont affichés en Mégaoctets (Mo).

Modèles	Poids minimum (Mo)	Poids moyen (Mo)	Poids maximum (Mo)
(■) AE et ED	0,316	0,598	2,347
(■) DAE et DED	0,316	0,588	2,346
(■) RAE et RED	2,968	2,973	2,999
(■) SAE et SED	0,664	0,902	2,555
(▲) Dense-p	0,080	0,270	1,241
(▲) LSTM-p	0,957	0,967	0,990

De plus, la Figure 55 illustre l'évolution du poids moyen de l'ensemble des modèles en fonction de la taille de fenêtre employée, tout en prenant en compte le nombre de combinaisons et de transformations, ce dernier étant approximé par l'écart type. Cette figure met en évidence une augmentation linéaire du poids des modèles en fonction de la taille de fenêtre utilisée pour toutes les combinaisons implémentées, à l'exception des modèles basés sur les réseaux de neurones récurrents. Il semble que la taille de la fenêtre mobile ait plus d'impact sur le poids des modèles que le nombre de combinaisons de

variables employées lorsque la taille de la fenêtre est faible ($N < 48$), comme l'indique l'écart type limité associé à chaque modèle et taille de fenêtre.

Cependant, une tendance inverse se manifeste lorsque la taille de la fenêtre augmente, avec un écart-type plus important observé dans ce cas.

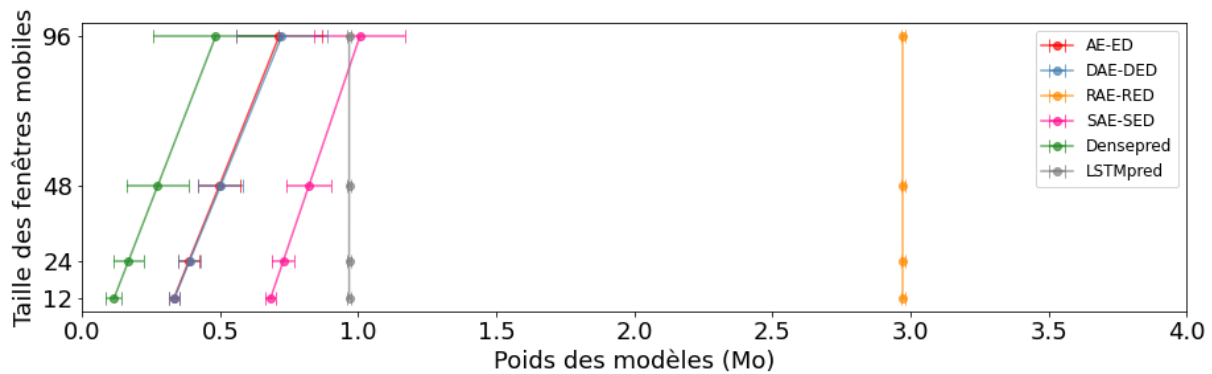


Figure 55 Évolution du poids moyen des modèles selon la taille des fenêtres de reconstruction et d'observations utilisées. L'impact du nombre de combinaisons et de transformations employées peut être estimé grâce à l'écart type.

5.5.2. Le cas des mesures de température d'air intérieur

5.5.2.1. Contexte

Tous les modèles proposés ont été appliqués aux mesures de température d'air intérieur et évalués sur différents types d'anomalies, telles que les pics, les bruits, les écarts, les écarts plats et les plateaux. Au total, 17 ensembles d'évaluation ont été créés, à partir de l'ensemble d'observations de base saines, qui est illustré dans la Figure 56. L'encadré intérieur présente un zoom sur une période de cinq jours parmi les 45 que compte le jeu de données. Cette vue rapprochée permet d'observer plus en détail la dynamique des données. Les brusques chutes négatives correspondent à des moments où les fenêtres ont été ouvertes, et certains de ces exemples sont annotés O1, O2, O3 et O4 (comme expliqué dans la partie 4.4.2). La période en gris représente une période d'absence d'occupants, caractérisée par une baisse de la température de 2 °C. Ces ouvertures de fenêtres et les changements de tendance dans les données sont des événements rares et intéressants à étudier afin d'évaluer leur impact sur les performances des méthodes employées dans ces travaux.

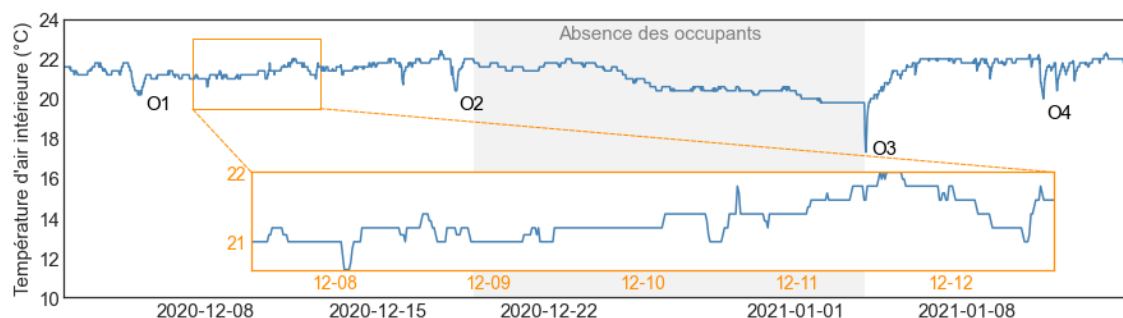


Figure 56 Mesures de température d'air intérieure saine avant contamination.

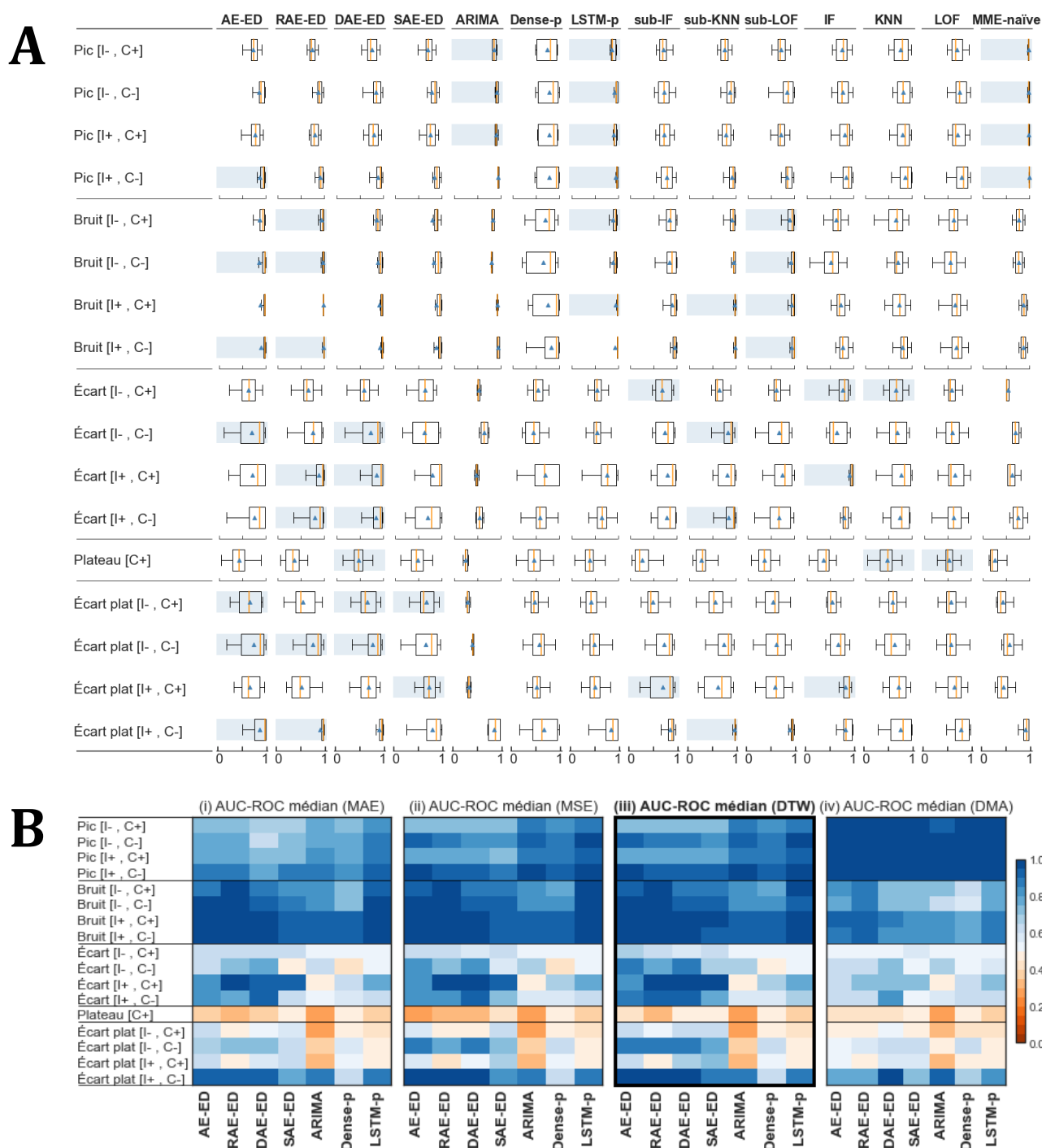
5.5.2.2. Performances des modèles pour chaque ensemble d'évaluation

Les résultats de l'indice AUC-ROC obtenus sont présentés sous forme de boîtes à moustaches pour tous les modèles, types d'anomalies et niveaux d'intensité ou taux de contamination dans la Figure 57 (A). Par exemple, « Pic [I+ ; C-] » désigne l'ensemble d'évaluation composé d'anomalies de type pics, d'une intensité forte (I+) et d'un taux de contamination faible (C-). Les trois meilleurs modèles pour chaque ensemble d'évaluation, en matière de médiane de l'indice AUC-ROC, sont mis en évidence en bleu. Pour simplifier, les modèles Encodeur-Décodeur (ED) sont regroupés avec leurs versions d'Auto-Encodeur respectives (AE). Par exemple, les Auto-Encodeur et les Encodeur-Décodeur récurrents sont notés « RAE-ED ».

Dans cette figure, l'indice AUC-ROC des modèles de reconstruction et de prédiction a été obtenu en utilisant les résultats du score d'anormalité calculé par la DTW. Les scores médians de l'indice AUC-ROC obtenus pour chaque mesure de similarité (MSE, MAE, DTW et DMA) sont présentés dans la Figure 57 (B) sous forme de cartes thermiques. Bien que les différences entre les mesures de similarité semblent minimales, l'approche basée sur la DTW produit en moyenne des résultats de 5 à 10 % supérieurs à ceux obtenus avec la MSE ou la MAE. Cependant, une distinction notable apparaît lorsque ces trois mesures sont comparées à la DMA, montrant une nette amélioration des performances de cette dernière pour les anomalies ponctuelles (pics). Cependant, cette amélioration semble se faire au détriment des anomalies collectives, avec une baisse d'environ 10 % de la médiane des indices AUC-ROC.

En observant la Figure 57 (A), notons tout d'abord que pour les pics, les modèles de MME-naïve et de LSTM-p se distinguent par leur performance et leur stabilité, affichant les meilleurs résultats avec des médianes et des moyennes des indices AUC-ROC très proches de 1. En revanche, les autres modèles, notamment ceux de la famille de reconstruction ainsi que les modèles de voisinage séquentiels (sub-IF, sub-LOF et sub-KNN), obtiennent généralement des résultats inférieurs, à quelques exceptions près.

Ce constat est particulièrement flagrant pour les taux de contamination élevés (C+). Cette différence s'explique exclusivement par l'approche séquentielle employée.



Contrairement aux modèles de prédiction, qui prédisent au maximum une séquence de 3 heures, ou aux modèles de voisinage classiques (IF, KNN et LOF), qui traitent chaque

observation individuellement, les modèles de reconstruction ou de voisinage séquentiels appliquent des fenêtres de taille conséquente, pouvant inclure jusqu'à 24 heures d'observations. Ainsi, lorsque plusieurs anomalies sont présentes dans une fenêtre, l'ensemble des observations normales situées entre elles sont également considérées comme anormales. Cela a pour conséquence de détériorer considérablement les performances de détection pour ces modèles dans de tels cas.

En effet, l'approche par score d'anormalité d'observation employant la DMA, qui permet de passer d'une approche en séquence à une approche par observation, présente de meilleurs résultats médians que les approches basées sur des séquences (B.iv). Cependant, cet inconvénient semble se transformer en un avantage considérable pour les anomalies collectives de bruit. Nous constatons une amélioration notable des résultats des modèles précédents, qui deviennent parmi les plus performants, quel que soit le taux de contamination et donc la taille de l'anomalie.

Les anomalies d'écarts et d'écarts plats, en particulier celles de faible intensité (I-), semblent être parmi les plus difficiles à détecter, tout comme les anomalies de plateau. Ce constat est d'autant plus visible en examinant les médianes (i), les troisièmes quartiles (ii) et les maximums (iii) des indices AUC-ROC, comme illustré dans la Figure 58, qui traduisent les performances des modèles pour chaque ensemble d'évaluation. Ces trois types d'anomalies présentent les pires résultats, avec des médianes ou des moyennes proches de 0,5, voire inférieures, ce qui signifie que les modèles ont des performances comparables à celles d'un modèle aléatoire.

Plusieurs explications sont possibles pour ces résultats. Tout d'abord, les faibles variations observées dans les mesures de température intérieure telles que montrées dans la Figure 56, associées à la précision des capteurs et à l'agrégation des données, pourraient rendre difficile la détection de séquences d'observations identiques ou très similaires sur plusieurs heures. Une telle séquence pourrait ne pas être considérée comme une anomalie compte tenu des données normales observées habituellement.

Cependant, une autre explication pourrait être liée à la taille des fenêtres glissantes utilisées. Parmi tous les modèles séquentiels employés, les modèles prédictifs semblent être les moins performants pour détecter les écarts et les écarts plats, en particulier la méthode ARIMA, comme le montrent les encadrés blancs (i, ii et iii) de la Figure 58. Cette performance médiocre pourrait s'expliquer par la taille des fenêtres de prédiction employées (resp. 1 et 3 heures), qui est plus courte que celle des modèles de reconstruction et des modèles séquentiels de voisinage (resp. 3, 6, 12 et 24 heures). Cette hypothèse est soutenue par l'amélioration notable des performances de détection observée pour le troisième quartile de ces deux familles de modèles, contrairement aux modèles prédictifs.

Enfin, il est également possible que les mauvaises performances de l'ARIMA soient dues à une fréquence d'actualisation trop courte. Dans le cas des anomalies collectives, cela conduit le modèle à s'actualiser trop fréquemment sur un comportement anormal, notamment sur les anomalies de grande taille. Cette actualisation excessive entraîne des prédictions qui présentent un score d'anormalité faible sur les observations anormales suivantes, mais un score plus élevé pour les observations normales immédiatement après l'anomalie. En conséquence, les premières sont considérées comme normales, tandis que les secondes sont à tort considérées comme anormales.

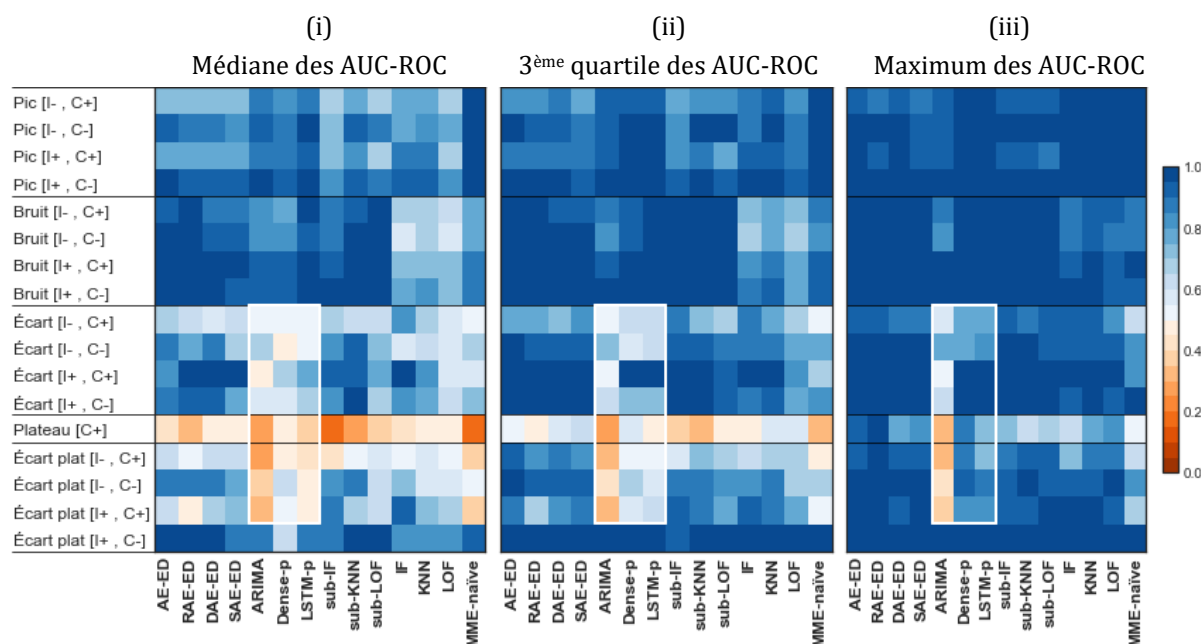


Figure 58 Carte thermique des médianes (i), des troisièmes quartiles (ii) et des maximums (iii) des indices AUC-ROC de l'ensemble des modèles, classés par type d'anomalies et appliqués aux mesures de température d'air intérieure. L'encadré blanc met en évidence les performances des modèles prédictifs sur trois types d'anomalies spécifiques.

En résumé, la plupart des méthodes employées, à l'exception des modèles prédictifs et de la MME-naïve, présentent des performances globalement satisfaisantes pour la plupart des types d'anomalies, avec au moins 25 % de leurs résultats affichant un indice AUC-ROC supérieur à 0,7 voire 0,8 (ii). De plus, de nombreux modèles montrent qu'au moins un d'entre eux atteint un indice AUC supérieur à 0,9, voire très proche de 1 (iii).

Les modèles d'Auto-Encodeur et d'Encodeur-Décodeur, à l'exception des SAE-ED, semblent être des choix polyvalents pour les modèles semi-supervisés. Pour les méthodes non-supervisées, il est plus délicat de faire des distinctions, mais les modèles séquentiels de voisinage tels que la sub-IF et la sub-KNN semblent à la fois performants et stables par rapport aux autres modèles.

Cependant, il est important de noter que les modèles prédictifs et la MME-naïve montrent des performances moins bonnes pour la détection de la plupart des anomalies collectives (iii. encadré blanc). Par ailleurs, concernant les ouvertures de fenêtres évoquées en introduction de cette partie, il est intéressant de noter que plusieurs de ces ouvertures sont souvent considérées comme anormales et ce, dès les premiers seuils de détection. Ce phénomène est particulièrement prononcé pour l'ouverture O3, présentée dans la Figure 56, qui est souvent détectée avant la plupart des autres anomalies par l'ensemble des modèles. En revanche, le changement de tendance provoqué par l'absence des occupants ne semble pas poser de problème de détection pour les modèles.

5.5.3. Le cas des mesures d'humidité relative d'air intérieur

Les mesures d'humidité relative intérieures ont également été évaluées en utilisant l'ensemble des modèles proposés. Différents types d'anomalies, tels que les pics, les bruits, les écarts, les écarts plats et les plateaux, ont été pris en compte, ce qui a conduit à la création de 17 ensembles d'évaluation. Le jeu de données de base, qui a servi de référence pour la création des ensembles anormaux, est présenté dans la Figure 59. Un encadré intérieur offre un aperçu détaillé de la dynamique des données au cours d'une période de cinq jours, parmi les 45 composant l'ensemble. Les pics significatifs, à la fois négatifs et positifs, correspondent à des moments où les fenêtres ont été ouvertes, par temps sec pour les premiers et par temps pluvieux pour les seconds. Quelques exemples de ces ouvertures sont annotés O1, O2, O3 et O4 (cf. partie 4.4.2). De plus, le changement de dynamique dans les données provoqué par l'absence des occupants est clairement visible dans la zone grise de la figure.

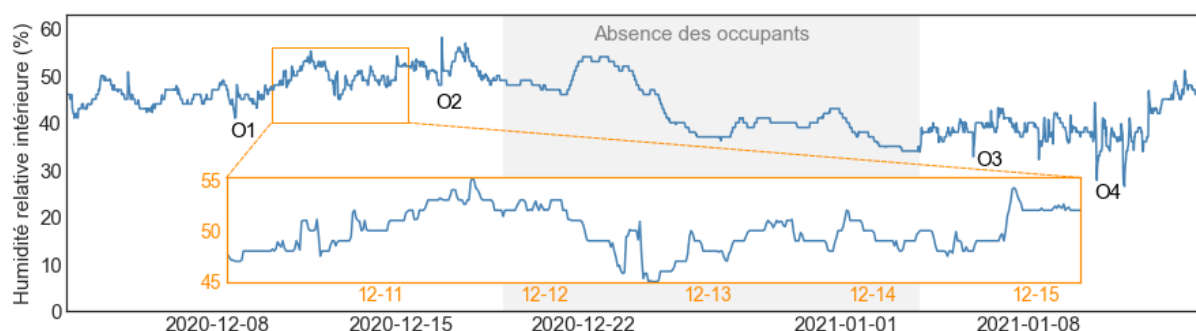


Figure 59 Mesures de l'humidité relative intérieure saine avant contamination.

Les tests réalisés sur l'humidité relative ont produit des résultats très proches de ceux obtenus pour la température, tant en matière de performances que de tendances observées. Par conséquent, nous ne les aborderons pas spécifiquement. Un aperçu de ces résultats est donné dans la Figure 60, qui présente des cartes thermiques montrant les médianes (i), les troisièmes quartiles (ii) et les maximums (iii) des indices AUC pour

chaque modèle et par type d'anomalie. Des ressources similaires aux précédentes, comprenant des structures et des approches similaires sont consultables en Annexe A. Ces performances étaient attendues car les deux variables, la température et l'humidité relative, présentent des tendances et des dynamiques globales similaires. Cela peut s'expliquer par l'influence de la température de l'air sur l'humidité relative, qui est prise en compte dans l'expression de la pression de vapeur saturante utilisée pour calculer l'humidité relative. Ces résultats renforcent donc les tendances observées précédemment. Une étude équivalente sur des mesures d'humidité absolue constituerait donc un cas pertinent à considérer. Cependant, compte tenu du temps nécessaire pour entraîner des modèles semi-supervisés sur une autre variable (*cf.* partie 5.5.1), elles n'ont pu être traitées dans ces travaux.

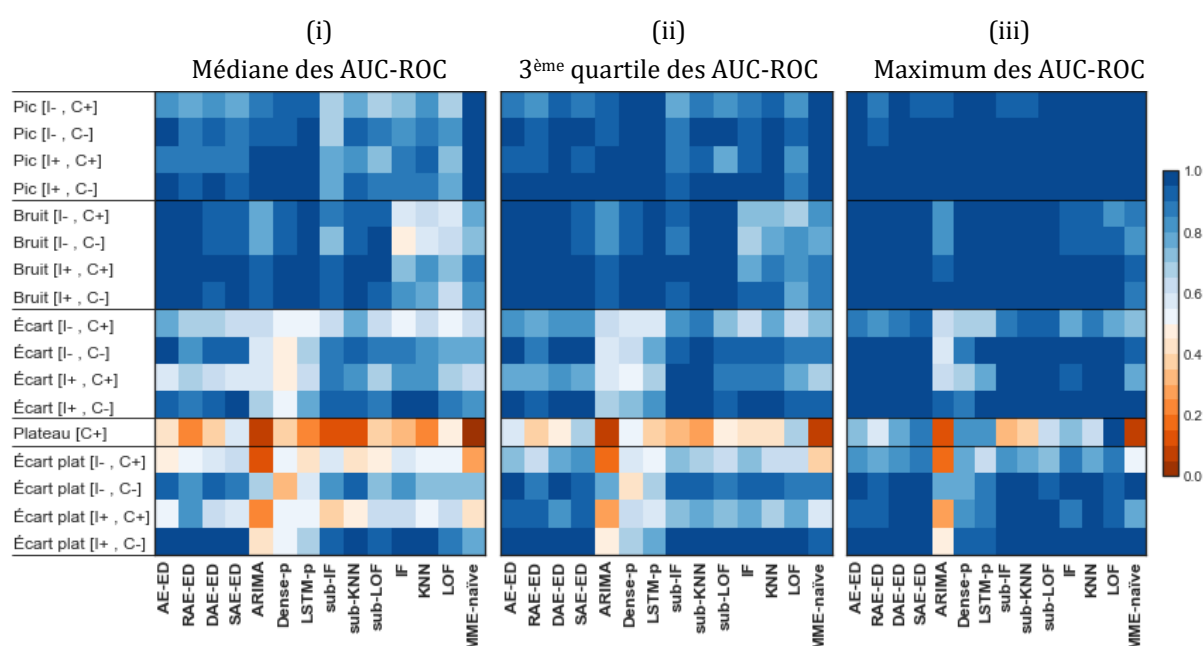


Figure 60 Carte thermique des médianes (i), des troisièmes quartiles (ii) et des maximums (iii) des indices AUC-ROC de l'ensemble des modèles, classés par type d'anomalies et appliqués aux mesures d'humidité relative intérieure. La DTW a été utilisée comme mesure de similarité pour les modèles de reconstruction et de prédiction.

5.5.4. Les cas des mesures de concentration de CO₂ intérieure

5.5.4.1. Contexte

Pour la mesure de concentration de CO₂ intérieure, l'ensemble des modèles proposés ont également été appliqués. Au total, 17 ensembles d'évaluation ont été créés en injectant des anomalies de plusieurs types, notamment les pics, les bruits, les écarts, les écarts plats et les plateaux. En raison de la plus grande variabilité de la concentration de CO₂ par rapport aux autres variables, l'écart-type utilisé pour générer les anomalies (noté σ_x dans la partie 5.3.1) a été divisé par deux pour cette mesure. Ce choix a été effectué pour que

les anomalies étudiées, d'une variable à l'autre, présentent des intensités et donc des impacts sur les données aussi similaires que possible.

Le jeu d'observations servant de base saine pour la création des ensembles anormaux est présenté dans la Figure 61. L'encadré intérieur montre une période de cinq jours sur les 45 jours le composant, afin d'avoir un aperçu plus détaillé de la dynamique des données. Comme précisé dans le chapitre 4 (*cf.* partie 4.4.2), cette mesure permet de distinguer deux tendances différentes, celle de l'occupation et celle de l'inoccupation (représentées en gris). Ces tendances sont importantes pour évaluer l'impact de tels phénomènes, propres à certaines mesures, sur les performances des méthodes employées.

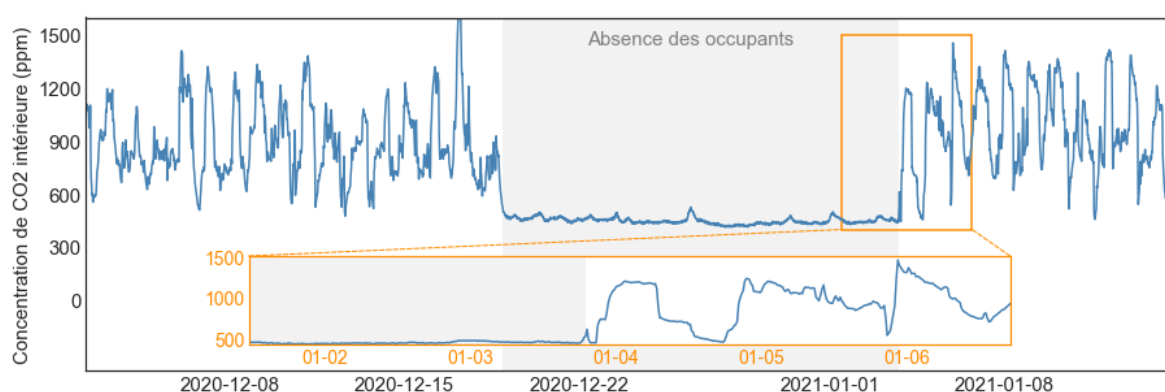


Figure 61 Mesures de concentration de CO₂ intérieure saine avant contamination.

5.5.4.2. Performances des modèles pour chaque ensemble d'évaluation

L'ensemble des résultats obtenus la mesure de concentration de CO₂, présentés de façon similaire à ceux de la Figure 57 pour la température d'air intérieure, sont consultables en Annexe B. Pour simplifier, nous examinerons uniquement les trois cartes thermiques de la Figure 62, qui illustrent les médianes (i), les troisièmes quartiles (ii) et les maximums (iii) des indices AUC-ROC.

La nomenclature suivie est la même que celle présentée dans la partie précédente. Pour cette figure, les indices AUC-ROC des modèles de reconstruction et de prédiction ont été obtenus en utilisant les résultats provenant du score d'anormalité calculé par la DTW. Les tendances observées pour les scores d'anormalité, en utilisant les mesures de MSE, de MAE, de DTW et de DMA, sont similaires à celles obtenues pour la température.

Globalement, il est important de noter une légère baisse des performances générales par rapport à l'application sur les mesures de température et d'humidité relative intérieures. Pour les jeux anormaux composés de pics et de bruits, cela se traduit par une baisse d'environ 10 à 20 % de la médiane de l'ensemble des indices AUC-ROC, ainsi qu'une

variabilité accrue des résultats. Cependant, les modèles les plus performants restent les mêmes pour les pics, avec à nouveau la MME-naïve et les LSTM-p. Un constat similaire peut être formulé pour les bruits, pour lesquels les meilleurs modèles sont essentiellement les modèles de reconstruction et de prédiction, à l'exception à nouveau de l'ARIMA.

Ces résultats moins satisfaisants peuvent s'expliquer par la variabilité nettement plus importante des mesures de concentration de CO₂ comparée aux mesures de température ou d'humidité relative. De plus, des tendances similaires aux variations de contamination, telles que celles observées sur les mesures de température pour les mêmes types d'anomalies, sont observées. En effet, une contamination élevée (C+) semble détériorer les performances de détection des pics.

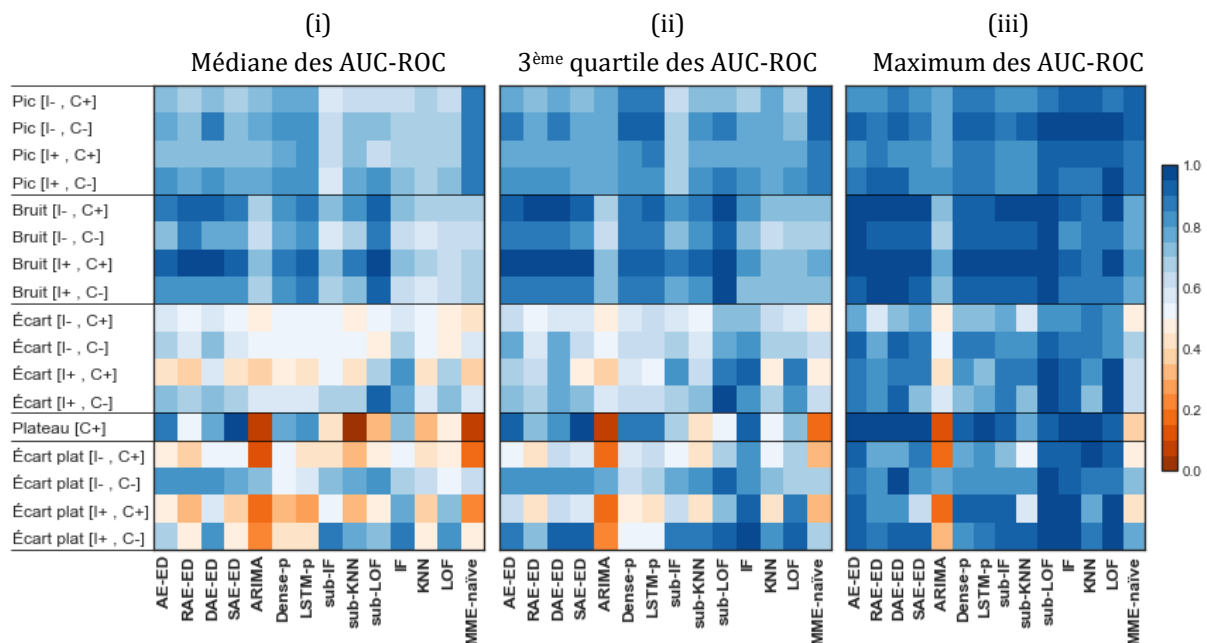


Figure 62 Carte thermique des médianes (i), des troisièmes quartiles (ii) et des maximums (iii) des indices AUC-ROC de l'ensemble des modèles, classés par type d'anomalies et appliqués aux mesures de concentration de CO₂. La DTW a été utilisée comme mesure de similarité pour les modèles de reconstruction et de prédiction.

Pour les autres types d'anomalies collectives, il est clair que les performances chutent de façon drastique, avec plus du double de modèles ayant une médiane proche ou inférieure à 0,5 (i) par rapport aux observations faites sur les autres mesures. De façon équivalente, les résultats pour le troisième quartile et le maximum sont bien plus bas, avec une diminution d'environ 0,2 de l'indice l'AUC-ROC. En général, indépendamment de leur famille ou de leur méthode d'apprentissage, les modèles semblent avoir du mal à détecter des anomalies sur la variable CO₂.

Ce constat est d'autant plus fort pour les modèles de reconstruction, qui semblent ici produire des résultats inférieurs aux méthodes classiques de voisinage. Cette tendance est en contradiction avec ce qui a été observé précédemment pour la température et l'humidité. De plus, les résultats de deux ensembles d'évaluation semblent incohérents et suscitent des doutes quant à la présence d'un facteur externe pouvant potentiellement expliquer cette baisse d'efficacité généralisée.

En effet, les performances de détection obtenues par les modèles de reconstruction, de prédiction et de voisinage séquentiel pour le jeu anormal « Écart [I+ ; C+] » sont similaires à celles observés pour le même type d'anomalie, mais avec une influence plus faible à contamination équivalente, « Écart [I- ; C+] ». Des exceptions similaires sont également visibles en comparant les performances obtenues sur le jeu « Écart plat [I+ ; C-] » à celles du jeu « Écart plat [I- ; C-] », pour lesquels la tendance devrait être inversée. Ces résultats contredisent ce qui peut être observé dans les parties précédentes et les performances des modèles classiques de voisinage IF, KNN et LOF.

L'étude approfondie a révélé que des incohérences étaient liées à la fois au positionnement des anomalies et au score d'anormalité. Contrairement aux mesures de température pendant la même période, la concentration de CO₂ présente deux tendances marquées, avec occupation et sans occupation, comme illustré dans la Figure 61. Les modèles de reconstruction (ou de prédiction) ont tendance à reconstruire (ou prédire) avec plus d'erreurs la période d'occupation du fait de la plus grande variabilité des observations. Ces observations dépendent de plusieurs facteurs, tels que la position de l'occupant par rapport au capteur, le type d'activité ou les mouvements d'air, et sont par conséquent plus stables sur de longues périodes d'inoccupation.

Ainsi, ces modèles « détectent » toutes les anomalies mais celles positionnées pendant la période d'inoccupation ont régulièrement un score d'anormalité inférieur à celui de plusieurs observations normales de la période d'occupation. Par exemple, cela peut se produire après une baisse soudaine de la concentration de CO₂ suite à l'ouverture d'une fenêtre. Cette complexité dans la dynamique de la concentration de CO₂ peut expliquer les résultats contradictoires et la baisse d'efficacité observée pour les modèles de reconstruction et de prédiction par rapport aux méthodes de voisinage classiques.

5.5.4.3. L'influence des changements des tendances sur les performances des modèles

Afin de vérifier ce constat, le jeu d'observations original a été séparé en deux ensembles distincts. Le premier ensemble se compose de 15 jours de données lorsque les occupants étaient présents, tandis que le deuxième ensemble comprend les 15 jours où ils étaient absents. En conséquence, 18 nouveaux ensembles anormaux ont été créés en suivant les mêmes procédures que celles utilisées pour le jeu de données original, en ce qui concerne

les anomalies d'écart, d'écart plat et de plateau. Pour faciliter la comparaison entre le jeu de données original et les nouveaux jeux, les taux de contamination ont été ajustés pour correspondre à ceux initialement utilisés. Les résultats de la tâche de détection sont résumés en matière de médiane (i), troisième quartile (ii) et maximum (iii) de l'indice AUC-ROC dans la Figure 63. Les ensembles anormaux originaux, composés des deux périodes d'occupation et d'inoccupation (A), sont utilisés pour évaluer les différences résultant de la séparation des périodes d'occupation (B) et de l'absence prolongée des occupants (C).

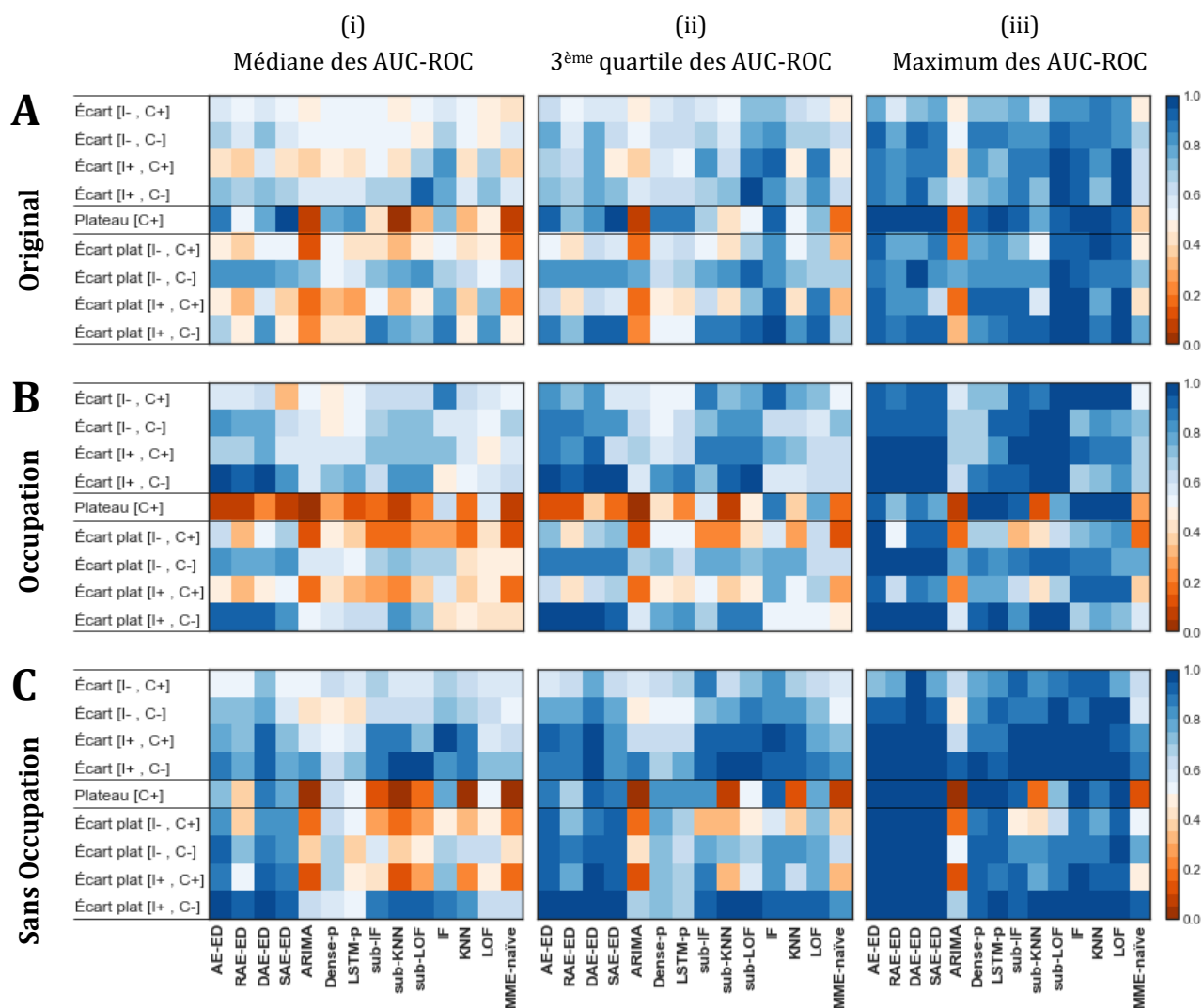


Figure 63 Carte thermique des médianes (i), des troisièmes quartiles (ii) et des maximums (iii) des indices AUC-ROC de l'ensemble des modèles, classés par type d'anomalies et appliqués aux mesures de concentration de CO₂. (A) les neuf ensembles anormaux originaux, (B) les neuf ensembles anormaux avec uniquement la période d'occupation et (C) les neuf ensembles anormaux avec uniquement la période d'absence des occupants. La DTW a été utilisée comme mesure de similarité pour les modèles de reconstruction et de prédiction.

Les résultats montrent que la séparation des deux périodes, chacune présentant une tendance différente, a eu un impact globalement positif sur les performances de détection

des modèles. Les améliorations les plus significatives des performances sont observées pour les écarts, avec une amélioration moyenne de 20 à 40 % en matière de médiane et de troisième quartile des indices AUC-ROC. Il est également remarqué que les meilleures performances obtenues pour les modèles de reconstruction sur les deux nouveaux jeux de données (B.C.iii) sont similaires pour l'ensemble des profils d'anomalies et sont proches du score maximal, contrairement aux jeux de données originaux (A.iii).

Cependant, il est important de noter que les anomalies diffèrent entre chaque ensemble, notamment pour les plateaux et les écarts plats. Néanmoins, on remarque des tendances relatives à l'intensité et à la contamination similaires à celles observées pour la mesure de température pour les nouveaux ensembles (B) et (C). Les anomalies de faible intensité semblent plus difficiles à détecter que celles de forte intensité, et l'inverse semble également être vérifié pour le taux de contamination. Dans ce cas d'application, la variation de tendance semble alors avoir une influence notable indirecte sur la détection d'anomalies. Cette influence reste indirecte car elle dépendra du positionnement de l'anomalie et non de la tendance en elle-même, qui ne semble pas être considérée comme anormale par les modèles. Cette observation peut s'expliquer par l'ajout d'une quantité importante de mesures de concentration de CO₂ avec des profils variés lors de l'entraînement des modèles.

De plus, il est notable que les tendances observées dans la Figure 63 (C) sont très proches de celles observées pour les mêmes types d'anomalies pour les mesures d'humidité (Figure 60) ou de température (Figure 58), bien que les données de concentration de CO₂ soient plus bruitées. En fin de compte, il semble que la plupart des modèles ne soient pas en mesure de détecter convenablement ces anomalies, quelle que soit la tendance observée.

Il est intéressant de noter que, malgré la grande variabilité naturelle des observations de la mesure de CO₂, les modèles de prédiction et de reconstruction ne sont pas en mesure de détecter de manière fiable les anomalies de plateau et d'écart plat. Cette conclusion peut sembler surprenante, mais elle s'explique par le fait que ces anomalies ne présentent pas un score d'anormalité particulièrement élevé, même dans un contexte de données variables. En d'autres termes, même si ces cas d'anomalies sont absents, dans l'ensemble des données d'apprentissage et que les données de CO₂ présentent une variabilité importante, seuls les modèles qui utilisent une fenêtre permettant d'englober l'intégralité de l'anomalie et de la séparer clairement du reste des données semblent obtenir de bons résultats. Ainsi, il ne s'agit pas tant de la dynamique des données en elle-même qui influence cet aspect, mais plutôt de la manière dont les modèles sont capables de traiter les anomalies en fonction de la fenêtre d'observation et de reconstruction qu'ils utilisent. Une explication plus détaillée de ce phénomène est prévue dans la partie 5.5.6.2.

5.5.5. Le cas des mesures de consommation électrique

5.5.5.1. Contexte

Tous les modèles, à l'exception de l'ARIMA, ont été appliqués aux mesures de consommation électrique cumulées pour détecter différents types d'anomalies tels que les pics, les écarts, les écarts plats et les changements de niveau. Un total de 14 jeux de tests distincts a été créé. La Figure 64 illustre la dynamique caractéristique des données de consommation électrique collectées, marquées par de vastes zones relativement plates, mais échelonnées, et des changements de niveaux plus marqués, dont certains sont annotés E1, E2 et E3. Ces changements de niveaux ne sont pas seulement brusques mais ils peuvent être composés d'une série d'augmentations très rapprochées les unes des autres, survenant lors de la consommation d'énergie. Cette particularité s'explique par deux facteurs. Tout d'abord, il s'agit d'un compteur qui s'incrémente à différentes fréquences, généralement toutes les 1 à 2 minutes en l'absence de changement de valeurs, mais pouvant être aussi courtes que quelques secondes en cas de consommation électrique. Ensuite, la méthode de collecte des données est particulière, car chaque prise est instrumentée individuellement, ce qui signifie que les consommations sont récoltées de manière individuelle. Plusieurs approches ont été envisagées pour ces travaux, telles que l'association de l'ensemble des consommations en une seule mesure ou la conversion de la consommation cumulée en consommation instantanée. Cependant, en raison du nombre d'erreurs et de données manquantes rencontrées, il n'a pas été possible d'exploiter plus de quelques semaines consécutives de données de la sorte, et ces approches n'ont donc pas été appliquées.

Compte tenu du profil spécifique de ces mesures, tel qu'illustré dans la Figure 64, les anomalies de plateau ne se démarquent pas clairement des autres observations et n'ont donc pas été incluses. Il est important de noter que, pour des raisons similaires, les écarts plats apparaissaient très similaires aux anomalies écarts. Cependant, les deux types d'anomalies ont été inclus dans le but de réduire, à moindre frais, la variabilité due à l'effet stochastique de l'introduction d'anomalies. Concernant les anomalies de bruits, elles n'ont pas été observées dans nos données ni identifiées dans la littérature pour les capteurs de consommation électrique cumulée, et par conséquent, elles n'ont pas été considérées.

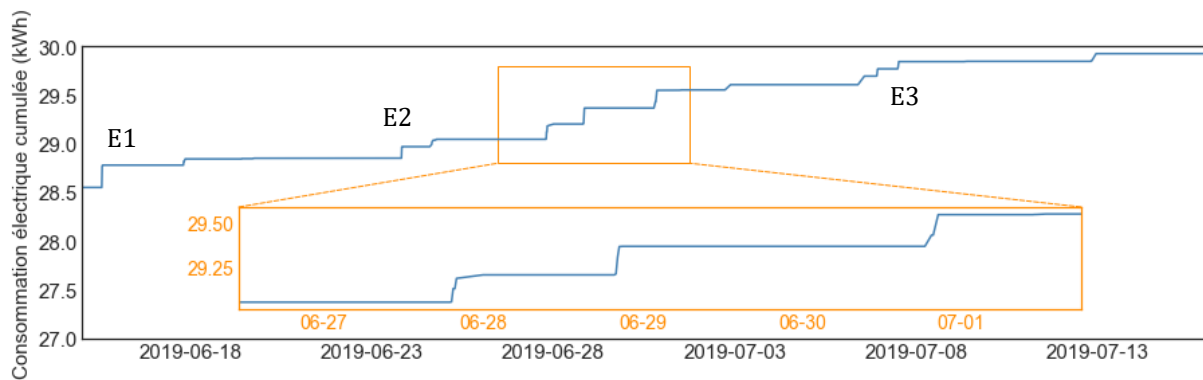


Figure 64 Mesure de consommation électrique cumulée saine avant contamination.

5.5.5.2. Performances des modèles pour chaque anomalie

Les performances des modèles pour chaque type d'anomalie sont présentées de façon similaire à la Figure 57, pour la température d'air intérieure, et sont consultables en Annexe C. Pour simplifier la discussion des résultats liés à la consommation électrique cumulée, nous utiliserons seulement trois cartes thermiques de la Figure 65, qui illustrent les médianes (i), les troisièmes quartiles (ii) et les maximums (iii) des indices AUC-ROC. La nomenclature suivie est identique à celle présentée dans la partie précédente. Cependant, pour cette figure, l'indice AUC-ROC des modèles de reconstruction et de prédiction a été calculé en utilisant les résultats provenant du score d'anormalité calculé par la DMA, compte tenu de ses meilleures performances.

Contrairement aux données précédentes, les mesures de consommation électrique sont très lisses et présentent peu de bruit ou de variabilité. Par conséquent, l'évaluation du score d'anormalité observation par observation, avec la DMA apparaît globalement plus stable que l'approche par séquençage employée avec les autres mesures. Les performances moins élevées des mesures de similarité MSE, MAE et DTW semblent s'expliquer en premier lieu par l'approche de détection par fenêtre glissante utilisée. Cette approche offre une forme de robustesse aux mauvaises reconstructions ou prédictions, car plusieurs séquences anormales consécutives sont nécessaires pour détecter une anomalie. Elle convient donc particulièrement aux données présentant des dynamiques variées ou du bruit naturel. *A contrario*, il suffit qu'une seule séquence soit considérée comme normale dans un ensemble de séquences anormales pour sensiblement détériorer les performances de détection (cf. partie 5.3.4.2).

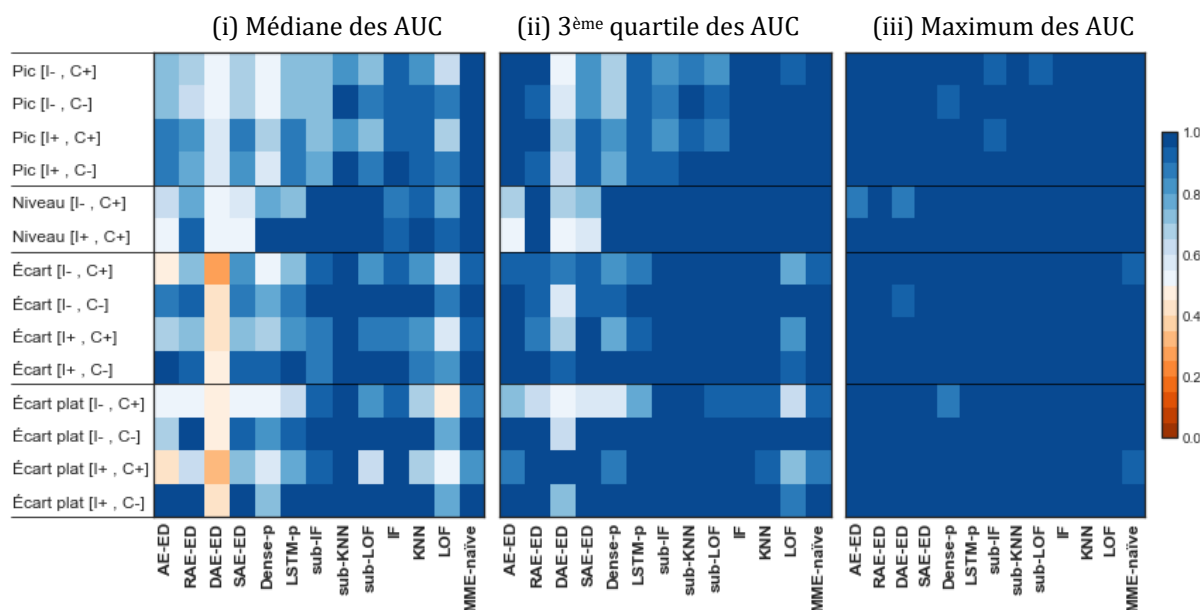


Figure 65 Carte thermique des médianes (i), des troisièmes quartiles (ii) et des maximums (iii) des indices AUC-ROC de l'ensemble des modèles, classés par type d'anomalies et appliqués aux mesures de consommation électrique. La DMA a été utilisée comme mesure de similarité pour les modèles de reconstruction et de prédiction.

Premièrement, contrairement aux résultats précédemment présentés pour les mesures de température, d'humidité et de concentration de CO₂, il y a une distinction claire entre les modèles non supervisés et semi-supervisés. Les techniques non supervisées obtiennent la quasi-totalité des meilleurs indices AUC-ROC médians pour chaque type d'anomalies, de contamination ou d'intensité. Les modèles séquentiels de voisinage, sub-IF, sub-KNN et sub-KNN, se distinguent particulièrement par leur stabilité et leurs performances élevées pour l'ensemble des types d'anomalies, à l'exception des pics. Cela est une caractéristique commune à tous les ensembles des données anormales précédemment présentés qui présentent ce type d'anomalies. Les valeurs d'indices AUC-ROC médians (i) sont presque équivalentes aux valeurs maximales atteintes (iii) pour l'ensemble des jeux d'évaluation, et donc des types d'anomalies. Ce constat est d'autant plus remarquable au troisième quartile (ii).

Le modèle classique de voisinage d'IF semble également particulièrement efficace, tout comme les modèles KNN et LOF, bien que l'on puisse observer une plus grande variabilité dans les anomalies collectives, surtout pour le LOF (ii). Cependant, il convient de noter que ces bonnes performances doivent être interprétées avec précaution, compte tenu des excellents résultats également obtenus par la MME-naïve. Ces résultats semblent découler de la dynamique spécifique de ces données. Dans le cas des techniques classiques de voisinage (IF, KNN et LOF), à l'exception des anomalies, les observations normales avec le score d'anormalité le plus élevé se situent juste avant, pendant et peu après les changements de niveaux naturels, qui peuvent souvent être considérés anormaux. Cela

explique la plus grande variabilité des résultats, notamment pour les anomalies d'écart et d'écart plat, qui apparaissent alors similaires aux changements de niveaux. Pour les modèles de voisinage employant des séquences, sub-IF, sub-KNN et sub-LOF, cet inconvénient se produit moins fréquemment en raison de l'approche par fenêtre glissante et de la méthode de calcul du score d'anormalité. Les anomalies semblent donc plus faciles à détecter. Enfin, pour la MME-naïve, des explications similaires à celles des techniques classiques de voisinages peuvent être avancées. Cependant, contrairement aux cas précédents, la stabilité générale des données semble mieux permettre à la MME-naïve « d'absorber » les variations induites par les anomalies. Il est toutefois notable que plus l'anomalie est d'une intensité forte (I+) et de longue durée (C+), plus cela a un impact sur les résultats. Dans de tels cas, les MME avec les coefficients de lissage les plus élevés semblent les plus performantes.

Les méthodes semi-supervisées présentent, pour l'ensemble des modèles, à l'exception des LSTM-p, une variabilité notable des indices AUC-ROC médians, avec des valeurs oscillant entre 0,5 et 0,9. (i). Ces médianes affichent des performances nettement inférieures à celles des méthodes non supervisées, et montrent une variabilité plus marquée que les autres modèles. Toutefois, il est important de noter que ces performances sont comparables à celles obtenues pour les autres mesures, et à partir du troisième quartile (ii), elles sont proches des modèles non supervisés, voire deviennent identiques pour les valeurs maximales (iii).

Le pire modèle est le DAE et sa version ED, ce qui n'est pas surprenant compte tenu du fait qu'ils sont surtout adaptés aux données présentant du bruit. D'une manière générale, les modèles semi-supervisés semblent moins adaptés à des données de cette nature, présentant peu de variations ou de bruits, notamment en raison de la quantité de données ou du temps nécessaire pour les entraîner. Les méthodes non supervisées de voisinage comme l'IF, la MME, pouvant être appliquées sans *a priori* sur les données et plus simples à utiliser, semblent constituer un meilleur compromis.

5.5.6. Observations communes à l'ensemble des variables

Dans cette partie, nous utiliserons les résultats des mesures de température pour mettre en évidence deux aspects importants (1) l'influence des combinaisons, des associations et des transformations ainsi que (2) l'impact de la taille des fenêtres mobiles sur les performances de détection d'anomalies. Il convient de noter que les tendances présentées ici sont globalement similaires à celles observées pour l'humidité relative, la concentration de CO₂ et les consommations électriques cumulées. Les résultats détaillés pour chaque variable sont accessibles en Annexe D et en Annexe E, dans un format similaire à celui utilisé ci-après.

5.5.6.1. Influence des combinaisons, des associations et des transformations sur les performances de détection

L'influence des différentes combinaisons, associations, et transformations, sur la tâche de détection peut être observée en examinant les variations des valeurs médianes, du troisième quartile et du maximum de l'indice AUC-ROC dans la Figure 66. Cette figure compare les résultats des modèles employant uniquement la variable de température sans combinaison, ni associations, ni transformations, aux résultats des modèles lorsque, (A) les combinaisons de variables sont utilisées ou (B) les associations et les transformations sont appliquées.

L'objectif de cette figure est d'évaluer si l'ajout de variables ou de transformations améliore (ou rend plus difficile) la tâche de détection d'anomalies. Cela peut se traduire par un meilleur apprentissage de la normalité pour les modèles de reconstruction et de prédiction, ou par une meilleure distinction entre les observations normales et anormales pour les modèles basés sur la méthode de voisinage. Seuls les modèles utilisant plusieurs variables ou transformations sont inclus dans cette analyse. Une valeur positive (ou négative) signifie que l'apport des combinaisons ou des transformations implique une amélioration (ou une détérioration) correspondante de la médiane (i), du troisième quartile (ii) ou du maximum (iii) des indices AUC-ROC. Cependant, il convient de rester prudent dans l'interprétation de ces résultats, car d'autres facteurs tels que les entraînements et la position aléatoire des anomalies doivent également être pris en considération.

La Figure 66 permet, en partie, d'illustrer deux profils de modèles : d'une part, ceux qui sont faiblement impactés par l'apport de combinaisons, d'associations et de transformations de variables, et d'autre part, ceux qui sont sensibles à ces modifications. Les modèles basés sur les méthodes de prédiction ou de reconstruction semblent moins dépendant des combinaisons (A), associations ou transformations de variables (B). Leur apport, semble avoir un impact soit nulle, soit faible (entre $\pm 0,05$ et $\pm 0,1$) sur les valeurs médiane (i) des indices AUC-ROC correspondants. Ce constat est d'autant plus visible dès le troisième quartile mais surtout pour le maximum (iii) puisque, qu'à quelques exceptions près, aucun impact n'est induit par l'apport des combinaisons et transformations.

Les modèles de reconstructions semblent être les plus stables, avec peu de variations significatives entre la médiane, le troisième quartile et le maximum. Ces résultats suggèrent que les Auto-Encodeur sont capables, *a minima*, d'offrir des performances équivalentes aux Encodeur-Décodeur (ii et iii) et, dans certains cas, meilleures (A.i) pour une grande partie des types d'anomalies étudiée. Dans le cas de l'apport de transformations, la différence entre les variations constatées pour les valeurs médianes

(B.i) et les valeurs maximales (B.iii) signifie que l'emploi de transformations améliore légèrement la stabilité générale des modèles mais ne permet pas d'obtenir les meilleures performances. Cela s'explique par le fait que chaque modèle de reconstruction est entraîné quatre fois successive, une pour chaque taille de fenêtre, et que l'une de ces tailles permet aux AE d'atteindre les meilleurs résultats. Les autres, moins performantes, compensent légèrement en tirant parti des transformations.

Les modèles de prédictions semblent également relativement stables et peu dépendant des combinaisons ou transformations. Nous observons cependant une légère dépendance aux combinaisons pour les anomalies collectives d'écart, d'écart plat et de plateau, bien qu'uniquement visible sur les performances maximales (A.iii).

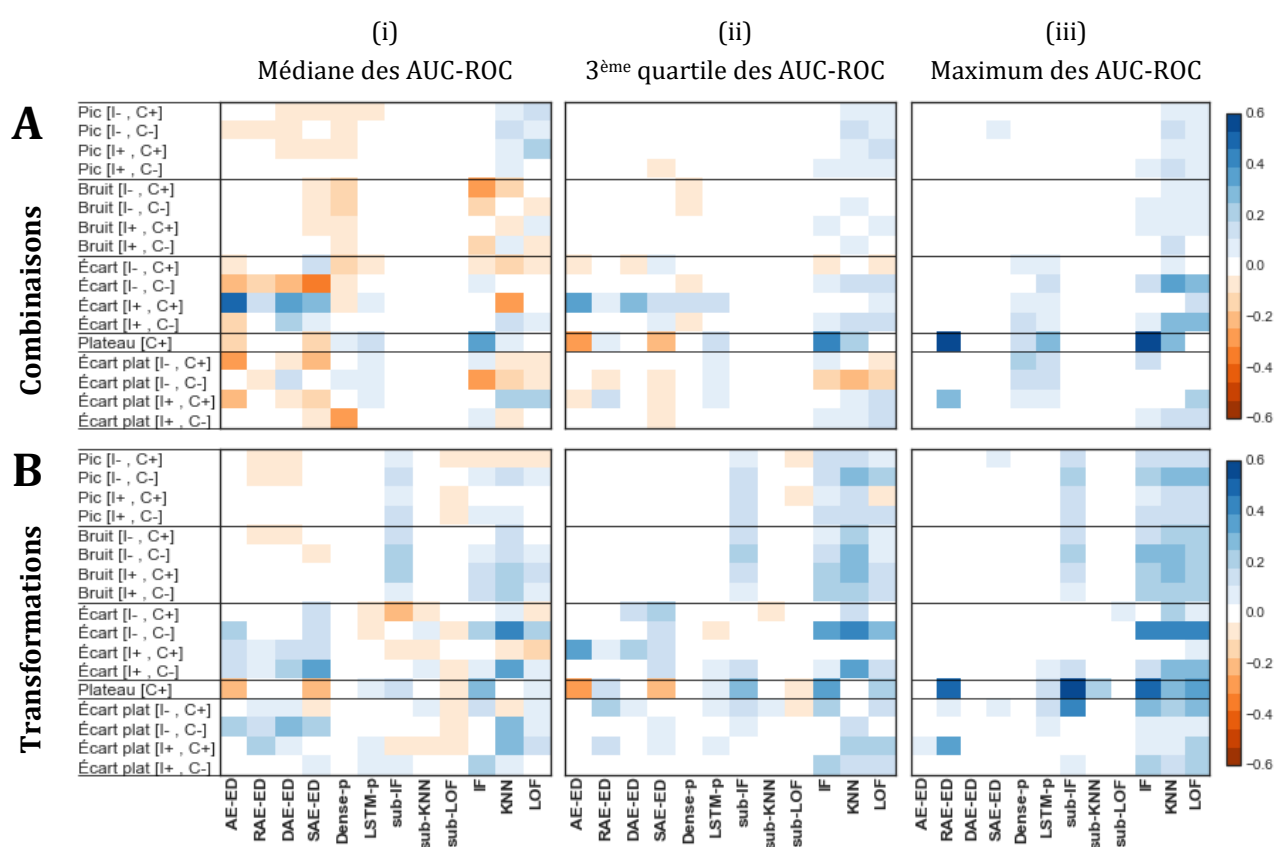


Figure 66 Carte thermique des variations des médianes (i), des troisièmes quartiles (ii) et des maximums (iii) des indices AUC-ROC de l'ensemble des modèles, classés par type d'anomalies et appliqués aux mesures de température d'air intérieure. Variations des indices AUC-ROC induites par l'apport de combinaisons (A) ou de transformations (B), par rapport aux performances obtenues en utilisant seulement la température sans combinaison, associations ou transformations. Une valeur supérieure à 0 (resp. inférieur) présente une amélioration (resp. réduction).

Au contraire des observations précédentes, les modèles de voisinage classiques, tels que IF, KNN et LOF, mais également la variante séquentielle sub-IF, semblent être particulièrement impactés par l'utilisation de combinaisons ou de transformations. L'utilisation exclusive de la variable de température ne semble pas suffisante pour atteindre des performances optimales. En effet, pour les transformations, de légères

augmentations de l'indice AUC-ROC sont observées dès la médiane (B.i), et ces augmentations deviennent plus conséquentes et généralisées, allant de 0,2 à 0,3, pour le troisième quartile (B.ii), pour toutes les méthodes classiques de voisinage (IF, KNN et LOF). Cette tendance est encore plus marquée pour les valeurs maximales (B.iii). Concernant l'apport de combinaisons de légères améliorations peuvent également être constatées bien qu'elles soient plus proches des 0,1 et moins systématiques (A.iii). De plus, elles tendent à détériorer de façon plus prononcée les résultats médians (A.i).

Ainsi, les meilleurs modèles obtenus en utilisant uniquement la température (sans transformations ni combinaisons) présenteront en moyenne des performances inférieures de 30 % par rapport aux versions avec combinaisons et transformations, et ce, pour l'ensemble des types d'anomalies. Les variantes séquentielles des méthodes de voisinages, sub-IF, sub-KNN et sub-LOF, semblent moins sensibles à de tels écarts, à l'exception de la sub-IF.

Parmi les transformations, trois se démarquent par leur impact significatif : la dérivée, le résidu_{STL} et l'association de la moyenne_{Fenêtre} et de l'écart type_{Fenêtre}. La dérivée améliore légèrement la détection des pics et des bruits. Cette amélioration peut s'expliquer par le fait que ces anomalies se caractérisent par une dispersion locale ou collective plus importante que les autres observations, ce qui peut être traduit par la dérivée. Cependant, l'application de la dérivée détériore sensiblement les résultats de détection des écarts, écarts plats et plateaux, avec une diminution d'environ 20 à 30 %. Pour les écarts, une explication possible pourrait être que la séquence d'observation anormale présente la même variabilité que les autres observations, bien qu'elle soit décalée d'un biais. Pour les deux autres types, les résultats sont plus surprenants. L'hypothèse, à ce stade, serait que ces deux anomalies sont relativement proches de la dynamique habituelle des mesures de température intérieure, qui présente également de très faibles variations sur de longues durées. Ainsi, une dérivée nulle sur une période de temps prolongée ne serait pas considérée comme une anormale.

Dans le cas des écarts, écarts plats et plateaux, l'utilisation de la transformation résidu_{STL} mais aussi de l'association de la moyenne_{Fenêtre} et de l'écart type_{Fenêtre}, semblent plus pertinentes. Elles permettent de rendre compte de l'écart entre la tendance attendue des données et celle observée. Pour les mesures de températures d'air intérieur, ces transformations semblent être polyvalentes et conduisent à une amélioration moyenne de la détection de 10 à 15 %. Cependant, l'association de la moyenne_{Fenêtre} et de l'écart type_{Fenêtre} détériore sensiblement les performances de détection des pics, en raison de l'effet d'agrégation qui absorbe les anomalies de très courtes durées. En revanche, cet effet permet de mieux caractériser les changements de moyenne ou de variance sur la durée, ce qui est globalement bénéfique pour la détection des anomalies collectives.

L'ensemble des tendances observées dans cette partie se retrouvent également pour les autres variables. Elles sont très proches mais moins prononcées pour les mesures d'humidité relatives mais bien plus prononcées et influentes, pour les mesures de concentrations de CO₂. Celles associées aux consommations électrique cumulées sont très faibles, voire inexistantes. Pour la concentration de CO₂, cela peut s'expliquer par la variabilité naturelle des données qui semble compliquer la tâche de détection, notamment des méthodes de voisinage. Dans le cas des consommations électriques, cela s'explique également par la nature spécifique des données. L'absence notable de variabilité semble rendre la tâche de détection d'anomalies moins complexe puisque les observations anormales ont des comportements qui dénotent relativement bien des observations normales.

5.5.6.2. Influences de la taille des fenêtres glissantes sur les performances de détection

L'influence de la taille des fenêtres glissantes sur les performances de détection est présentée dans la Figure 67. Les modèles sont regroupés en deux catégories :

1. Les modèles de reconstruction (AE, RAE, DAE et SAE et leurs versions d'ED), ainsi que les variantes séquentielles des méthodes de voisinages (sub-IF, sub-KNN et sub-LOF) qui utilisent des fenêtres de 12, 24, 48 et 96 observations.
2. Les modèles ARIMA, Dense-p et LSTM-p qui effectuent des prédictions basées sur des fenêtres de 4 et 12 observations.

Chaque jeu d'évaluation anormal comporte environ 450 résultats individuels pour chaque modèle, et les médianes sont représentées par des traits pleins de couleurs et de symboles distincts pour chaque taille de fenêtre. Les moyennes sont représentées par des lignes discontinues de la même couleur et du même symbole. La DTW a été employée pour mesurer le score d'anormalité.

Cette analyse permet de comprendre comment la taille des fenêtres glissantes affecte les performances de détection de chaque modèle.

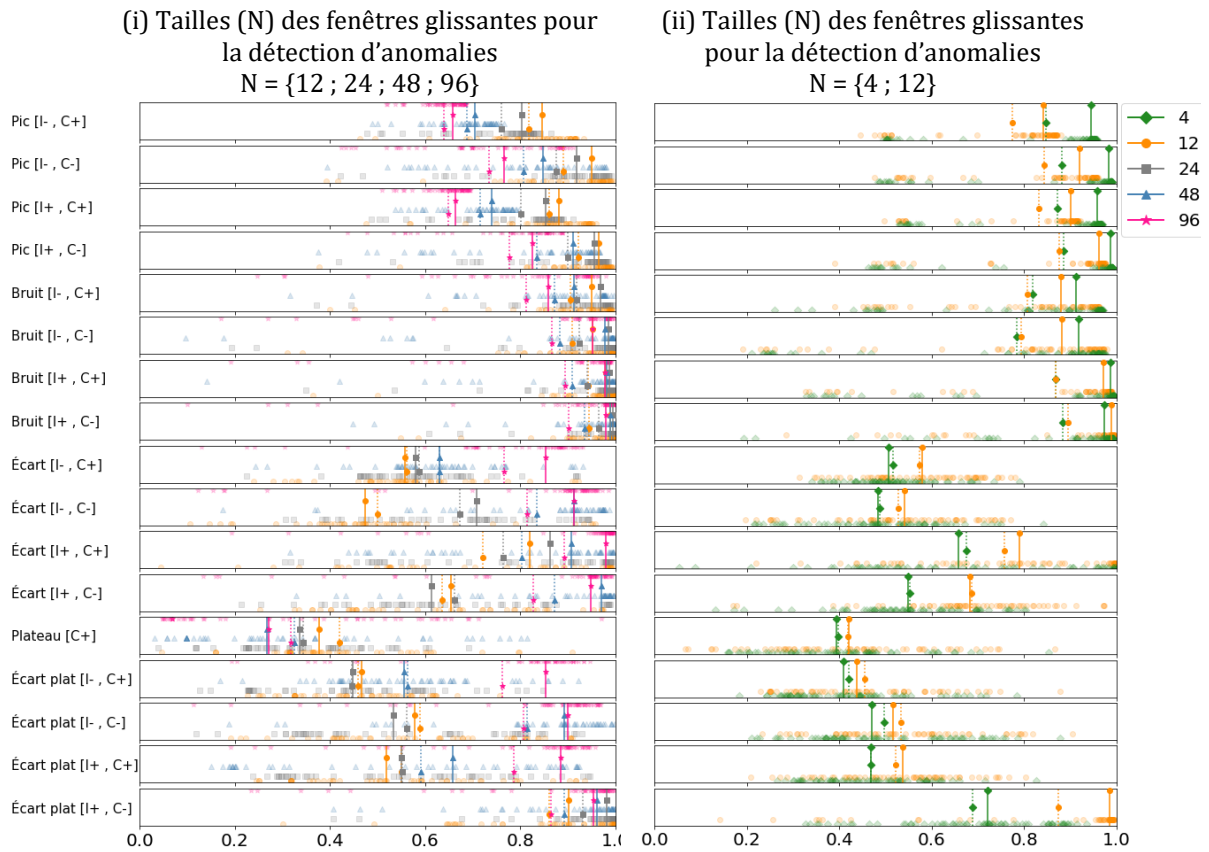


Figure 67 Influence de la taille des fenêtres glissantes de détection, prédiction et reconstruction, sur les performances de détection d'anomalies. Les modèles de reconstruction et de voisinage séquentiel sont présentés en (i) et les modèles de prédiction en (ii). Les médianes sont représentées par les traits pleins dont les couleurs et les sigles respectifs sont annotés en légende. Les moyennes sont représentées par des traits en pointillés de même couleur et même sigle.

La Figure 67 met en évidence plusieurs points importants qui confirment les observations précédentes. Tout d'abord, les valeurs médianes de l'AUC-ROC sont inversement proportionnelles à la taille de la fenêtre utilisée pour les pics (i et ii). De plus, il existe une certaine constance entre les modèles lorsqu'ils utilisent des fenêtres de même taille, comme c'est le cas avec une fenêtre de 12 observations, qui présente des valeurs médianes similaires dans les deux cas (i et ii). De plus, il est clair que le taux de contamination a une influence significative sur les performances, indépendamment de la taille des fenêtres.

Pour les bruits, la faible variabilité des résultats dans la Figure 67 pourrait s'expliquer par le manque d'influence que présente la taille des fenêtres glissantes. Cependant, il est important de noter que les modèles en (i) ont de meilleures performances que ceux en (ii) pour ce type d'anomalies. Cela suggère que les modèles en (ii) sont sensibles à l'intensité du bruit, mais pas à sa taille (exprimé par le taux de contamination). Il est possible que les modèles de prédiction soient impactés par la présence des observations bruitées lors de leur passage.

Pour les autres anomalies collectives, une tendance inverse se manifeste, et la taille de la fenêtre apparaît alors comme particulièrement discriminante. On observe une disparité très importante des résultats pour l'ensemble des anomalies d'écart et d'écart plat. Lorsque de plus petites tailles de fenêtres sont utilisées (4 et 12) (i et ii), les indices AUC-ROC médians et moyens sont d'environ 0,5. En revanche, lorsque les tailles de fenêtres plus grandes sont utilisées (48 et 96) (i), les performances augmentent considérablement, atteignant environ 0,9 (i).

L'impact du taux de contamination sur les performances des modèles, et par conséquent, de la taille des anomalies, apparaît également plus marqué que celui de l'intensité, notamment pour les plus petites tailles de fenêtres (4, 12 et 24). Lorsque la fenêtre employée est trop petite par rapport à la taille de l'anomalie, il arrive un moment où l'ensemble de la fenêtre est constituée d'observations anormales. Cependant, si ces observations anormales ne divergent que légèrement des tendances observées (p. ex un écart de faible intensité), les scores d'anormalité associés seront trop similaires à ceux des observations normales pour être considérés comme anormaux.

Cela explique les résultats faibles obtenus pour les modèles appliquant de petites fenêtres de détection, de reconstruction ou de prédiction, car seules les limites de l'anomalie seront correctement détectées. En illustrant cette idée, le jeu « Écart [I-, C-] » a été contaminé par deux anomalies de faible intensité (respectivement 33 et 17 observations), avec un taux de contamination faible de 1,25 %. Les modèles utilisant une fenêtre de taille 12 ne peuvent détecter l'intégralité de ces anomalies, d'où leur faible score AUC-ROC de 0,5. En revanche, les modèles utilisant une fenêtre de taille 24 détectent la deuxième anomalie mais pas la première. Les modèles dont la fenêtre dépasse 48 observations détectent les deux anomalies et présentent donc des performances similaires. En revanche, le jeu « Écart [I+, C+] » montre de meilleurs résultats malgré un taux de contamination plus élevé. Dans ce cas, les anomalies sont de plus forte intensité, rendant la taille de la fenêtre moins déterminante, bien qu'une différence notable persiste.

Pour ces travaux, les anomalies ne dépassent pas 96 observations (équivalent à 24 heures) ce qui est également la taille maximale des fenêtres employées. En conséquence, les fenêtres sont au moins aussi grandes que les anomalies. Cependant, il est important de noter que les médianes des résultats des AUC-ROC varient peu pour l'ensemble des anomalies collectives (à l'exception des plateaux) et des taux de contamination, malgré les variations d'intensité. Ainsi, dans le cas de la détection d'anomalies collectives, le choix d'une taille de fenêtre adaptée pourrait permettre de réduire la versatilité induite par l'intensité des anomalies. Bien que la taille de l'anomalie ne soit pas le seul facteur à considérer, elle demeure l'un des plus importants, voire le plus important, car elle

détermine la taille de fenêtre utilisée pour la détection et influence les performances potentielles.

À titre d'exemple des variations induites par la taille des fenêtres de détection, la Figure 68 illustre les résultats pour les anomalies écart, écart plat et plateau. Dans la partie (A), les résultats originaux sont présentés, tels qu'exposés dans la partie 5.5.2. Dans la partie (B), seuls les résultats correspondant aux tailles de fenêtres de 48 et 96 observations sont inclus pour les modèles d'AE-ED, RAE-ED, DAE-ED et SAE-ED ainsi que les variantes séquentielles sub-IF, sub-KNN et sub-LO. F. On remarque une amélioration notable des performances médianes entre (A.i encadré blanc) et (B.i encadré blanc) mais surtout au troisième quartile, qui atteint presque les valeurs maximales (ii encadré blanc). Par ailleurs, aucun changement n'est constaté pour les deux maximales, ce qui indique que l'ensemble des meilleurs modèles se compose soit de modèles de fenêtre de taille 48, soit de taille 96.

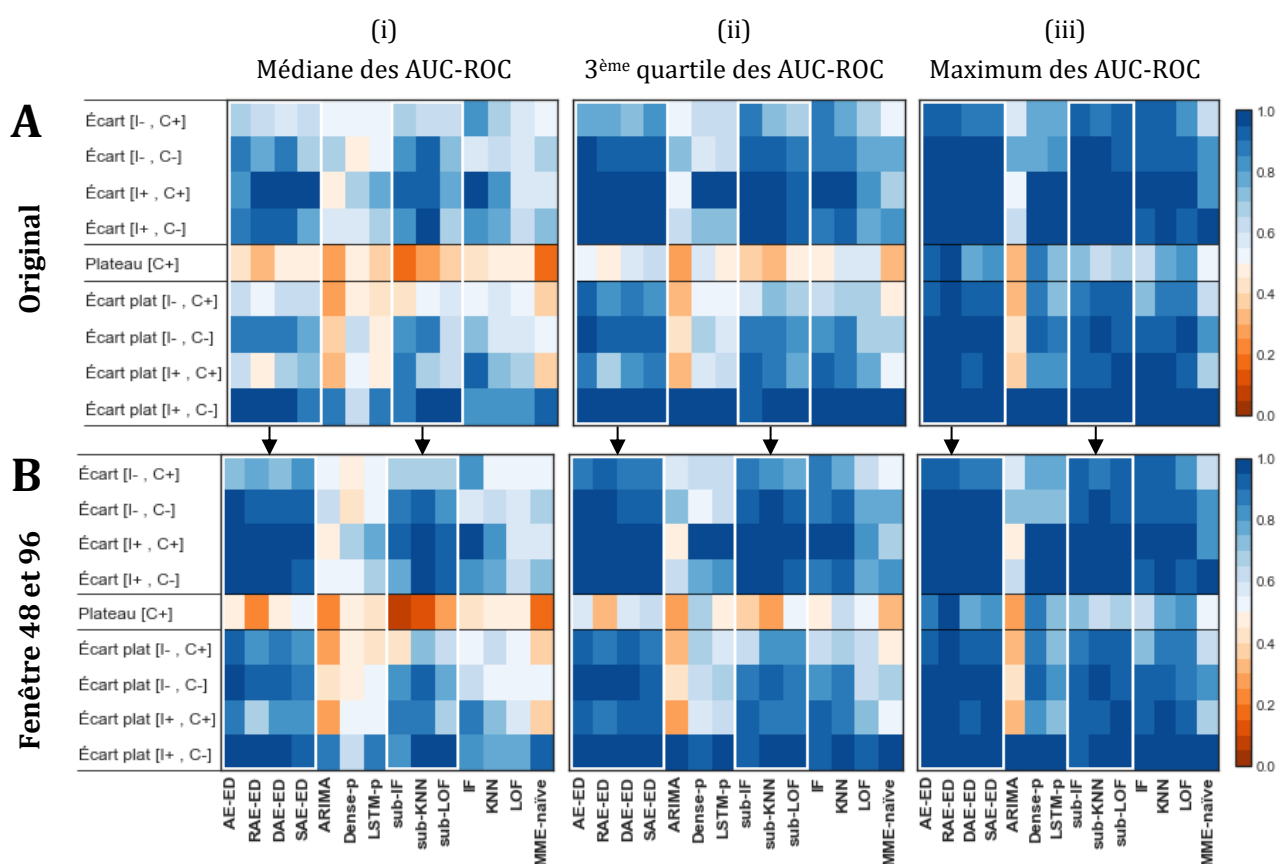


Figure 68 Carte thermique des médianes (i), des troisièmes quartiles (ii) et des maximums (iii) des indices AUC-ROC de l'ensemble des modèles, classés par type d'anomalies et appliqués aux mesures de consommation électrique. (A) les 9 jeux d'observations anormales originaux de la température et (B) les neuf jeux d'observations anormales composés uniquement des résultats des modèles de fenêtres d'une taille de 48 ou 96 observations.

5.5.7. Synthèse des résultats

Ces travaux ont permis de tirer plusieurs conclusions importantes. Concernant les variables, il a été observé que les performances de détection d'anomalies sont assez proches entre les modèles lorsque les variables partagent des tendances et des dynamiques générales similaires, comme c'est le cas avec les mesures de température d'air et d'humidité relative intérieures. De plus, l'influence conséquente des changements de tendance à long terme sur ces performances a été mise en évidence, ainsi que l'impact de la dynamique des données. Dans le premier cas, les changements de tendance ont provoqué des disparités de détection selon les modèles et le positionnement des anomalies. La séparation des jeux de données en deux jeux distincts avec des tendances différentes a finalement conduit à une amélioration des performances de détection pour plusieurs modèles. Dans le deuxième cas, des différences marquées de performances ont pu être observées entre les mesures de température, humidité et CO₂ d'une part, et les mesures de consommation électrique cumulée d'autre part. Ces dernières se différencient par des données présentant très peu de variations par rapport aux autres mesures, mais surtout par un profil difficile à prévoir. Sur de telles données, où les observations normales et anormales sont plus distinctes, les méthodes non supervisées de voisinage classiques et séquentielles comme l'IF, la sub-IF, la sub-KNN ou même la MME, qui peuvent être appliquées sans *a priori* sur les données et sont plus simples à utiliser, semblent particulièrement efficaces et fiables. Au contraire, sur des données présentant une dynamique plus complexe, des tendances courtes et variées, voire du bruit, il est préférable de privilégier les modèles de reconstruction et les modèles séquentiels de voisinage. Cependant, compte tenu de la grande diversité des données issues du domaine du bâtiment, des recherches plus approfondies sur d'autres variables sont nécessaires pour approfondir ou généraliser les observations faites dans ces travaux.

Concernant les modèles, bien que les résultats n'indiquent pas qu'un modèle soit globalement plus performant qu'un autre, plusieurs tendances semblent émerger. Tout d'abord, il est à noter que la plupart des modèles employés permettent de détecter presque tous les types d'anomalies, quelle que soit la variable considérée. En effet, les valeurs maximales des AUC-ROC atteintes sont constamment supérieures à 0,9 et se rapprochent de 1. Si l'objectif de ces travaux était uniquement de trouver le meilleur modèle, il serait difficile de faire des comparaisons significatives compte tenu de ces résultats.

Cependant, l'utilisation de diverses transformations, associations et combinaisons de données, ainsi que de différentes méthodes d'entraînement avec des fenêtres glissantes de tailles variables ou de calcul de scores d'anormalité, permet de faire des distinctions plus claires entre les modèles.

Premièrement, il est important de noter que les performances de détection des modèles de voisinage classiques tels que IF, KNN et LOF, ainsi que celles des variantes séquentielles comme sub-IF, semblent être relativement influencées par les associations et les transformations de données. La manière dont ces transformations fonctionnent et s'appliquent à la détection d'anomalie est étroitement liée aux caractéristiques des anomalies elles-mêmes, notamment leur taille, leur intensité, leur moyenne ou leur variance, ainsi qu'aux caractéristiques des données sur lesquelles elles sont appliquées.

Ce constat suggère que l'application de ces modèles de voisinage classiques nécessite une connaissance *a priori* des types d'anomalies recherchés afin d'éviter l'utilisation de transformations inadéquates. Parmi les transformations qui se sont révélées pertinentes pour ces modèles, il apparaît adéquat de citer l'utilisation de la dérivée pour les types d'anomalies de pics et de bruits mais aussi le résidu_{STL} ou la moyenne_{Fenêtre} et l'écart type_{Fenêtres}, qui se sont avérés particulièrement adaptés aux types d'anomalies d'écart et d'écart plat. Ces méthodes sont simples à mettre en oeuvre, rapides à exécuter et faciles à exploiter. Par conséquent, il est possible d'effectuer plusieurs essais, de corrélérer les résultats ou simplement de les utiliser en complément d'autres techniques pour cibler spécifiquement certaines anomalies.

Deuxièmement, il est important de noter que les performances de détection des modèles qui utilisent une approche basée sur des fenêtres mobiles semblent être les plus robustes et moins sensibles aux combinaisons, associations et transformations de données. Ces modèles sont particulièrement adaptés à la détection d'anomalies collectives. Dans ce contexte, la taille de la fenêtre employée se révèle être le critère le plus crucial pour optimiser les performances de détection. Il est généralement recommandé d'utiliser une taille de fenêtre légèrement supérieure à la taille des anomalies, ce qui semble être un compromis adéquat.

Cependant, il convient de noter que les modèles de prédiction basés sur des fenêtres mobiles peuvent avoir des applications limitées dans le cas d'application présenté dans ces travaux, en raison de la contrainte liée à la taille des anomalies. Ces modèles pourraient être plus adéquats dans des cas d'applications réels de prédiction ou à des situations impliquant des anomalies de petites tailles. Les modèles basés sur des réseaux de neurones denses (Dense-p) se distinguent par leur légèreté, avec un poids moyen de 0,2 Mo, voire moins de 0,1 Mo lorsqu'ils utilisent peu de combinaisons ou d'associations. Toutefois, ces modèles peuvent être moins robustes que les LSTM-p, lesquels peuvent nécessiter un entraînement plus long et avoir des limitations d'utilisation en raison de leur poids et de leur temps d'entraînement plus importants. En fin de compte, les modèles de reconstruction se démarquent en tant que meilleurs parmi les modèles semi-supervisés

et affichent généralement les meilleures performances globales et la plus grande stabilité, tous modèles confondus.

Leur emploi semble être privilégiée dans des situations impliquant des données complexes pour lesquelles peu d'informations complémentaires sont disponibles. Ces modèles semblent moins dépendants de l'ajout d'informations, et bénéficient peu de ces ajouts. En général, il y a peu d'écart de performances observées entre les différentes variantes d'Autoencoder et d'Encoder-Decoder. Les performances des DAE-ED semblent légèrement meilleures lorsque des variations naturelles peuvent être observées dans les mesures, mais elles sont sensiblement détériorées dans le cas contraire. La principale différence tient donc dans le temps d'entraînement et la taille des modèles. Les RAE-ED, qui incluent des couches LSTM, sont bien plus lourds et longs à entraîner que les autres variantes (AE-ED, SAE-ED et DAE-ED). Ainsi, compte tenu des performances globales, ces variantes semblent moins pertinentes, bien qu'elles permettent d'obtenir de très bons résultats en matière de détection.

Concernant le modèle ARIMA, il aurait été possible d'explorer une approche basée sur des taux d'actualisation plus variés afin d'évaluer leur impact sur les performances de détection. Le taux d'actualisation initial, choisi pour adapter à des variables présentant des dynamiques différentes, s'est révélé trop court et a été préjudiciable pour la détection d'anomalies collectives, notamment celles de grandes tailles. Cependant, en raison du temps d'exécution important nécessaire par ce modèle et des résultats obtenus par les autres modèles de prédiction, il a été décidé de ne pas répéter l'expérimentation.

En ce qui concerne l'évaluation des scores d'anormalités de séquences utilisés pour les modèles de reconstruction et de prédiction, l'emploi de la DTW semble permettre d'obtenir des résultats légèrement plus performants que la MSE et la MAE. Cependant, il convient de noter que son usage est principalement adapté aux données présentant des dynamiques complexes, comme la température, l'humidité relative ou la concentration de CO₂.

Il est important de souligner que l'utilisation de la DTW demande des temps de calcul 5 à 50 fois plus long. Ces temps de calcul augmentent considérablement à mesure que la taille des fenêtres de séquence augmente. Ainsi, bien que la DTW puisse offrir des performances légères meilleures dans certaines conditions, son utilisation est réservée à des cas spécifiques où la précision de la mesure de similarité est cruciale malgré les coûts de calcul plus élevés. Enfin, l'approche basée sur le score d'anormalité d'observation employant la DMA apparaît comme un complément intéressant à exploiter en plus de l'approche basée sur les séquences, notamment pour des cas où une évaluation plus fine de l'anormalité est nécessaire.

Enfin, concernant les anomalies, il existe un constat mitigé. Les pics et les bruits semblent être les plus simples à détecter, quelles que soient leur taille ou leur intensité. Pour les écarts et écarts plats l'intensité et la taille semblent être deux composantes déterminantes à considérer conjointement selon l'approche envisagée. Concernant la dynamique des données, les performances moins bonnes observées pour les anomalies de type écarts plats de faible intensité, ainsi que celles de type plateau, laissent supposer que soit la dynamique des anomalies n'a pas l'impact aussi important que supposé dans le chapitre 3, soit que les données utilisées, les méthodes d'entraînement réalisées, les modèles employés, ou encore les transformations appliquées n'étaient pas adéquats ou suffisants.

Il pourrait être intéressant à l'avenir d'évaluer la représentation des espaces latents des autoencoders face à ce genre d'anomalies pour voir s'il est possible de les détecter autrement. Néanmoins il est important de considérer que, selon les données étudiées, ces anomalies pourraient éventuellement être détectées à l'aide de règles simples du type IF-THEN, associées à une expertise métier. En rétrospective, il aurait peut-être été judicieux de limiter l'aspect aléatoire de la localisation et de la taille des anomalies collectives. Un cadre plus contrôlé aurait pu permettre de faire des distinctions plus claires quant à l'influence d'autres caractéristiques, inhérentes aux données et aux anomalies, sur les performances de détection des modèles, comme cela a pu être observé pour la concentration de CO₂.

5.6. Exemple d'application sur des anomalies issues de mesures

5.6.1. Contexte

Dans le cadre de cet exemple d'application sur des anomalies provenant de mesures, nous utilisons les données de température d'air intérieur provenant du séjour de la maison, illustrée dans la Figure 69. Ces données ont un profil similaire à celui des mesures de température précédemment utilisées, mais elles comportent un nombre variable et conséquent d'anomalies. Nous avons annotés quelques exemples représentatifs de ces anomalies.

La première anomalie, qui est également la plus importante, est une anomalie prenant la forme d'un écart plat (EP) causée par l'arrêt du capteur pendant une semaine. De plus, nous avons identifié de nombreuses d'anomalies de type écart (p. ex., E1 et E2), liés à la position du capteur. En effet, le capteur est situé sur un mur à 1,5 mètres au-dessus du sol et est exposé directement aux rayons du soleil en début d'après-midi, ce qui a un impact conséquent sur la température ambiante à proximité du capteur. Ces anomalies se caractérisent par des variations positives d'environ 4 °C et d'une courte durée,

généralement 1 à 3 heures. Ces écarts conséquents sont liés à des phénomènes d'intérêts, similaires à ceux causés par une ouverture de fenêtre (p. ex., O1 et O2), tels que présentés dans la partie 2.2.

De manière similaire, nous observons la présence d'anomalies qui pourraient se rapprocher de bruits (p. ex., B1 et B2), potentiellement provoquées par l'effet combiné de l'ouverture de la baie vitrée et du rayonnement solaire. Enfin, quelques anomalies similaires à des pics (p. ex., P1, P2 et P3) sont également présentes de manière isolée.

Il est important de noter que ces données comportent un nombre considérable de valeurs manquantes, avec des périodes de 10 à 30 minutes sans données, plusieurs fois par jour. La plupart d'entre elles se situent une semaine avant l'arrêt de transmission (EP) et pendant la totalité de la période suivante. Pour permettre l'analyse de ces données, ces valeurs manquantes ont été corrigées par extrapolation linéaire. Compte tenu de l'agrégation des données réalisées, cela revient à remplacer une à deux observations seulement. Enfin, il convient de souligner que toutes les anomalies identifiées dans ces mesures se distinguent de celles étudiées précédemment par leur intensité beaucoup plus élevée, notamment l'anomalie de type écart plat, qui peut être considérée comme une anomalie de nature indépendante.

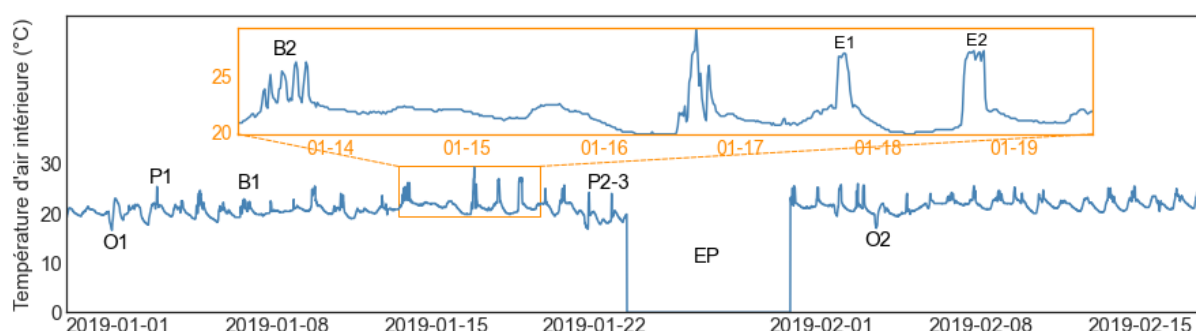


Figure 69 Mesure de température d'air intérieure comportant des anomalies (séjour de la maison).

5.6.2. Visualisation des résultats

Dans le cadre de cette application réelle, plusieurs modèles sont utilisés pour détecter les anomalies dans la mesure de température d'air intérieure (en suivant la même approche qu'en 5.3.4.2). Comme nous n'avons pas de données étiquetées (normales ou anormales), nous utilisons un *score de suspicion* pour évaluer, avec une valeur comprise entre 0 et 1, le degré d'anormalité de chaque observation. Ce score est calculé en comptant le nombre de fois où une observation est considérée comme anormale à chaque changement de seuil de décision. Ensuite, la somme est divisée par le nombre total de seuils employés. Ainsi, plus le score est proche de 1, plus l'observation est susceptible

d'être une anomalie. Pour les besoins de cette partie, nous considérons qu'une observation est suspecte si son score est supérieur à 0,5, tandis qu'une observation avec un score inférieur à 0,5 est considérée comme normale. Ce seuil de suspicion a été choisi de sorte à qualifier de suspecte toute observation qui est considérée comme anormale pour au moins 50 % des seuils de décision.

Pour cette application sur la mesure de température d'air intérieur, plusieurs modèles ont été utilisés. Pour simplifier l'application, nous n'avons pas effectué de combinaisons de variables, et le nombre de modèles présentés a été limité. Nous avons utilisé uniquement la variable originale ainsi que des associations et des transformations. Une taille de fenêtre de 12 heures (48 observations) a été utilisée pour les modèles de reconstruction et les modèles de voisinage séquentiel. Pour les modèles prédictifs, 12 heures d'observations passées sont utilisées pour les Dense-p et LSTM-p, afin de prédire des fenêtres de respectivement 1 et 3 heures (resp. 4 et 12 observations).

Les modèles, les associations et les transformations ont été choisis sur la base des tendances observées et des conclusions tirées de la partie 5.5. Les performances de tous les modèles, en fonction des tailles de fenêtres et des associations et transformations employées, sont présentées dans la Figure 70. Dans cette figure, une observation avec un score de suspicion proche de 1 est représentée en rouge, tandis qu'un score proche de 0,5 est représenté en jaune. Les observations non suspectes (score < 0,5) ne sont pas affichées. La taille des barres représente la durée des observations anormales. Le modèle de MME-naïve est utilisé comme référence de pour une méthode simple à appliquer.

Il est possible de tirer plusieurs observations de l'application des modèles sur la température de l'air intérieur.

Tout d'abord, l'utilisation de l'AE semble produire les meilleurs résultats parmi tous les modèles. Les résultats obtenus sont précis et présentent peu de bruit. De plus, il est relativement simple de corrélérer chaque écart aux scores dépassant le seuil de 0,5. Les modèles prédictifs semblent également présenter des résultats proches de ceux de l'AE, bien que légèrement plus bruités, notamment pour le Dense-p avec la fenêtre de prédiction la plus courte. Cependant, leurs scores semblent être plus discriminants que ceux de l'AE, car on constate plus fréquemment des scores proches de 1.

Contrairement aux résultats observés précédemment, les modèles parviennent à détecter l'écart plat, même avec une taille de fenêtre de reconstruction ou de prédiction beaucoup plus courte. Cette détection s'explique par l'intensité particulièrement importante de l'anomalie, comme détaillé dans la partie 5.5.6.2. Dans le cadre de l'expérimentation, toutes les valeurs nulles de cette anomalie ont été remplacées par la

dernière valeur enregistrée avant l'arrêt du capteur (19,8 °C). Les modèles n'étaient alors plus en mesure de détecter cette anomalie.

En résumé, les modèles AE et les modèles prédictifs, malgré un certain niveau de bruit, semblent bien adaptés à la détection d'anomalies de cette mesure de la température de l'air intérieur. Ils sont capables de détecter des anomalies de différentes natures, y compris celles de grande intensité comme les écarts plats.

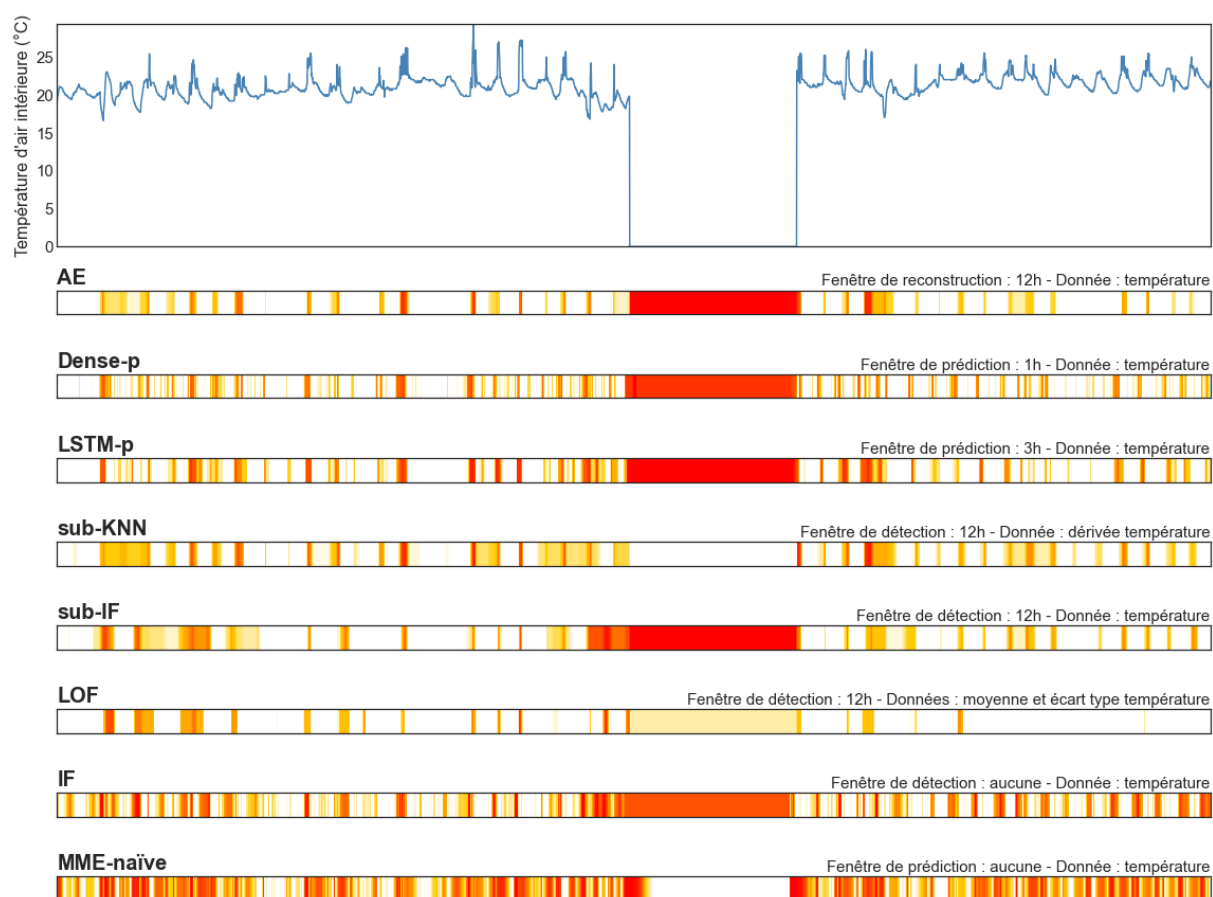


Figure 70 Exemple d'application de modèles de détection d'anomalies sur une mesure de température non étiquetée. Une couleur proche du rouge représente une observation ayant un score de suspicion proche de 1 tandis qu'une couleur proche du jaune indique un score proche de 0,5.

Pour les modèles non supervisés, la méthode de sous-échantillonnage par *Isolation Forest* (sub-IF) semble offrir des performances globalement satisfaisantes, bien qu'un peu moins précises que celles des modèles semi-supervisés. Par ailleurs, tous les modèles de voisinage ont également été testés, et ont tous obtenus des scores similaires dans la plupart des cas, à une exception près. Les modèles d'IF et de sub-IF sont les seuls capables de détecter l'écart plat sans nécessiter l'utilisation de transformations. Cependant, en raison de la taille conséquente de cette anomalie, seule l'association des transformations moyenne_{Fenêtre} et écart type_{Fenêtre} permet une détection modérée de cette anomalie. Les performances de cette association sont illustrées par les résultats du modèle LOF. Par

ailleurs, cette association semble permettre au modèle de réduire les erreurs de détection. Comme détaillé dans la partie 5.5.6.1, elle a tendance à réduire la sensibilité aux anomalies de petite taille ou de plus faible intensité en raison de l'effet supposé d'agrégation, ce qui est particulièrement visible après la période d'arrêt de fonctionnement du capteur. L'utilisation de la transformation de dérivée, en combinaison avec le modèle KNN, facilite la discrimination des scores, bien qu'elle introduise également davantage de bruit. Les scores observés pour le modèle IF sont très bruités, et cela peut donner l'impression qu'ils sont moins bons que ceux des autres modèles, car environ deux tiers du jeu d'observation sont considérés comme suspects. Cependant, il est important de noter qu'il existe des distinctions très nettes, avec pour la plupart des anomalies collectives et ponctuelles obtenant des scores très proches de 1. Dans ce cas, il pourrait être pertinent d'envisager un seuil de suspicion supérieur à la valeur de 0,5 qui a été fixée.

D'une manière générale, nous observons deux tendances dans les performances de détection des modèles. D'une part, il y a des performances relativement précises, qui sont obtenues lorsque les anomalies sont bien définies, ces performances produisent des scores de suspicion peu bruités mais moins discriminants. D'autre part, il y a des performances de détection moins précises, mais plus sensibles et discriminantes. Le premier cas semble surtout concerner les approches permettant l'emploi des fenêtres de grande taille, comme l'AE ou les transformations $\text{moyenne}_{\text{Fenêtre}}$ et $\text{écart type}_{\text{Fenêtre}}$ et les versions séquentielles des méthodes de voisinage, comme la sub-IF. Le deuxième cas concerne principalement les modèles sans approche séquentielle ou avec des fenêtres de faible taille, tels que les modèles de voisinage classiques ou les modèles prédictifs. Les résultats obtenus dans la section précédente, souvent similaires entre modèles malgré les différences notées dans cette partie, peuvent s'expliquer par le fonctionnement de l'indice AUC-ROC qui « récompense » davantage les modèles sensibles que les modèles précis. Par conséquent, en utilisant un seuil de suspicion différent pour les deux cas (p. ex., légèrement plus élevé pour les modèles de voisinage classiques et les modèles de prédiction), il serait possible d'obtenir des résultats plus similaires. En fin de compte, il semble que les modèles permettant une approche avec une fenêtre de grande taille, comme les Autoencoder, sont capables de fournir une approche de détection d'anomalie fiable et polyvalente, bien que cela puisse nécessiter l'ajustement du seuil de suspicion en fonction des besoins spécifiques de détection.

En résumé, cet exemple d'application sur des données réelles sans étiquetage a été précieux pour mettre pratique les modèles utilisés précédemment. En plus de procurer l'opportunité de vérifier certaines des observations et tendances identifiés dans la section 5.5, il a démontré la véritable applicabilité et pertinence de l'utilisation de certains modèles. Bien que le score de suspicion ait principalement une valeur visuelle, il a permis d'observer les performances des différents modèles dans un contexte réel et non contrôlé.

De plus, il est également envisageable d'utiliser ce score comme moyen de caractériser de nouveaux jeux de données, dans le but d'évaluer leur taux de contamination *a priori*. Cela pourrait s'avérer utile pour la détection d'anomalies dans de nouveaux jeux de données où l'étiquetage des données est difficile ou impossible à obtenir.

5.7. Conclusion

Ce chapitre a permis d'appliquer différentes méthodes de détection d'anomalies semi-supervisées et non supervisées à savoir, des modèles de reconstruction (plusieurs variantes d'Autoencoders et d'Encoder-Decoder), différents modèles de prédiction (*Auto-Regressive moving Average*, Dense et *Long Short-Term Memory*), des méthodes basées sur le voisinage (*Isolation Forest*, K plus proches voisins et *Local Outlier Factor*) ainsi que des variantes adaptées à l'utilisation de séquences. Il a ainsi été possible de les appliquer sur des jeux de données réels de différentes natures, dans lesquels des anomalies avaient été injectées de façon artificielle. Cette approche a permis de comparer les modèles sur des tâches de détection d'anomalies aux profils variés, tant par leur type (pics, bruits, écarts, écarts plats, plateaux et changement de niveaux) que par leur intensité (faible ou forte), leurs tailles (ponctuelles ou collectives de différentes longueurs) ainsi qu'à différentes échelles de contamination (faible ou forte). Ces anomalies ont été construites de sorte à être de nature contextuelle univariée ou multivariée afin d'évaluer la sensibilité de détection des modèles sur des observations anormales pouvant tant se distinguer que se confondre des observations normales.

Il a été constaté que la plupart des modèles, quelles que soit leurs méthodes d'entraînement ou leurs spécificités de fonctionnement, pouvaient proposer des performances médianes et maximales similaires. Cependant des distinctions ont pu être réalisées quant aux méthodes devant être employées afin d'atteindre de tels résultats. Il apparaît que les performances de détection des modèles de voisinages sont très dépendantes des variables mais aussi des associations ou des transformations de données, lesquelles sont à appliquer selon les anomalies recherchées. Une connaissance *a priori* de la nature des données employées et des anomalies visées semble indispensable pour ces méthodes. L'utilisation de fenêtres de détection de différentes tailles a cependant permis de réduire cette dépendance aux associations et transformations dans certains cas. Les performances de détection des modèles semi-supervisés de reconstruction et de prédiction semblent, au contraire, bien moins dépendantes des combinaisons, associations ou transformations de données. Elles semblent plutôt corrélées à la taille de la fenêtre de reconstruction ou de prédiction employée. Cette dernière permet, dans une majorité de cas, de faire abstraction de l'intensité de l'anomalie si elle est en mesure d'inclure l'ensemble de l'anomalie. Avec une taille de fenêtre adéquate, des anomalies

collectives présentant un profil proche des observations normales peuvent alors être détectées.

De plus, et bien que des travaux complémentaires apparaissent nécessaires, plusieurs impacts relatifs à la nature des données ont pu être constatés sur les performances de détection et donc, sur l'applicabilité des modèles. La dynamique et le bruit naturel des données semblent avoir un impact significatif sur le choix des modèles à employer au contraire de celui induit par les changements de tendances qui apparaît plus généralisé. Dans le premier cas, le choix des modèles à appliquer sera impacté tandis que dans le deuxième cas, il apparaît bénéfique d'évaluer séparément des données qui présentent des tendances de longue durée trop différentes. Un tel cas a pu être observé avec un profil d'occupation d'une pièce différent sur une longue période, lequel se répercute de façon prononcée sur les mesures de concentration de CO₂. Cependant, il est important de noter que des performances de détection similaires ont pu être constatées sur des données présentant une nature analogue telle que la température d'air intérieure et l'humidité relative. Cependant des règles explicites restent encore à définir, lesquelles pourraient faire l'objet de futurs travaux plus approfondis.

Conclusion et perspectives

Le secteur du bâtiment, représentant une part considérable de la consommation énergétique en France et en Europe, ainsi qu'une source importante d'émissions de gaz à effet de serre, est soumis à des directives nationales et internationales visant à promouvoir l'efficacité énergétique. Dans cette perspective, l'implication des occupants dans la gestion énergétique des bâtiments prend une importance croissante. L'émergence de l'Internet des Objets a ouvert la voie aux bâtiments connectés et intelligents, générant une multitude de données. Cette thèse s'est concentrée sur la détection d'anomalies dans ces données, une étape essentielle pour assurer leur fiabilité et optimiser l'utilisation énergétique des bâtiments. Après avoir présenté les principales contributions, nous dresserons également les perspectives clés qui découlent de cette thèse.

Bilan du travail réalisé

Nous avons commencé par clarifier les conceptions de « bâtiments connectés » et « bâtiments intelligents », mettant en évidence leur nature commune en tant que générateur de données grâce à des dispositifs de l'Internet des Objets (IdO). Nous avons ensuite exploré leur structure et leur fonctionnement similaires à ceux de l'IdO, mettant en avant le rôle central des capteurs, et des dispositifs connectés dans la collecte de données diverses et abondantes. Ces données alimentent une gamme variée de services visant à (1) répondre aux enjeux énergétiques et environnementaux, ainsi qu'à (2) améliorer l'expérience des utilisateurs dans le bâtiment.

Cependant, la qualité et la quantité des données sont essentielles pour le bon fonctionnement de ces services, d'où l'importance de garantir leur fiabilité. Face à divers événements anormaux pouvant perturber la disponibilité et l'accessibilité des données, nous avons exploré l'utilisation de techniques d'apprentissage automatiques pour détecter ces anomalies, offrant ainsi une solution prometteuse.

Nous avons examiné les enjeux relatifs à l'application de techniques d'apprentissage automatique à travers un état de l'art. Dans cette optique, nous avons clarifié la notion « d'anomalie » en identifiant plusieurs définitions courantes dans la littérature, soulignant leur caractère rare et unique, et distinguant entre les anomalies parasites et les anomalies d'intérêt.

Nous avons ensuite présenté les défis inhérents à la détection d'anomalies, qui découlent en partie de la variabilité et de la subjectivité associés aux anomalies, ainsi que des nombreux problèmes spécifiques à chaque cas d'étude. En explorant les principaux domaines d'application et de recherche en détection d'anomalies, nous avons mis en évidence la diversité des problèmes, des données et des types d'anomalies étudiés dans la littérature, tout en soulignant l'influence considérable que ces facteurs ont sur le choix des techniques de détection.

De plus, nous avons souligné la complexité de la sélection des méthodes de détection, en particulier pour des données aussi variées que celles issues du bâtiment, en raison de la multitude de techniques disponibles, chacune adaptée à des cas d'usage spécifiques et rarement normalisés, ainsi qu'à des contraintes inhérentes à chaque cas. Cette tâche est rendue encore plus délicate par l'absence d'études évaluant conjointement plusieurs techniques de détection d'anomalies sur des mesures et des types d'anomalies variés provenant du bâtiment.

Nous avons élaboré une méthodologie de sélection de techniques de détection d'anomalies basées sur l'apprentissage automatique, adaptée aux données et aux anomalies provenant de mesures dans le domaine du bâtiment. Notre choix de technique a principalement été guidé par les caractéristiques spécifiques des données et des types d'anomalies rencontrés. Nous avons réalisé un état de l'art pour identifier les techniques couramment utilisées dans la littérature. En utilisant ces informations, nous avons élaboré une taxonomie basée sur les caractéristiques distinctives des données et une typologie d'anomalies adaptée de la littérature. Cette taxonomie nous a servi de base pour évaluer diverses méthodes de détection d'anomalies sur différentes problématiques, afin de sélectionner celles qui conviennent le mieux à une mesure particulière, une taille d'anomalie ou à une quantité d'informations disponible.

Nous avons également mis en lumière certaines contraintes liées aux données, comme la dynamique ou le bruit, aux types d'anomalies et à la disponibilité des données, qui compliquent le processus de détection. Pour surmonter ces contraintes, nous avons suggéré que l'expertise métier, reposant sur une connaissance préalable des données et des mécanismes qui les sous-tendent, pourrait être une solution pertinente à considérer.

Nous avons réalisé une preuve de concept de notre méthodologie en appliquant des techniques d'apprentissage supervisé à un type spécifique d'anomalies étiquetées, à savoir la détection d'ouvertures de fenêtres. Ce problème équivaut à une classification binaire de motifs récurrents, similaire à la détection d'anomalies. Nous avons testé plusieurs modèles de classification supervisée, notamment les *Long Short-Term Memory* (LSTM), les *Gated Recurrent Unit* (GRU), les Séparateurs à Vaste Marge (SVM), l'Analyse Discriminante Linéaire (ADL) et les forêts aléatoires.

Nous avons évalué ce processus de détection en utilisant trois mesures différentes : la température intérieure, l'humidité relative intérieure et la concentration de CO₂ intérieure. Pour examiner l'importance de la quantité d'informations disponibles, nous avons modifié le contexte de l'étude en variant les combinaisons de ces mesures parmi un ensemble de cinq. De plus, nous avons appliqué différentes transformations aux données, issues d'une expertise métier, pour mieux caractériser les motifs spécifiques à détecter et améliorer ainsi les performances de classification.

Nos travaux ont montré que, pour la plupart des techniques, une combinaison de deux mesures, incluant la température intérieure, était nécessaire pour obtenir des performances optimales. Les modèles LSTM et GRU ont obtenu les meilleures performances avec un F₁-score de 0,80, tandis que les autres modèles se situaient autour d'un score de 0,73. Cependant, nous avons également constaté que l'application de transformations appropriées pouvait permettre d'atteindre d'excellents résultats avec la seule mesure de température. Par exemple, les modèles GRU et LSTM ont obtenu des scores respectifs de 0,77 et 0,75 en utilisant la dérivée de la température, contre un F₁-score de 0,60 et 0,57 pour la température sans transformations.

En outre, nous avons introduit six métriques d'évaluation additionnelles basées sur une expertise métier, offrant ainsi une perspective différente sur les performances des modèles. Ces métriques pourraient faciliter la sélection d'un modèle en fonction des besoins spécifiques, car elles ont révélé des différences importantes dans le nombre d'ouvertures détectées et leur précision en fonction des modèles ou des transformations appliqués.

Enfin, nous avons confirmé que deux caractéristiques relatives à la nature des données et aux types d'anomalies avaient un impact considérable sur les performances de classification des modèles : les changements rares de tendance de longue durée et la taille/intensité des ouvertures, ce qui incluent les événements anormaux.

Nous avons développé un nouveau cas d'étude basé sur la détection d'ouvertures de fenêtres, construit dans le but d'évaluer l'application de techniques d'apprentissage automatique semi-supervisées et non supervisées. Nous avons utilisé plusieurs familles de modèles, notamment des modèles de reconstruction tels que les Auto-Encodeur (AE) et les Encodeur-Décodeur (ED) - composées de réseaux de neurones denses (AE et ED), de LSTM (RAE et RED), des variants débruiteurs (DAE et DED) et des variants épars (SAE et SED) -, des modèles de prédiction tels que l'*Auto-Regressive moving Average* (ARIMA), des réseaux de neurones denses (Dense-p) et des LSTM (LSTM-p) ou encore, des modèles de voisinage tels que, l'*Isolation Forest* (IF), les K plus proches voisins (KNN) et le *Local Outlier Factor* (LOF), ainsi que leurs variantes adaptées à l'utilisation de séquences annotées sub-IF, sub-KNN et sub-LOF.

Ces modèles ont été appliqués sur des jeux de données variés qui ont été artificiellement contaminés par un générateur d'anomalies, avec des anomalies de différentes tailles et des taux de contamination variables (faible ou élevé). Les anomalies ont également été créées avec différentes intensités pour évaluer la sensibilité de détection des modèles à différents niveaux d'anormalité. Les mesures utilisées incluent la température d'air intérieure, l'humidité relative intérieure, la concentration de CO₂ intérieure et la consommation électrique.

Nos résultats ont montré que, bien que la plupart des modèles offrent des performances globalement comparables, il existe des distinctions importantes dans l'approche à adopter pour obtenir ces performances. Les modèles non supervisés de voisinage semblent fortement dépendre des variables, des associations et des transformations de données en fonction des anomalies ciblées, ce qui souligne l'importance de la connaissance *a priori* de ces anomalies. Toutefois, l'usage de fenêtres de détection variées réduit cette dépendance pour les variantes séquentielles de ces modèles.

En revanche, les performances des modèles semi-supervisés de reconstruction et de prédiction semblent plus stables et moins dépendantes des transformations ou des combinaisons de données, bien que la taille de la fenêtre employée puisse avoir un impact significatif. Une fenêtre adéquate peut permettre de détecter des anomalies de faible intensité, même si elles sont proches des observations normales.

La nature des données a une influence significative sur les performances de détection, bien que des études complémentaires semblent nécessaires pour consolider ces observations. Les données similaires, telles que la température d'air intérieure et l'humidité relative intérieure, présentent des performances similaires. La prise en compte de la dynamique et du bruit des données sont essentiels dans le choix des modèles, tandis que les changements de tendances détériorent les performances de manière générale, indépendamment des modèles. Les AE se montrent performants, précis et stables sur des données complexes et bruitées, tandis que les modèles de voisinage sont plus adaptés aux données avec peu de bruits ou de variations.

La séparation des données en deux jeux de tendances différentes peut améliorer les performances de détection, des recherches complémentaires sont nécessaires pour tirer des conclusions définitives. Enfin, un exemple d'application sur des anomalies réelles a été présenté, utilisant un score de suspicion pour évaluer le degré d'anormalité d'un jeu par un aspect visuel, lequel pourrait également être employé afin de caractériser un jeu de données.

En résumé, cette thèse nous a permis d'identifier et de formaliser de nombreux éléments distinctifs des données et des anomalies provenant de mesures du bâtiment,

impactant ainsi notre choix et nos performances des modèles d'apprentissage automatiques utilisés pour la détection d'anomalies. Nous avons développé une méthodologie de sélection basée sur une taxonomie pour caractériser les problématiques liés à des données et anomalies, ce qui nous a permis d'évaluer divers modèles de différentes familles d'apprentissage automatique. L'analyse statistique des résultats nous a révélé certaines tendances liées aux données, aux anomalies, aux modèles et aux approches employées.

Nous avons également mis en évidence l'importance de l'expertise métier dans le processus de détection et d'évaluation des résultats. Plutôt que de chercher la meilleure technique, notre objectif était de comprendre les avantages et les limites pour notre domaine d'étude spécifique. Nous avons ainsi cherché à éclairer les observations, les tendances, et les réflexions pour orienter les futurs travaux dans ce domaine.

Perspectives

Les perspectives de ces travaux de thèse ouvrent la voie à de nombreuses opportunités de recherche dans le domaine de la détection d'anomalies appliquée aux mesures du bâtiment. Tout d'abord, il serait pertinent de procéder à des évaluations similaires sur d'autres types de données quantitatives présentant des dynamiques et des profils différents. Cela pourrait inclure les mesures de température d'air extérieure, de rayonnement solaire, et autres données couramment utilisées dans le secteur du bâtiment. Une étude plus approfondie des données qualitatives serait également essentielle car elles sont largement utilisées dans le bâtiment, notamment pour la gestion des systèmes (par exemple, les mesures d'ouverture de vannes, les mesures d'occupation ou de détection de mouvement). Cette démarche serait bénéfique tant pour le domaine du bâtiment que d'autres domaines, car il semble y avoir un manque de travaux traitant de données qualitatives dans la littérature, quelle que soit la discipline.

En outre, il serait intéressant d'évaluer l'impact de l'ajout de combinaisons de données variées dans la tâche de détection d'anomalies. Ces travaux se sont principalement concentrés sur des combinaisons limitées de données, et il existe probablement des corrélations intéressantes qui n'ont pas encore été explorées. Par exemple, l'étude de la corrélation entre les mesures de concentration de CO₂ et les mesures d'occupation ou de mouvement, ou entre les mesures de température d'air intérieure et les mesures de consommation d'un système de chauffage, pourrait être très instructive. Une recherche plus axée sur les données, plutôt que sur les anomalies ou les modèles, avec des données plus variées et des combinaisons complémentaires, pourrait permettre de mieux comprendre le rôle des différents composants liés à la nature des données dans le

processus de détection d'anomalies. Cela pourrait également aider à formaliser des critères distinctifs pour déterminer la similarité entre différentes mesures lorsqu'elles sont utilisées dans la détection d'anomalies.

Concernant les modèles, bien que plusieurs approches aient été testées, il reste de nombreuses possibilités à explorer, notamment l'utilisation de méthodes de *clustering*. Dans cette perspective, l'application conjointe de plusieurs modèles, aussi appelés modèles hybrides, semble être une piste prometteuse à explorer. Cette approche consiste à combiner différents modèles, par exemple en utilisant l'un pour prétraiter les données et l'autre pour détecter les anomalies, ce qui pourrait ouvrir de nouvelles perspectives.

Un exemple couramment rencontré dans la littérature, bien que non abordé dans cette thèse, est l'association de méthodes semi-supervisées de reconstruction avec des méthodes non supervisées de *clustering* ou de voisinage. Par exemple, l'entraînement d'un modèle de d'Auto-Encodeur variationnel (VAE) selon une approche semi-supervisée a été préparé dans le but de réaliser la détection d'anomalies en utilisant plusieurs méthodes non supervisées sur son espace latent. D'autres modèles comme les *Generative Adversarial Networks* (GAN) pourraient également être explorés, ainsi que des variantes des méthodes déjà utilisées. En effet, seuls les modèles représentatifs d'une méthode et d'une famille de techniques ont été utilisés jusqu'à présent, mais il existe des variantes performantes ou différentes de ces modèles qui méritent d'être évaluées, comme les méthodes d'*Extended Isolation Forest* (EIF) ou la *Vector Auto-Regression Integrated Moving Average* (VARIMA).

En ce qui concerne les anomalies, une approche logique à poursuivre serait de répondre à la demande croissante dans la littérature en matière de détection d'anomalies globales, ou plus précisément de séries temporelles anormales. Cependant, il est important de noter que cette approche peut être délicate à mettre en œuvre en raison de la quantité considérable de données requises pour détecter de telles anomalies, ainsi que des besoins en puissance de calcul associés.

Une autre approche intéressante à considérer serait d'évaluer plus clairement la précision des modèles de détection en fonction de l'intensité ou de la taille des anomalies. Cela permettrait de privilégier certains modèles par rapport à d'autres en fonction de l'impact attendu des anomalies dans des scénarios d'application spécifiques. Par exemple, dans le cas d'un retour d'expérience suite à une rénovation, il pourrait être plus préjudiciable de ne pas détecter une légère baisse anormale de la température de 0,5 °C sur une période de 24 heures que de ne pas détecter une augmentation anormale de 2 °C sur 2 heures. La mise en place de métriques d'évaluation similaires à celles utilisées pour la détection d'ouverture de fenêtres pourrait donc représenter une contribution pertinente pour des travaux futurs dans ce domaine.

Enfin, deux derniers éléments majeurs méritent d'être pris en compte pour l'avenir de ces travaux. Tout d'abord, il serait essentiel d'appliquer ces méthodes à des cas réels et monitorés pour évaluer pleinement leur applicabilité et de déterminer leur pertinence dans des situations concrètes. Dans le contexte de l'énergétique du bâtiment, l'emploi de techniques de détection pourrait contribuer à évaluer la qualité des données, ce qui pourrait, le cas échéant, conduire à des améliorations. Par exemple, l'utilisation des données d'un capteur influencé par le rayonnement solaire dans le cadre d'un retour d'expérience après à une rénovation aurait pu conduire à des erreurs fréquentes dans les consignes de chauffage. Ainsi, pour mettre en pratique les connaissances acquises au cours de ces travaux sur un cas d'application réel, les modèles préalablement entraînés ainsi que de nouveaux modèles seront utilisés pour prétraiter les données utilisées dans le cadre d'une autre thèse visant à développer des recommandations personnalisées pour les occupants afin de réaliser des économies d'énergie en se basant sur l'estimation de leur comportement à partir de données mesurées.

D'autre part, une perspective majeure consiste à quantifier le degré d'anormalité d'un jeu d'observations, comme cela a été brièvement abordé. Couplé à des métriques d'évaluations adaptées, cela pourrait permettre de mieux quantifier la qualité d'un jeu d'observations au regard des applications envisagées. Une telle perspective pourrait constituer un apport conséquent pour le domaine. De plus, il serait pertinent d'explorer la possibilité d'appliquer une correction adéquate à chaque type de mesure et d'anomalies. En effet, certains modèles utilisés dans ces travaux, tels que ceux de reconstruction et de prédiction, semblent *a priori* pertinents pour cette tâche. Cependant, il convient de noter que les modèles employés jusqu'à présent semblent trop sensibles aux anomalies pour proposer une correction adéquate. Par conséquent, des travaux approfondis visant à élaborer un modèle capable de détecter les anomalies tout en proposant une correction appropriée représentent une perspective prometteuse à explorer.

Publications

Certaines des contributions de cette thèse ont été publiées, dans des journaux ou des conférences. Ces publications sont listées ci-dessous.

Article de journal

de Rautlin de la Roy, E., Recht, T., Zemmari, A., Bourreau, P., Mora, L., 2023. *Deep learning models for building window-openings detection in heating season*. Building and Environment 231, 110019. <https://doi.org/10.1016/j.buildenv.2023.110019>.

Articles de conférence

de Rautlin de la Roy, E., Recht, T., Zemmari, A., Bourreau, P., Mora, L., 2021. *Data analytics for smart buildings: a classification method for anomaly detection for measured data*. J. Phys.: Conf. Ser. 2042, 012015. <https://doi.org/10.1088/1742-6596/2042/1/012015>.

de Rautlin de la Roy, E., Recht, T., Zemmari, A., Bourreau, P., Houchat, K., Mora L., 2022. Détection d'ouvertures de fenêtres au moyen de réseaux de neurones artificiels LSTM. IBPSA France, Châlons en Champagne.

de Rautlin de la Roy, E., Recht, T., Mora L., 2020. Vérification de performance énergétique de maisons « à énergie positive ». IBPSA France.

Bibliographie

- Adhikari, D., Jiang, W., Zhan, J., He, Z., Rawat, D.B., Aickelin, U., Khorshidi, H.A., 2023. A Comprehensive Survey on Imputation of Missing Data in Internet of Things. *ACM Comput. Surv.* 55, 1–38. <https://doi.org/10.1145/3533381>.
- Aggarwal, C.C., 2017. *Outlier Analysis*. Springer International Publishing, Cham. <https://doi.org/10.1007/978-3-319-47578-3>.
- Ahmed, O., Sezer, N., Ouf, M., Wang, L. (Leon), Hassan, I.G., 2023. State-of-the-art review of occupant behavior modeling and implementation in building performance simulation. *Renewable and Sustainable Energy Reviews* 185, 113558. <https://doi.org/10.1016/j.rser.2023.113558>.
- Alduailij, Mona A., Petri, I., Rana, O., Alduailij, Mai A., Aldawood, A.S., 2021. Forecasting peak energy demand for smart buildings. *J Supercomput* 77, 6356–6380. <https://doi.org/10.1007/s11227-020-03540-3>.
- Anastasiadou, M., Santos, V., Dias, M.S., 2021. Machine Learning Techniques Focusing on the Energy Performance of Buildings: A Dimensions and Methods Analysis. *Buildings* 12, 28. <https://doi.org/10.3390/buildings12010028>.
- Angiulli, F., Fassetti, F., Ferragina, L., 2022. LatentOut : an unsupervised deep anomaly detection approach exploiting latent space distribution. *Machine Learning*. <https://doi.org/10.1007/s10994-022-06153-4>.
- Araya, D.B., Grolinger, K., ElYamany, H.F., Capretz, M.A.M., Bitsuamlak, G., 2016. Collective contextual anomaly detection framework for smart buildings, in: 2016 International Joint Conference on Neural Networks (IJCNN). IEEE, Vancouver, BC, Canada, pp. 511–518. <https://doi.org/10.1109/IJCNN.2016.7727242>.
- Ardakanian, O., Bhattacharya, A., Culler, D., 2016. Non-Intrusive Techniques for Establishing Occupancy Related Energy Savings in Commercial Buildings, in: *Proceedings of the 3rd ACM International Conference on Systems for Energy-Efficient Built Environments*. ACM, Palo Alto CA USA, pp. 21–30. <https://doi.org/10.1145/2993422.2993574>.
- Ariyaluran Habeeb, R.A., Nasaruddin, F., Gani, A., Amanullah, M.A., Abaker Targio Hashem, I., Ahmed, E., Imran, M., 2022. Clustering-based real-time anomaly detection—A breakthrough in big data technologies. *Trans Emerging Tel Tech* 33. <https://doi.org/10.1002/ett.3647>.
- Asch, G., 2010. *Les capteurs en instrumentation industrielle*. Dunod, Paris.
- Assegond, C., 2015. *Affichage des consommations et réflexivité des ménages : construire une culture domestique de l'énergie par l'information*. Université de Tours.

- Aste, N., Manfren, M., Marenzi, G., 2017. Building Automation and Control Systems and performance optimization: A framework for analysis. *Renewable and Sustainable Energy Reviews* 75, 313–330. <https://doi.org/10.1016/j.rser.2016.10.072>.
- Audibert, J., 2021. Unsupervised Anomaly Detection in Time-Series (Thèse de doctorat). Université de la Sorbonne. <https://theses.hal.science/tel-03681871>.
- Ayadi, A., Ghorbel, O., Obeid, A.M., Abid, M., 2017. Outlier detection approaches for wireless sensor networks: A survey. *Computer Networks* 129, 319–333. <https://doi.org/10.1016/j.comnet.2017.10.007>.
- Bachlin, M., Plotnik, M., Roggen, D., Maidan, I., Hausdorff, J.M., Giladi, N., Troster, G., 2010. Wearable Assistant for Parkinson's Disease Patients With the Freezing of Gait Symptom. *IEEE Trans. Inform. Technol. Biomed.* 14, 436–446. <https://doi.org/10.1109/TITB.2009.2036165>.
- Bailey, K.D., 1994. *Typologies and taxonomies: an introduction to classification techniques, Quantitative applications in the social sciences*. Sage Publications, Thousand Oaks, Calif.
- Baloglu, O., Latifi, S.Q., Nazha, A., 2021. What is machine learning? *Arch Dis Child Educ Pract Ed.* 2022 Oct;107(5):386–388. doi: 10.1136/archdischild-2020-319415. Epub 2021 Feb 8. PMID: 33558304.
- Bank, D., Koenigstein, N., Giryes, R., 2021. Autoencoders. <https://doi.org/10.48550/arXiv.2003.05991>.
- Bansal, M., Chana, I., Clarke, S., 2021. A Survey on IoT Big Data: Current Status, 13 V's Challenges, and Future Directions. *ACM Comput. Surv.* 53, 1–59. <https://doi.org/10.1145/3419634>.
- Barnett, V., Lewis, T., 1994. *Outliers in statistical data*, Repr. ed, Wiley series in probability and mathematical statistics Applied probability and statistics. Wiley, New York.
- Bashir, M.R., Gill, A.Q., Beydoun, G., 2022. A Reference Architecture for IoT-Enabled Smart Buildings. *SN COMPUT. SCI.* 3, 493. <https://doi.org/10.1007/s42979-022-01401-9>.
- Bataineh, B., 2022. Fast Component Density Clustering in Spatial Databases: A Novel Algorithm. *Information* 13, 477. <https://doi.org/10.3390/info13100477>.
- Belgiu, M., Drăguț, L., 2016. Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing* 114, 24–31. <https://doi.org/10.1016/j.isprsjprs.2016.01.011>.
- Bemdt, D.J., Clifford, J., 1994. Using dynamic time warping to find patterns in time series. *AAAI Workshop on Knowledge Discovery in Databases*.
- Ben-Ari, M., Mondada, F., 2018. *Elements of Robotics*. Springer International Publishing, Cham. <https://doi.org/10.1007/978-3-319-62533-1>.
- Bengio, Y., Courville, A., Vincent, P., 2014. Representation Learning: A Review and New Perspectives. *IEEE Trans. Pattern Anal. Mach. Intell.* 35 1798–828.
- Bengio, Y., Simard, P., Frasconi, P., 1994. Learning Long-Term Dependencies with Gradient Descent is Difficult. *IEEE Trans Neural Network* 1994;5:157–66. <https://doi.org/10.1109/72.279181>.

- Benkabou, S.-E., 2018. Détection d'anomalies dans les séries temporelles: application aux masses de données sur les pneumatiques (Thèse de doctorat). Université de Lyon. <https://theses.hal.science/tel-01839074>.
- Berrar, D., 2019. Cross-Validation, in: Encyclopedia of Bioinformatics and Computational Biology. Elsevier, pp. 542–545. <https://doi.org/10.1016/B978-0-12-809633-8.20349-X>.
- Bhuyan, M.H., Bhattacharyya, D.K., Kalita, J.K., 2014. Network Anomaly Detection: Methods, Systems and Tools. IEEE Commun. Surv. Tutorials 16, 303–336. <https://doi.org/10.1109/SURV.2013.052213.00046>.
- Bianco, A.M., García Ben, M., Martínez, E.J., Yohai, V.J., 2001. Outlier Detection in Regression Models with ARIMA Errors using Robust Estimates: Outlier Detection in Regression Models with ARIMA Errors. J. Forecast. 20, 565–579. <https://doi.org/10.1002/for.768>.
- Bishop, C.M., 2006. Pattern recognition and machine learning, Information science and statistics. Springer, New York.
- Bishop, C.M., 1994. Neural networks and their applications. Neural networks 65.
- Biswas, J., Kobayashi, H., Wong, L., Abdulrazak, B., Mokhtari, M. (Eds.), 2013. Inclusive Society: Health and Wellbeing in the Community, and Care at Home: 11th International Conference on Smart Homes and Health Telematics, ICOST 2013, Singapore, June 19-21, 2013. Proceedings, Lecture Notes in Computer Science. Springer Berlin Heidelberg, Berlin, Heidelberg. <https://doi.org/10.1007/978-3-642-39470-6>.
- Blázquez-García, A., Conde, A., Mori, U., Lozano, J.A., 2022. A Review on Outlier/Anomaly Detection in Time Series Data. ACM Comput. Surv. 54, 1–33. <https://doi.org/10.1145/3444690>.
- Bonte, M., Thellier, F., Lartigue, B., 2014. Impact of occupant's actions on energy building performance and thermal sensation. Energy and Buildings 76, 219–227. <https://doi.org/10.1016/j.enbuild.2014.02.068>.
- Borhani Afuosi, M., Zoghi, M.R., 2020. Indoor positioning based on improved weighted KNN for energy management in smart buildings. Energy and Buildings 212, 109754. <https://doi.org/10.1016/j.enbuild.2019.109754>.
- Boriah, S., Chandola, V., Kumar, V., 2008. Similarity Measures for Categorical Data: A Comparative Evaluation, in: Proceedings of the 2008 SIAM International Conference on Data Mining. Society for Industrial and Applied Mathematics, pp. 243–254. <https://doi.org/10.1137/1.9781611972788.22>.
- Bosman, H., 2016. Anomaly detection in networked embedded sensor systems (Ph.D. dissertation). Technische Universiteit Eindhoven, Department of Electrical Engineering, Proefschrift.
- Boudellal, M., 2014. Smart home: habitat connecté, installations domotiques et multimédia, Collection Technique et ingénierie. "L'Usine nouvelle" Dunod, Paris.
- Bourdon, M., 2021. Détection d'intrusion basée sur l'analyse de compteurs matériels pour des objets connectés (Thèse de doctorat). INSA de Toulouse. <https://laas.hal.science/tel-03572845v2>.

- Braei, M., Wagner, S., 2020. Anomaly Detection in Univariate Time-series: A Survey on the State-of-the-Art. <https://doi.org/10.48550/arXiv.2004.00433>.
- Breiman, L., 2001. Random Forest. *Machine Learning* 45, 5–32. <https://doi.org/10.1023/A:1010933404324>.
- Breiman, L., 1996. Bagging predictors. *Mach Learn* 24, 123–140. <https://doi.org/10.1007/BF00058655>.
- Breunig, M.M., Kriegel, H.-P., Ng, R.T., Sander, J., 2000. LOF: Identifying Density-Based Local Outliers. In *Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data*(pp. 93–104).
- Brockwell, P.J., Davis, R.A., 2010. *Introduction to time series and forecasting*, Second edition, (corrected at 8th. printing 2010). ed, Springer texts in statistics. Springer, New York, NY.
- Buckman, A.H., Mayfield, M., B.M. Beck, S., 2014. What is a Smart Building? *Smart and Sustainable Built Environment* 3, 92–109. <https://doi.org/10.1108/SASBE-01-2014-0003>.
- Burgas, L., Colomer, J., Melendez, J., Gamero, F.I., Herraiz, S., 2021. Integrated Unfold-PCA Monitoring Application for Smart Buildings: An AHU Application Example. *Energies* 14, 235. <https://doi.org/10.3390/en14010235>.
- Candanedo, L.M., Feldheim, V., 2016. Accurate occupancy detection of an office room from light, temperature, humidity and CO₂ measurements using statistical learning models. *Energy and Buildings* 112, 28–39. <https://doi.org/10.1016/j.enbuild.2015.11.071>.
- Candelieri, A., 2017. Clustering and Support Vector Regression for Water Demand Forecasting and Anomaly Detection. *Water* 9, 224. <https://doi.org/10.3390/w9030224>.
- Cao, C., Blaise, A., Verwer, S., Rebecchi, F., 2022. Learning State Machines to Monitor and Detect Anomalies on a Kubernetes Cluster, in: *Proceedings of the 17th International Conference on Availability, Reliability and Security*. pp. 1–9. <https://doi.org/10.1145/3538969.3543810>.
- Carreño, A., Inza, I., Lozano, J.A., 2020. Analyzing rare event, anomaly, novelty and outlier detection terms under the supervised classification framework. *Artif Intell Rev* 53, 3575–3594. <https://doi.org/10.1007/s10462-019-09771-y>.
- Casado-Vara, R., Vale, Z., Prieto, J., Corchado, J., 2018. Fault-Tolerant Temperature Control Algorithm for IoT Networks in Smart Buildings. *Energies* 11, 3430. <https://doi.org/10.3390/en11123430>.
- Chalapathy, R., Chawla, S., 2019. Deep Learning for Anomaly Detection: A Survey. <https://doi.org/10.48550/arXiv.1901.03407>.
- Chandola, V., Banerjee, A., Kumar, V., 2012. Anomaly Detection for Discrete Sequences: A Survey. *IEEE Trans. Knowl. Data Eng.* 24, 823–839. <https://doi.org/10.1109/TKDE.2010.235>.
- Chandola, V., Banerjee, A., Kumar, V., 2009. Anomaly detection: A survey. *ACM Comput. Surv.* 41, 1–58. <https://doi.org/10.1145/1541880.1541882>.

- Chawla, N., Wang, W. (Eds.), 2017. Proceedings of the 2017 SIAM International Conference on Data Mining. Society for Industrial and Applied Mathematics, Philadelphia, PA. <https://doi.org/10.1137/1.9781611974973>.
- Chen, J., Li, J., Chen, W., Wang, Y., Jiang, T., 2020. Anomaly detection for wind turbines based on the reconstruction of condition parameters using stacked denoising autoencoders. *Renewable Energy* 147, 1469–1480. <https://doi.org/10.1016/j.renene.2019.09.041>.
- Chen, Z., Masood, M.K., Soh, Y.C., 2016. A fusion framework for occupancy estimation in office buildings based on environmental sensor data. *Energy and Buildings* 133, 790–798. <https://doi.org/10.1016/j.enbuild.2016.10.030>.
- Cheng, H., Tan, P.-N., Potter, C., Klooster, S., 2009. Detection and Characterization of Anomalies in Multivariate Time Series, in: Proceedings of the 2009 SIAM International Conference on Data Mining. Society for Industrial and Applied Mathematics, pp. 413–424. <https://doi.org/10.1137/1.9781611972795.36>.
- Cheng, Z., Zou, C., Dong, J., 2019. Outlier detection using isolation forest and local outlier factor, in: Proceedings of the Conference on Research in Adaptive and Convergent Systems, RACS '19. Association for Computing Machinery, New York, NY, USA, pp. 161–168. <https://doi.org/10.1145/3338840.3355641>.
- Chou, J.-S., Telaga, A.S., 2014. Real-time detection of anomalous power consumption. *Renewable and Sustainable Energy Reviews* 33, 400–411. <https://doi.org/10.1016/j.rser.2014.01.088>.
- Chouchane, A., 2018. Une nouvelle approche de détection des anomalies dans les réseaux multidimensionnels. Montréal (Québec, Canada), Université du Québec à Montréal, Maîtrise en informatique. <http://archipel.uqam.ca/11333/1/M15400.pdf>.
- Chung, J., Gulcehre, C., Cho, K., Bengio, Y., 2014. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. <https://doi.org/10.48550/arXiv.1412.3555>.
- Cleveland, R.B., Cleveland, W.S., Jean E., M., Irma J., T., 1990. STL: A seasonal-trend decomposition procedure based on loess (with discussion). <https://api.semanticscholar.org/CorpusID:64570714>.
- Coma-Puig, B., Carmona, J., Gavalda, R., Alcoverro, S., Martin, V., 2016. Fraud Detection in Energy Consumption: A Supervised Approach, in: 2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA). IEEE, Montreal, QC, Canada, pp. 120–129. <https://doi.org/10.1109/DSAA.2016.19>.
- Cortes, C., Vapnik, V., 1995. Support-vector networks. *Machine Learning* 20, 273–297. <https://doi.org/10.1007/BF00994018>.
- Costa-jussà, M.R., Nuez, Á., Segura, C., 2018. Experimental Research on Encoder-Decoder Architectures with Attention for Chatbots. *CyS* 22. <https://doi.org/10.13053/cys-22-4-3060>.
- Cramer, J.S., 2003. The Origins of Logistic Regression. *SSRN Journal*. <https://doi.org/10.2139/ssrn.360300>.
- Csillik, O., Belgiu, M., Asner, G.P., Kelly, M., 2019. Object-Based Time-Constrained Dynamic Time Warping Classification of Crops Using Sentinel-2. *Remote Sensing* 11, 1257. <https://doi.org/10.3390/rs11101257>.

- Dai, X., Liu, J., Zhang, X., 2020. A review of studies applying machine learning models to predict occupancy and window-opening behaviours in smart buildings. *Energy and Buildings* 223, 110159. <https://doi.org/10.1016/j.enbuild.2020.110159>.
- Delzendeh, E., Wu, S., Lee, A., Zhou, Y., 2017. The impact of occupants' behaviours on building energy analysis: A research review. *Renewable and Sustainable Energy Reviews* 80, 1061–1071. <https://doi.org/10.1016/j.rser.2017.05.264>.
- Deng, H., Runger, G., 2013. Gene selection with guided regularized random forest. *Pattern Recognition* 46, 3483–3489. <https://doi.org/10.1016/j.patcog.2013.05.018>.
- Desforges, M.J., Jacob, P.J., Cooper, J.E., 1998. Applications of probability density estimation to the detection of abnormal conditions in engineering. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science* 212, 687–703. <https://doi.org/10.1243/0954406981521448>.
- Ding, D., Cooper, R.A., Pasquina, P.F., Fici-Pasquina, L., 2011. Sensor technology for smart homes. *Maturitas* 69, 131–136. <https://doi.org/10.1016/j.maturitas.2011.03.016>.
- Directive 2010/31/UE, 2010. Directive 2010/31/UE du Parlement européen et du Conseil du 19 mai 2010 sur la performance énergétique des bâtiments.
- Djenouri, D., Laidi, R., Djenouri, Y., Balasingham, I., 2020. Machine Learning for Smart Building Applications: Review and Taxonomy. *ACM Comput. Surv.* 52, 1–36. <https://doi.org/10.1145/3311950>.
- Do, H., Cetin, K.S., 2018. Evaluation of the causes and impact of outliers on residential building energy use prediction using inverse modeling. *Building and Environment* 138, 194–206. <https://doi.org/10.1016/j.buildenv.2018.04.039>.
- Doersch, C., 2021. Tutorial on Variational Autoencoders. Preprint at <https://arxiv.org/abs/1606.05908>.
- Dong, B., Prakash, V., Feng, F., O'Neill, Z., 2019. A review of smart building sensing system for better indoor environment control. *Energy and Buildings* 199, 29–46. <https://doi.org/10.1016/j.enbuild.2019.06.025>.
- Ducharlet, K., Travé-Massuyès, L., Lann, M.-V.L., Miloudi, Y., 2022. Etude des méthodes de détection d'anomalies non supervisées appliquées aux flux de données. 20èmes Rencontres des Jeunes Chercheurs en Intelligence Artificielle (RJCI 2022), Jun 2022, Saint-Etienne, France. hal-03765550.
- Duda, R.O., Hart, P.E., Stork, D.G., 2001. *Pattern classification*, 2. ed. ed, A Wiley-Interscience publication. Wiley, New York Weinheim.
- Emmott, A., Das, S., Dietterich, T., Fern, A., Wong, W.-K., 2016. A Meta-Analysis of the Anomaly Detection Problem. <https://arxiv.org/abs/1503.01158>.
- Erhan, L., Ndubuaku, M., Di Mauro, M., Song, W., Chen, M., Fortino, G., Bagdasar, O., Liotta, A., 2021. Smart anomaly detection in sensor systems: A multi-perspective review. *Information Fusion* 67, 64–79. <https://doi.org/10.1016/j.inffus.2020.10.001>.
- European Climate Law, 2021. European Climate Law.
- European Environment Agency, 2015. Final energy consumption by sector and fuel.

- Fan, C., Xiao, F., Yan, C., 2015. A framework for knowledge discovery in massive building automation data and its application in building diagnostics. *Automation in Construction* 50, 81–90. <https://doi.org/10.1016/j.autcon.2014.12.006>.
- Fan, C., Xiao, F., Zhao, Y., Wang, J., 2018. Analytical investigation of autoencoder-based methods for unsupervised anomaly detection in building energy data. *Applied Energy* 211, 1123–1135. <https://doi.org/10.1016/j.apenergy.2017.12.005>.
- Fernando, T., Gammulle, H., Denman, S., Sridharan, S., Fookes, C., 2022. Deep Learning for Medical Anomaly Detection – A Survey. *ACM Comput. Surv.* 54, 1–37. <https://doi.org/10.1145/3464423>.
- Foorthuis, R., 2021. On the nature and types of anomalies: a review of deviations in data. *Int J Data Sci Anal* 12, 297–331. <https://doi.org/10.1007/s41060-021-00265-1>.
- Foulon, L., 2020. Détection d'anomalies dans les flux de données par structure d'indexation et approximation. Application à l'analyse en continu des flux de messages du système d'information de la SNCF (Thèse de doctorat). Université de Lyon. <https://theses.hal.science/tel-03125747>.
- Fourure, D., Javaid, M.U., Posocco, N., Tihon, S., 2021. Anomaly Detection: How to Artificially Increase your F1-Score with a Biased Evaluation Protocol. <https://doi.org/10.48550/arXiv.2106.16020>.
- Fujioka, T., Kubota, K., Mori, M., Kikuchi, Y., Katsuta, L., Kimura, M., Yamaga, E., Adachi, M., Oda, G., Nakagawa, T., Kitazume, Y., Tateishi, U., 2020. Efficient Anomaly Detection with Generative Adversarial Network for Breast Ultrasound Imaging. *Diagnostics* 10, 456. <https://doi.org/10.3390/diagnostics10070456>.
- Gallagher, C.V., Leahy, K., O'Donovan, P., Bruton, K., O'Sullivan, D.T.J., 2018. Development and application of a machine learning supported methodology for measurement and verification (M&V) 2.0. *Energy and Buildings* 167, 8–22. <https://doi.org/10.1016/j.enbuild.2018.02.023>.
- Ghojogh, B., Crowley, M., 2019. Linear and Quadratic Discriminant Analysis: Tutorial. <https://doi.org/10.48550/arXiv.1906.02590>.
- GIEC, 2022. CLIMATE CHANGE 2022 IMPACTS, ADAPTATION AND VULNERABILITY, 1ST ED. ed. CAMBRIDGE UNIV PRESS UK, CAMBRIDGE.
- Girin, L., Leglaive, S., Bie, X., Diard, J., Hueber, T., Alameda-Pineda, X., 2021. Dynamical Variational Autoencoders: A Comprehensive Review. *FNT in Machine Learning* 15, 1–175. <https://doi.org/10.1561/22000000089>.
- Gokcen, I., Peng, J., 2002. Comparing Linear Discriminant Analysis and Support Vector Machines, in: Yakhno, T. (Ed.), *Advances in Information Systems, Lecture Notes in Computer Science*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 104–113. https://doi.org/10.1007/3-540-36077-8_10.
- Goldsborough, P., 2016. A Tour of TensorFlow. <https://doi.org/10.48550/arXiv.1610.01178>.
- Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y., 2014. Generative Adversarial Networks. *Commun. ACM* 63, 11 (November 2020), 139–144. <https://doi.org/10.1145/3422622>.

- Greenwald, S.D., Patil, R.S., Mark, R.G., 1992. Improved detection and classification of arrhythmias in noise-corrupted electrocardiograms using contextual information within an expert system. *Biomed Instrum Technol* 26, 124–132.
- Greff, K., Srivastava, R.K., Koutník, J., Steunebrink, B.R., Schmidhuber, J., 2017. LSTM: A Search Space Odyssey. *IEEE Trans. Neural Netw. Learning Syst.* 28, 2222–2232. <https://doi.org/10.1109/TNNLS.2016.2582924>.
- Greis, R., Ries, T., Nguyen, C.D., 2018. Comparing Prediction Methods in Anomaly Detection: An Industrial Evaluation.
- Grillone, B., Danov, S., Sumper, A., Cipriano, J., Mor, G., 2020. A review of deterministic and data-driven methods to quantify energy efficiency savings and to predict retrofitting scenarios in buildings. *Renewable and Sustainable Energy Reviews* 131, 110027. <https://doi.org/10.1016/j.rser.2020.110027>.
- Grubbs, F.E., 1969. Procedures for Detecting Outlying Observations in Samples. *Technometrics* 11, 1–21. <https://doi.org/10.1080/00401706.1969.10490657>.
- Gunay, H.B., Shi, Z., 2020. Cluster analysis-based anomaly detection in building automation systems. *Energy and Buildings* 228, 110445. <https://doi.org/10.1016/j.enbuild.2020.110445>.
- Gundersen, O.E., Coakley, K., Kirkpatrick, C., Gil, Y., 2023. Sources of Irreproducibility in Machine Learning: A Review. <https://doi.org/10.48550/arXiv.2204.07610>.
- Gupta, M., Gao, J., Aggarwal, C.C., Han, J., 2014. Outlier Detection for Temporal Data: A Survey. *IEEE Trans. Knowl. Data Eng.* 26, 2250–2267. <https://doi.org/10.1109/TKDE.2013.184>.
- Hair, J.F. (Ed.), 2010. *Multivariate data analysis: a global perspective*, 7. ed., global ed. ed. Pearson, Upper Saddle River, NJ Munich.
- Haldi, F., Robinson, D., 2009. Interactions with window openings by office occupants. *Building and Environment* 44, 2378–2395. <https://doi.org/10.1016/j.buildenv.2009.03.025>.
- Haldi, F., Robinson, D., 2008. On the behaviour and adaptation of office occupants. *Building and Environment* 43, 2163–2177. <https://doi.org/10.1016/j.buildenv.2008.01.003>.
- Hammi, M.T., 2019. *Sécurisation de l'Internet des objets (Thèse de doctorat)*. Université Paris Saclay. <https://theses.hal.science/tel-01997261>.
- Handelman, G.S., Kok, H.K., Chandra, R.V., Razavi, A.H., Huang, S., Brooks, M., Lee, M.J., Asadi, H., 2019. Peering Into the Black Box of Artificial Intelligence: Evaluation Metrics of Machine Learning Methods. *American Journal of Roentgenology* 212, 38–43. <https://doi.org/10.2214/AJR.18.20224>.
- Haq, E.U., Huang, J., Xu, H., Li, K., Ahmad, F., 2021. A hybrid approach based on deep learning and support vector machine for the detection of electricity theft in power grids. *Energy Reports* 7, 349–356. <https://doi.org/10.1016/j.egyr.2021.08.038>.
- Hariri, S., Kind, M.C., Brunner, R.J., 2021. Extended Isolation Forest. *IEEE Trans. Knowl. Data Eng.* 33, 1479–1489. <https://doi.org/10.1109/TKDE.2019.2947676>.

- Hasebrook, N., Morsbach, F., Kannengießer, N., Zöller, M., Franke, J., Lindauer, M., Hutter, F., Sunyaev, A., 2023. Practitioner Motives to Select Hyperparameter Optimization Methods.
- Hassan, Q.F., Khan, A.R., Madani, S.A., 2020. Internet of Things: Challenges, Advances and Applications. Chapman & Hall/CRC computer and informationscience series. CRC Press: Boca Raton, FL, USA,
- Hastie, T., Tibshirani, R., Friedman, J., 2009. The Elements of Statistical Learning, Springer Series in Statistics. Springer New York, New York, NY. <https://doi.org/10.1007/978-0-387-84858-7>.
- Hawkins, D.M., 1980. Identification of Outliers. Springer Netherlands, Dordrecht. <https://doi.org/10.1007/978-94-015-3994-4>.
- He, Z., Xu, X., Deng, S., 2003. Discovering cluster-based local outliers. Pattern Recognition Letters 24, 1641–1650. [https://doi.org/10.1016/S0167-8655\(03\)00003-5](https://doi.org/10.1016/S0167-8655(03)00003-5).
- Heiding, F., Süren, E., Olegård, J., Lagerström, R., 2023. Penetration testing of connected households. Computers & Security 126, 103067. <https://doi.org/10.1016/j.cose.2022.103067>.
- Himeur, Y., Alsalemi, A., Bensaali, F., Amira, A., 2021a. Smart power consumption abnormality detection in buildings using micromoments and improved K-nearest neighbors. Int J Intell Syst 36, 2865–2894. <https://doi.org/10.1002/int.22404>.
- Himeur, Y., Alsalemi, A., Bensaali, F., Amira, A., 2020. A Novel Approach for Detecting Anomalous Energy Consumption Based on Micro-Moments and Deep Neural Networks. Cogn Comput 12, 1381–1401. <https://doi.org/10.1007/s12559-020-09764-y>.
- Himeur, Y., Ghanem, K., Alsalemi, A., Bensaali, F., Amira, A., 2021b. Artificial intelligence based anomaly detection of energy consumption in buildings: A review, current trends and new perspectives. Applied Energy 287, 116601. <https://doi.org/10.1016/j.apenergy.2021.116601>.
- Hird, J.N., McDermid, G.J., 2009. Noise reduction of NDVI time series: An empirical comparison of selected techniques. Remote Sensing of Environment 113, 248–258. <https://doi.org/10.1016/j.rse.2008.09.003>.
- Hochreiter, S., Schmidhuber, J., 1997. Long Short-Term Memory. Neural Computation 9, 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>.
- Hoffman, D.L., Novak, T.P., 2015. Emergent Experience and the Connected Consumer in the Smart Home Assemblage and the Internet of Things. SSRN Journal. <https://doi.org/10.2139/ssrn.2648786>.
- Hong, T., Wang, Z., Luo, X., Zhang, W., 2020. State-of-the-art on research and applications of machine learning in the building life cycle. Energy and Buildings 212, 109831. <https://doi.org/10.1016/j.enbuild.2020.109831>.
- Huang, C., 2018. Featured Anomaly Detection Methods and Applications (Thèse de doctorat). University of Exeter. <http://hdl.handle.net/10871/34351>.
- Huchuk, B., Sanner, S., O'Brien, W., 2019. Comparison of machine learning models for occupancy prediction in residential buildings using connected thermostat data.

- Building and Environment 160, 106177.
<https://doi.org/10.1016/j.buildenv.2019.106177>.
- Jakkula, V., Cook, D., 2010. Outlier Detection in Smart Environment Structured Power Datasets, in: 2010 Sixth International Conference on Intelligent Environments. IEEE, Kuala Lumpur, Malaysia, pp. 29–33. <https://doi.org/10.1109/IE.2010.13>.
- Janda, K.B., 2011. Buildings don't use energy: people do. Architectural Science Review 54, 15–22. <https://doi.org/10.3763/asre.2009.0050>
- Javaid, M., Haleem, A., Rab, S., Pratap Singh, R., Suman, R., 2021. Sensors for daily life: A review. Sensors International 2, 100121. <https://doi.org/10.1016/j.sintl.2021.100121>.
- Jia, M., Komeily, A., Wang, Y., Srinivasan, R.S., 2019. Adopting Internet of Things for the development of smart buildings: A review of enabling technologies and applications. Automation in Construction 101, 111–126. <https://doi.org/10.1016/j.autcon.2019.01.023>.
- Johnson, J.M., Khoshgoftaar, T.M., 2019. Survey on deep learning with class imbalance. J Big Data 6, 27. <https://doi.org/10.1186/s40537-019-0192-5>.
- Josse, R., 2017. Méthode et outils pour l'identification de défauts des bâtiments connectés performants. Université de Grenoble Alpes. <https://theses.hal.science/tel-01743757>.
- Jozefowicz, R., Zaremba, W., Sutskever, I., 2015. An Empirical Exploration of Recurrent Network Architectures. International Conference on Machine Learning.
- Kaelbling, L.P., Littman, M.L., Moore, A.W., 1996. Reinforcement Learning: A Survey. jair 4, 237–285. <https://doi.org/10.1613/jair.301>.
- Kaiser, R., Maravall Herrero, A., 1999. Seasonal outliers in time series, Documento de trabajo / Banco de España, Servicio de Estudios. Banco de España, Madrid.
- Kalinichenko, L., Shanin, I., Taraban, I., 2014. Methods for anomaly detection: A survey. Proceedings of the All-Russian Conference Digital Libraries: Advanced Methods and Technologies.
- Karagiannis, V., Chatzimisios, P., Vazquez-Gallego, F., Alonso-Zarate, J., 2015. A Survey on Application Layer Protocols for the Internet of Things. Transaction on IoT and Cloud Computing, 1(1), 9-18.
- Karimi, R., Farahzadi, L., M.E. Sepasgozar, S., Sargolzaei, S., M. Ebrahimzadeh Sepasgozar, S., Zareian, M., Nasrolahi, A., 2021. Smart Built Environment Including Smart Home, Smart Building and Smart City: Definitions and Applied Technologies, in: Kaboli, A., Shirowzhan, S. (Eds.), Advances and Technologies in Building Construction and Structural Analysis. IntechOpen. <https://doi.org/10.5772/intechopen.95104>.
- Keogh, E., Lin, J., Fu, A.W., Van Herle, H., 2006. Finding Unusual Medical Time-Series Subsequences: Algorithms and Applications. IEEE Trans. Inform. Technol. Biomed. 10, 429–439. <https://doi.org/10.1109/TITB.2005.863870>.
- Ketkar, N., 2017. Deep Learning with Python. Apress, Berkeley, CA. <https://doi.org/10.1007/978-1-4842-2766-4>.
- Kieu, T., Yang, B., Guo, C., Jensen, C.S., 2019. Outlier Detection for Time Series with Recurrent Autoencoder Ensembles, in: Proceedings of the Twenty-Eighth

- International Joint Conference on Artificial Intelligence. International Joint Conferences on Artificial Intelligence Organization, Macao, China, pp. 2725–2732. <https://doi.org/10.24963/ijcai.2019/378>.
- Kim, H., Hong, T., Kim, J., 2019. Automatic ventilation control algorithm considering the indoor environmental quality factors and occupant ventilation behavior using a logistic regression model. *Building and Environment* 153, 46–59. <https://doi.org/10.1016/j.buildenv.2019.02.032>.
- Kim, S., Sung, Yujin, Sung, Yoondong, Seo, D., 2019. Development of a Consecutive Occupancy Estimation Framework for Improving the Energy Demand Prediction Performance of Building Energy Modeling Tools. *Energies* 12, 433. <https://doi.org/10.3390/en12030433>.
- Kim, Z.M., Oh, H., Kim, H.-G., Lim, C.-G., Oh, K.-J., Choi, H.-J., 2017. Modeling long-term human activeness using recurrent neural networks for biometric data. *BMC Med Inform Decis Mak* 17, 57. <https://doi.org/10.1186/s12911-017-0453-1>.
- Kingma, D.P., Ba, J., 2017. Adam: A Method for Stochastic Optimization. 3rd International Conference for Learning Representations, San Diego. <https://doi.org/10.48550/arXiv.1412.6980>.
- Kingma, D.P., Welling, M., 2013. Auto-Encoding Variational Bayes. <https://doi.org/10.48550/arXiv.1312.6114>.
- Koike-Akino, T., Wang, Y., 2020. Stochastic Bottleneck: Rateless Auto-Encoder for Flexible Dimensionality Reduction, in: 2020 IEEE International Symposium on Information Theory (ISIT). IEEE, Los Angeles, CA, USA, pp. 2735–2740. <https://doi.org/10.1109/ISIT44484.2020.9174523>.
- Kostelich, E.J., Yorke, J.A., 1990. Noise reduction: Finding the simplest dynamical system consistent with the data. *Physica D: Nonlinear Phenomena* 41, 183–196. [https://doi.org/10.1016/0167-2789\(90\)90121-5](https://doi.org/10.1016/0167-2789(90)90121-5).
- Kotsiantis, S.B., 2013. Decision trees: a recent overview. *Artif Intell Rev* 39, 261–283. <https://doi.org/10.1007/s10462-011-9272-4>.
- Kourou, K., Exarchos, T.P., Exarchos, K.P., Karamouzis, M.V., Fotiadis, D.I., 2015. Machine learning applications in cancer prognosis and prediction. *Computational and Structural Biotechnology Journal* 13, 8–17. <https://doi.org/10.1016/j.csbj.2014.11.005>.
- Kraiem, I.B., 2021. Détection d'Anomalies Multiples par Apprentissage Automatique de Règles dans les Séries Temporelles (Thèse de doctorat). Université de Toulouse-Jean Jaurès. <https://hal.science/tel-03137163v1>.
- Kriegel, H.-P., Kröger, P., Schubert, E., Zimek, A., 2009. LoOP: local outlier probabilities, in: Proceedings of the 18th ACM Conference on Information and Knowledge Management. ACM, Hong Kong China, pp. 1649–1652. <https://doi.org/10.1145/1645953.1646195>.
- Kullaa, J., 2013. Detection, identification, and quantification of sensor fault in a sensor network. *Mechanical Systems and Signal Processing* 40, 208–221. <https://doi.org/10.1016/j.ymssp.2013.05.007>.

- Laperre, B., Amaya, J., Lapenta, G., 2020. Dynamic Time Warping as a New Evaluation for Dst Forecast With Machine Learning. *Front. Astron. Space Sci.* 7, 39. <https://doi.org/10.3389/fspas.2020.00039>.
- Lawrynowicz, A., Tresp, V., 2014. Introducing Machine Learning. *Introduction to Machine Learning*. In *Perspectives on Ontology Learning*; Microsoft Press: Redmond, WA, USA, 2014; pp. 35–50.
- Li, J., Malialis, K., Polycarpou, M.M., 2023. Autoencoder-based Anomaly Detection in Streaming Data with Incremental Learning and Concept Drift Adaptation, in: 2023 International Joint Conference on Neural Networks (IJCNN). pp. 1–8. <https://doi.org/10.1109/IJCNN54540.2023.10191328>.
- Li, S., Li, F., Tang, S., Xiong, W., 2020. A Review of Computer-Aided Heart Sound Detection Techniques. *BioMed Research International* 2020, 1–10. <https://doi.org/10.1155/2020/5846191>.
- Li, S., Wen, J., 2014. A model-based fault detection and diagnostic methodology based on PCA method and wavelet transform. *Energy and Buildings* 68, 63–71. <https://doi.org/10.1016/j.enbuild.2013.08.044>.
- Li, T., Zhu, S., Ogiwara, M., 2006. Using discriminant analysis for multi-class classification: an experimental investigation. *Knowl Inf Syst* 10, 453–472. <https://doi.org/10.1007/s10115-006-0013-y>.
- Li, Y., Lei, Y., Wang, P., Jiang, M., Liu, Y., 2021. Embedded stacked group sparse autoencoder ensemble with L1 regularization and manifold reduction. *Applied Soft Computing* 101, 107003. <https://doi.org/10.1016/j.asoc.2020.107003>.
- Linder, L., Vionnet, D., Bacher, J.-P., Hennebert, J., 2017. Big Building Data - a Big Data Platform for Smart Buildings. *Energy Procedia* 122, 589–594. <https://doi.org/10.1016/j.egypro.2017.07.354>.
- Liu, F.T., Ting, K.M., Zhou, Z.-H., 2008. Isolation Forest, in: 2008 Eighth IEEE International Conference on Data Mining. IEEE, Pisa, Italy, pp. 413–422. <https://doi.org/10.1109/ICDM.2008.17>.
- Liu, J., Yang, D., Zhang, K., Gao, H., Li, J., 2023. Anomaly and change point detection for time series with concept drift. *World Wide Web*. <https://doi.org/10.1007/s11280-023-01181-z>.
- Liu, Y., Pang, Z., Karlsson, M., Gong, S., 2020. Anomaly detection based on machine learning in IoT-based vertical plant wall for indoor climate control. *Building and Environment* 183, 107212. <https://doi.org/10.1016/j.buildenv.2020.107212>.
- Lopes, M.A.R., Antunes, C.H., Reis, A., Martins, N., 2017. Estimating energy savings from behaviours using building performance simulations. *Building Research & Information* 45, 303–319. <https://doi.org/10.1080/09613218.2016.1140000>.
- Luczak, M., 2017. Univariate and multivariate time series classification with parametric integral dynamic time warping. *IFS* 33, 2403–2413. <https://doi.org/10.3233/JIFS-17523>.
- Lundervold, A.S., Lundervold, A., 2019. An overview of deep learning in medical imaging focusing on MRI. *Zeitschrift für Medizinische Physik* 29, 102–127. <https://doi.org/10.1016/j.zemedi.2018.11.002>.

- Madhulatha, T.S., 2012. An overview of clustering methods. IOSRJEN 02, 719–725. <https://doi.org/10.9790/3021-0204719725>.
- Mahmoud, S., Lotfi, A., Langensiepen, C., 2016. User Activities Outliers Detection; Integration of Statistical and Computational Intelligence Techniques: USER ACTIVITIES OUTLIERS DETECTION. Computational Intelligence 32, 49–71. <https://doi.org/10.1111/coin.12045>.
- Malhotra, P., Vig, L., Shroff, G., Agarwal, P., 2015. Long Short Term Memory Networks for Anomaly Detection in Time Series. The European Symposium on Artificial Neural Networks.
- Mariano-Hernández, D., Hernández-Callejo, L., Zorita-Lamadrid, A., Duque-Pérez, O., Santos García, F., 2021. A review of strategies for building energy management system: Model predictive control, demand side management, optimization, and fault detect & diagnosis. Journal of Building Engineering 33, 101692. <https://doi.org/10.1016/j.jobbe.2020.101692>.
- Marikyan, D., Papagiannidis, S., Alamanos, E., 2019. A systematic review of the smart home literature: A user perspective. Technological Forecasting and Social Change 138, 139–154. <https://doi.org/10.1016/j.techfore.2018.08.015>.
- Markou, M., Singh, S., 2003. Novelty detection: a review—part 1: statistical approaches. Signal Processing 83, 2481–2497. <https://doi.org/10.1016/j.sigpro.2003.07.018>.
- Markovic, R., 2018. Window opening model using deep learning methods. Building and Environment 11. <https://doi.org/10.48550/arXiv.1807.03610>.
- Markovic, R., Frisch, J., van Treeck, C., 2019. Learning short-term past as predictor of window opening-related human behavior in commercial buildings. Energy and Buildings 185, 1–11. <https://doi.org/10.1016/j.enbuild.2018.12.012>.
- Markovic, R., Wolf, S., Cao, J., Spinnraker, E., Wölki, D., Frisch, J., van Treeck, C., 2017. Comparison of Different Classification Algorithms for the Detection of User's Interaction with Windows in Office Buildings. Energy Procedia 122, 337–342. <https://doi.org/10.1016/j.egypro.2017.07.333>.
- Martí, L., Sanchez-Pi, N., Molina, J., Garcia, A., 2015. Anomaly Detection Based on Sensor Data in Petroleum Industry Applications. Sensors 15, 2774–2797. <https://doi.org/10.3390/s150202774>.
- Martin Nascimento, G.F., Wurtz, F., Kuo-Peng, P., Delinchant, B., Jhoé Batistela, N., 2021. Outlier Detection in Buildings' Power Consumption Data Using Forecast Error. Energies 14, 8325. <https://doi.org/10.3390/en14248325>.
- McFarlane, D., Sarma, S., Chirn, J.L., Wong, C.Y., Ashton, K., 2003. Auto ID systems and intelligent manufacturing control. Engineering Applications of Artificial Intelligence 16, 365–376. [https://doi.org/10.1016/S0952-1976\(03\)00077-0](https://doi.org/10.1016/S0952-1976(03)00077-0).
- Mehrotra, K.G., Mohan, C.K., Huang, H., 2017. Anomaly Detection Principles and Algorithms, Terrorism, Security, and Computation. Springer International Publishing, Cham. <https://doi.org/10.1007/978-3-319-67526-8>.
- Miljković, D., 2020. Review of Novelty Detection Methods. The 33rd International Convention MIPRO : 593-598.

- Mishra, A., Cohen, A., Reichherzer, T., Wilde, N., 2021. Detection of data anomalies at the edge of pervasive IoT systems. *Computing* 103, 1657–1675. <https://doi.org/10.1007/s00607-021-00927-9>.
- Moody, G.B., Mark, R.G., 2001. The impact of the MIT-BIH Arrhythmia Database. *IEEE Eng. Med. Biol. Mag.* 20, 45–50. <https://doi.org/10.1109/51.932724>.
- Moreno, M., Úbeda, B., Skarmeta, A., Zamora, M., 2014. How can We Tackle Energy Efficiency in IoT Based Smart Buildings? *Sensors* 14, 9582–9614. <https://doi.org/10.3390/s140609582>.
- Mori, U., Mendiburu, A., Lozano, J., A., 2016. Distance Measures for Time Series in R: The TSdist Package. *The R Journal* 8, 451. <https://doi.org/10.32614/RJ-2016-058>.
- Moudgil, V., Hewage, K., Hussain, S.A., Sadiq, R., 2023. Integration of IoT in building energy infrastructure: A critical review on challenges and solutions. *Renewable and Sustainable Energy Reviews* 174, 113121. <https://doi.org/10.1016/j.rser.2022.113121>.
- Mulongo, J., Atemkeng, M., Ansah-Narh, T., Rockefeller, R., Nguegnang, G.M., Garuti, M.A., 2020. Anomaly Detection in Power Generation Plants Using Machine Learning and Neural Networks. *Applied Artificial Intelligence* 34, 64–79. <https://doi.org/10.1080/08839514.2019.1691839>.
- Musall, E., Weiss, T., Voss, K., Lenoir, A., Donn, M., Cory, S., Garde, F., 2010. Net Zero Energy Solar Buildings: An Overview and Analysis on Worldwide Building Projects, in: *Proceedings of the EuroSun 2010 Conference*. International Solar Energy Society, Graz, Austria, pp. 1–9. <https://doi.org/10.18086/eurosun.2010.06.16>.
- Nabil, T., 2018. Building identification within a connected object ecosystem (Thèse de doctorat). Télécom ParisTech. <https://hal.science/tel-02983222>.
- Nandan Mohanty, S., Chatterjee, J.M., Satpathy, S. (Eds.), 2022. *Internet of Things and Its Applications*, EAI/Springer Innovations in Communication and Computing. Springer International Publishing, Cham. <https://doi.org/10.1007/978-3-030-77528-5>.
- Naspi, F., Arnesano, M., Zampetti, L., Stazi, F., Revel, G.M., D’Orazio, M., 2018. Experimental study on occupants’ interaction with windows and lights in Mediterranean offices during the non-heating season. *Building and Environment* 127, 221–238. <https://doi.org/10.1016/j.buildenv.2017.11.009>.
- Nazareth, N., Ramana Reddy, Y.V., 2023. Financial applications of machine learning: A literature review. *Expert Systems with Applications* 219, 119640. <https://doi.org/10.1016/j.eswa.2023.119640>.
- Ng, A., 2010. Sparse autoencoder. *Lecture Notes: Sparse Autoencoder*, Stanford, CA, USA: Stanford Univ. Press.
- Ni, K., Ramanathan, N., Chehade, M.N.H., Balzano, L., Nair, S., Zahedi, S., Kohler, E., Pottie, G., Hansen, M., Srivastava, M., 2009. Sensor network data fault types. *ACM Trans. Sen. Netw.* 5, 1–29. <https://doi.org/10.1145/1525856.1525863>.
- Nihtianov, S., Luque, A., 2018. *Smart Sensors and MEMs*, second edition. ed. Elsevier/Woodhead publishing.

- Norlander, E., Sopasakis, A., 2019. Latent space conditioning for improved classification and anomaly detection. <https://doi.org/10.48550/arXiv.1911.10599>.
- Noshad, Z., Javaid, N., Saba, T., Wadud, Z., Saleem, M., Alzahrani, M., Sheta, O., 2019. Fault Detection in Wireless Sensor Networks through the Random Forest Classifier. *Sensors* 19, 1568. <https://doi.org/10.3390/s19071568>.
- Öller, L.-E., 1985. Macroeconomic forecasting with a Vector ARIMA model : A case study of the finnish economy. *International Journal of Forecasting*, Volume 1, Issue 2, 1985, Pages 143-150, ISSN 0169-2070. [https://doi.org/10.1016/0169-2070\(85\)90019-6](https://doi.org/10.1016/0169-2070(85)90019-6).
- O'Reilly, C., Gluhak, A., Imran, M.A., Rajasegarar, S., 2014. Anomaly Detection in Wireless Sensor Networks in a Non-Stationary Environment. *IEEE Commun. Surv. Tutorials* 16, 1413–1432. <https://doi.org/10.1109/SURV.2013.112813.00168>.
- Padmanabhan, J., Johnson Premkumar, M.J., 2015. Machine Learning in Automatic Speech Recognition: A Survey. *IETE Technical Review* 32, 240–251. <https://doi.org/10.1080/02564602.2015.1010611>.
- Panjei, E., Gruenwald, L., Leal, E., Nguyen, C., Silvia, S., 2022. A survey on outlier explanations. *The VLDB Journal* 31, 977–1008. <https://doi.org/10.1007/s00778-021-00721-1>.
- Park, H., Rhee, S.-B., 2018. IoT-Based Smart Building Environment Service for Occupants' Thermal Comfort. *Journal of Sensors* 2018, 1–10. <https://doi.org/10.1155/2018/1757409>.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., 2011. Scikit-learn: Machine Learning in Python. *MACHINE LEARNING IN PYTHON*.
- Peña, M., Biscarri, F., Guerrero, J.I., Monedero, I., León, C., 2016. Rule-based system to detect energy efficiency anomalies in smart buildings, a data mining approach. *Expert Systems with Applications* 56, 242–255. <https://doi.org/10.1016/j.eswa.2016.03.002>.
- Pereira, J., Silveira, M., 2018. Unsupervised Anomaly Detection in Energy Time Series Data Using Variational Recurrent Autoencoders with Attention, in: 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA). IEEE, Orlando, FL, pp. 1275–1282. <https://doi.org/10.1109/ICMLA.2018.00207>.
- Pereira, M., Glisic, B., 2023. Detection and quantification of temperature sensor drift using probabilistic neural networks. *Expert Systems with Applications* 213, 118884. <https://doi.org/10.1016/j.eswa.2022.118884>.
- Petroșanu, D.-M., Căruțașu, G., Căruțașu, N.L., Pîrjan, A., 2019. A Review of the Recent Developments in Integrating Machine Learning Models with Sensor Devices in the Smart Buildings Sector with a View to Attaining Enhanced Sensing, Energy Efficiency, and Optimal Building Management. *Energies* 12, 4745. <https://doi.org/10.3390/en12244745>.
- Pham, H.V., Qian, S., Wang, J., Lutellier, T., Rosenthal, J., Tan, L., Yu, Y., Nagappan, N., 2020. Problems and Opportunities in Training Deep Learning Software Systems: An Analysis of Variance. 35th IEEE/ACM International Conference on Automated Software Engineering (ASE) : 771-783.

- Phan, L.H., 2022. Modélisation de l'activité des ménages dans le bâtiment résidentiel instrumenté (Thèse de doctorat). <https://theses.hal.science/tel-03796541>.
- Phoha, V.V., 2002. Internet security dictionary. Springer, New York.
- Prabhakaran, K., Dridi, J., Amayri, M., Bouguila, N., 2022. Explainable K-Means Clustering for Occupancy Estimation. *Procedia Computer Science* 203, 326–333. <https://doi.org/10.1016/j.procs.2022.07.041>.
- Prechelt, L., 2012. Early Stopping — But When?. In: Montavon, G., Orr, G.B., Müller, KR. (eds) *Neural Networks: Tricks of the Trade. Lecture Notes in Computer Science*, vol 7700. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-35289-8_5.
- Provost, F., Fawcett, T., 1999. Analysis and Visualization of Classifier Performance: Comparison Under Imprecise Class and Cost Distributions. *Proceedings of the third international conference on knowledge discovery and data mining*.
- Purarjomandlangrudi, A., Ghapanchi, A.H., Esmalifalak, M., 2014. A data mining approach for fault diagnosis: An application of anomaly detection algorithm. *Measurement* 55, 343–352. <https://doi.org/10.1016/j.measurement.2014.05.029>.
- Qiang, G., Tang, S., Hao, J., Di Sarno, L., Wu, G., Ren, S., 2023. Building automation systems for energy and comfort management in green buildings: A critical review and future directions. *Renewable and Sustainable Energy Reviews* 179, 113301. <https://doi.org/10.1016/j.rser.2023.113301>.
- Ramaswamy, S., Rastogi, R., Shim, K., 2000. Efficient Algorithms for Mining Outliers from Large Data Sets. *Association for Computing Machinery, Association for Computing Machinery*. <https://doi.org/10.1145/342009.335437>.
- Ramesh, T., Prakash, R., Shukla, K.K., 2010. Life cycle energy analysis of buildings: An overview. *Energy and Buildings* 42, 1592–1600. <https://doi.org/10.1016/j.enbuild.2010.05.007>.
- Razavi, R., Gharipour, A., Fleury, M., Akpan, I.J., 2019. Occupancy detection of residential buildings using smart meter data: A large-scale study. *Energy and Buildings* 183, 195–208. <https://doi.org/10.1016/j.enbuild.2018.11.025>.
- Recht, T., 2016. Étude de l'écoconception de maisons à énergie positive (Thèse de doctorat). Paris Sciences et Lettres.
- Recht, T., Goffart, J., Mora, L., Woloszyn, M., Buhé, C., 2018. Méthodologie pour la comparaison des performances simulées et mesurées de maisons « à énergie positive ». Bordeaux, France, p. 8.
- Rezende, D.J., Mohamed, S., Wierstra, D., 2014. Stochastic Backpropagation and Approximate Inference in Deep Generative Models. *Proceedings of the 31st International Conference on Machine Learning*, , in *Proceedings of Machine Learning Research* 32(2) : 1278-1286 Available from <https://proceedings.mlr.press/v32/rezende14.html>.
- Rijal, H.B., Humphreys, M.A., Nicol, J.F., 2018. Development of a window opening algorithm based on adaptive thermal comfort to predict occupant behavior in Japanese dwellings. *Japan Architectural Review* 1, 310–321. <https://doi.org/10.1002/2475-8876.12043>.

- Rijal, H.B., Tuohy, P., Humphreys, M.A., Nicol, J.F., Samuel, A., Clarke, J., 2007. Using results from field surveys to predict the effect of open windows on thermal comfort and energy use in buildings. *Energy and Buildings* 39, 823–836. <https://doi.org/10.1016/j.enbuild.2007.02.003>.
- Rose, K., Eldridge, S., Chapin, L., 2015. The Internet of Things: An Overview understanding the Issues and challenges of a more connected world”, Internet Society(ISOC).
- Rumelhart, D.E., Hintont, G.E., Williams, R.J., 1986. Learning representations by back-propagating errors. *Nature*, 323, 533-536.
- Sakoe, H., Chiba, S., 1978. Dynamic programming algorithm optimization for spoken word recognition. *IEEE Trans. Acoust., Speech, Signal Process.* 26, 43–49. <https://doi.org/10.1109/TASSP.1978.1163055>.
- Sanjay Kumar Ahuja, Kiran Kumar Ravulakollu, Manoj Kumar Shukla, 2019. Surface Corrosion Grade Classification using Convolution Neural Network. *IJRTE* 8, 7645–7649. <https://doi.org/10.35940/ijrte.C6196.098319>.
- Savage, D., Zhang, X., Yu, X., Chou, P., Wang, Q., 2014. Anomaly detection in online social networks. *Social Networks* 39, 62–70. <https://doi.org/10.1016/j.socnet.2014.05.002>.
- Scanu, L., 2017. Towards archetypes of self-tuned models for connected buildings (Thèse de doctorat). Université Grenoble Alpe. <https://theses.hal.science/tel-01736999>.
- Scanu, L., Ploix, S., Bernaud, P., Wurtz, E., 2017. Model Tuning Approach for Energy Management of Office and Apartment Settings. <https://doi.org/10.26868/25222708.2017.233>.
- Schmidhuber, J., 2015. Deep Learning in Neural Networks: An Overview. *Neural Networks* 61, 85–117. <https://doi.org/10.1016/j.neunet.2014.09.003>.
- Schmidl, S., Wenig, P., Papenbrock, T., 2022. Anomaly detection in time series: a comprehensive evaluation. *Proc. VLDB Endow.* 15, 1779–1797. <https://doi.org/10.14778/3538598.3538602>.
- Schneider, M., Ertel, W., Ramos, F., 2016. Expected similarity estimation for large-scale batch and streaming anomaly detection. *Mach Learn* 105, 305–333. <https://doi.org/10.1007/s10994-016-5567-7>.
- Seliya, N., Khoshgoftaar, T.M., Van Hulse, J., 2009. A Study on the Relationships of Classifier Performance Metrics. 2009 21st IEEE International Conference on Tools with Artificial Intelligence 59–66. <https://doi.org/10.1109/ICTAI.2009.25>.
- Senin, P., 2008. Dynamic Time Warping Algorithm Review. *Inf. Comput. Sci. Department University of Hawaii at Manoa Honolulu, USA*, vol. 855, no. 1-23, p. 40.
- Sethi, P., Sarangi, S.R., 2017. Internet of Things: Architectures, Protocols, and Applications. *Journal of Electrical and Computer Engineering* 2017, 1–25. <https://doi.org/10.1155/2017/9324035>.
- Shakya, A.K., Pillai, G., Chakrabarty, S., 2023. Reinforcement learning algorithms: A brief survey. *Expert Systems with Applications* 231, 120495. <https://doi.org/10.1016/j.eswa.2023.120495>.

- Sharma, A.B., Golubchik, L., Govindan, R., 2010. Sensor faults: Detection methods and prevalence in real-world datasets. *ACM Trans. Sen. Netw.* 6, 1–39. <https://doi.org/10.1145/1754414.1754419>.
- Shcherbakov, M.V., Brebels, A., Shcherbakova, N.L., Tyukov, A.P., Janovsky, T.A., Kamaev, V.A., 2013. A Survey of Forecast Error Measures. *World Applied Sciences Journal (Information Technologies in Modern Industry, Education & Society)* 24.
- Shi, Z., Qian, H., Zheng, X., Lv, Z., Li, Y., Liu, L., Nielsen, P.V., 2018. Seasonal variation of window opening behaviors in two naturally ventilated hospital wards. *Building and Environment* 130, 85–93. <https://doi.org/10.1016/j.buildenv.2017.12.019>.
- Siami-Namini, S., Tavakoli, N., Siami Namin, A., 2018. A Comparison of ARIMA and LSTM in Forecasting Time Series, in: 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA). IEEE, Orlando, FL, pp. 1394–1401. <https://doi.org/10.1109/ICMLA.2018.00227>.
- Sinopoli, J., 2010. *Smart Building Systems for Architects, Owners and Builders*. Elsevier. <https://doi.org/10.1016/C2009-0-20023-7>.
- Skrobek, D., Krzywanski, J., Sosnowski, M., Kulakowska, A., Zylka, A., Grabowska, K., Ciesielska, K., Nowak, W., 2020. Prediction of Sorption Processes Using the Deep Learning Methods (Long Short-Term Memory). *Energies* 13, 6601. <https://doi.org/10.3390/en13246601>.
- Škvára, V., Pevný, T., Šmídl, V., 2018. Are generative deep models for novelty detection truly better? <https://doi.org/10.48550/arXiv.1807.05027>.
- Soffer, S., Klang, E., Shimon, O., Nachmias, N., Eliakim, R., Ben-Horin, S., Kopylov, U., Barash, Y., 2020. Deep learning for wireless capsule endoscopy: a systematic review and meta-analysis. *Gastrointestinal Endoscopy* 92, 831–839.e8. <https://doi.org/10.1016/j.gie.2020.04.039>.
- Song, H., Jiang, Z., Men, A., Yang, B., 2017. A Hybrid Semi-Supervised Anomaly Detection Model for High-Dimensional Data. *Computational Intelligence and Neuroscience* 2017, 1–9. <https://doi.org/10.1155/2017/8501683>.
- Sovacool, B.K., Furszyfer Del Rio, D.D., 2020. Smart home technologies in Europe: A critical review of concepts, benefits, risks and policies. *Renewable and Sustainable Energy Reviews* 120, 109663. <https://doi.org/10.1016/j.rser.2019.109663>.
- Soydaner, D., 2020. A Comparison of Optimization Algorithms for Deep Learning. *Int. J. Patt. Recogn. Artif. Intell.* 34, 2052013. <https://doi.org/10.1142/S0218001420520138>.
- Stamatescu, G., Stamatescu, I., Arghira, N., Fagarasan, I., 2019. Data-Driven Modelling of Smart Building Ventilation Subsystem. *Journal of Sensors* 2019, 1–14. <https://doi.org/10.1155/2019/3572019>.
- Stazi, F., Naspi, F., D'Orazio, M., 2017. Modelling window status in school classrooms. Results from a case study in Italy. *Building and Environment* 111, 24–32. <https://doi.org/10.1016/j.buildenv.2016.10.013>.
- Su, S., Sun, Y., Gao, X., Qiu, J., Tian, Z., 2019. A Correlation-Change Based Feature Selection Method for IoT Equipment Anomaly Detection. *Applied Sciences* 9, 437. <https://doi.org/10.3390/app9030437>.

- Taha, A., Hadi, A.S., 2020. Anomaly Detection Methods for Categorical Data: A Review. *ACM Comput. Surv.* 52, 1–35. <https://doi.org/10.1145/3312739>.
- Tan, P.-N., Steinbach, M., Kumar, V., 2014. Introduction to data mining, New internat. edition. ed, Always learning. Pearson, Harlow.
- Tang, J., Chen, Z., Wai-chee Fu, A., Cheung, D.W., 2002. Enhancing Effectiveness of Outlier Detections for Low Density Patterns. *Proceedings of the Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining (PAKDD)*, *Proceedings of the Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining (PAKDD)*. https://doi.org/doi: 10.1007/3-54047887-6_53.
- Tang, S., Shelden, D.R., Eastman, C.M., Pishdad-Bozorgi, P., Gao, X., 2019. A review of building information modeling (BIM) and the internet of things (IoT) devices integration: Present status and future trends. *Automation in Construction* 101, 127–139. <https://doi.org/10.1016/j.autcon.2019.01.020>.
- Tharwat, A., 2021. Classification assessment methods. *ACI* 17, 168–192. <https://doi.org/10.1016/j.aci.2018.08.003>.
- Theodoridis, S., Koutroumbas, K., 2008. Pattern Recognition Ed. 4. Elsevier Science.
- Thibault, É., Désilets, F.L., Poulin, B., Chioua, M., Stuart, P., 2023. Comparison of signal processing methods considering their optimal parameters using synthetic signals in a heat exchanger network simulation. *Computers & Chemical Engineering* 178, 108380. <https://doi.org/10.1016/j.compchemeng.2023.108380>.
- Tomat, V., Ramallo-González, A.P., Skarmeta Gómez, A.F., 2020. A Comprehensive Survey about Thermal Comfort under the IoT Paradigm: Is Crowdsensing the New Horizon? *Sensors* 20, 4647. <https://doi.org/10.3390/s20164647>.
- Tran, L., Fan, L., Shahabi, C., 2019. Outlier Detection in Non-stationary Data Streams, in: *Proceedings of the 31st International Conference on Scientific and Statistical Database Management*. ACM, Santa Cruz CA USA, pp. 25–36. <https://doi.org/10.1145/3335783.3335788>.
- Trappey, A.J.C., Trappey, C.V., Hareesh Govindarajan, U., Chuang, A.C., Sun, J.J., 2017. A review of essential standards and patent landscapes for the Internet of Things: A key enabler for Industry 4.0. *Advanced Engineering Informatics* 33, 208–229. <https://doi.org/10.1016/j.aei.2016.11.007>.
- Vandewalle, N., Ausloos, M., Boveroux, P., 1999. The moving averages demystified. *Physica A: Statistical Mechanics and its Applications*, Elsevier, vol. 269(1), pages 170-176. <https://ideas.repec.org/a/eee/phsmap/v269y1999i1p170-176.html>.
- Vellei, M., Natarajan, S., Biri, B., Padget, J., Walker, I., 2016. The effect of real-time context-aware feedback on occupants' heating behaviour and thermal adaptation. *Energy and Buildings* 123, 179–191. <https://doi.org/10.1016/j.enbuild.2016.03.045>.
- Verma, A., Prakash, S., Srivastava, V., Kumar, A., Mukhopadhyay, S.C., 2019. Sensing, Controlling, and IoT Infrastructure in Smart Building: A Review. *IEEE Sensors J.* 19, 9036–9046. <https://doi.org/10.1109/JSEN.2019.2922409>.
- Vincent, P., Larochelle, H., Bengio, Y., Manzagol, P.-A., 2008. Extracting and composing robust features with denoising autoencoders, in: *Proceedings of the 25th International Conference on Machine Learning - ICML '08*. ACM Press, Helsinki, Finland, pp. 1096–1103. <https://doi.org/10.1145/1390156.1390294>.

- Wan Yen, S., Morris, S., Ezra, M.A.G., Jun Huat, T., 2019. Effect of smart meter data collection frequency in an early detection of shorter-duration voltage anomalies in smart grids. *International Journal of Electrical Power & Energy Systems* 109, 1–8. <https://doi.org/10.1016/j.ijepes.2019.01.039>.
- Wang, Y., Masoud, N., Khojandi, A., 2021. Real-Time Sensor Anomaly Detection and Recovery in Connected Automated Vehicle Sensors. *IEEE Trans. Intell. Transport. Syst.* 22, 1411–1421. <https://doi.org/10.1109/TITS.2020.2970295>.
- Wei, Y., Jang-Jaccard, J., Xu, W., Sabrina, F., Camtepe, S., Boulic, M., 2022. LSTM-Autoencoder based Anomaly Detection for Indoor Air Quality Time Series Data.
- Wei, Y., Yu, H., Pan, S., Xia, L., Xie, J., Wang, X., Wu, J., Zhang, W., Li, Q., 2019. Comparison of different window behavior modeling approaches during transition season in Beijing, China. *Building and Environment* 157, 1–15. <https://doi.org/10.1016/j.buildenv.2019.04.040>.
- Wilson, C., Hargreaves, T., Hauxwell-Baldwin, R., 2017. Benefits and risks of smart home technologies. *Energy Policy* 103, 72–83. <https://doi.org/10.1016/j.enpol.2016.12.047>.
- Wu, R., Keogh, E., 2021. Current Time Series Anomaly Detection Benchmarks are Flawed and are Creating the Illusion of Progress. *IEEE Trans. Knowl. Data Eng.* 1–1. <https://doi.org/10.1109/TKDE.2021.3112126>.
- Wythoff, B.J., 1993. Backpropagation neural networks A tutorial. *Chemom. Intell. Lab. Syst. Chemometrics and Intelligent Laboratory Systems*, Volume 18, Issue 2, Pages 115–155, ISSN 0169-7439. [https://doi.org/10.1016/0169-7439\(93\)80052-J](https://doi.org/10.1016/0169-7439(93)80052-J).
- Xu, D., Tian, Y., 2015. A Comprehensive Survey of Clustering Algorithms. *Ann. Data. Sci.* 2, 165–193. <https://doi.org/10.1007/s40745-015-0040-1>.
- Xu, H., Pang, G., Wang, Yijie, Wang, Yongjun, 2023. Deep Isolation Forest for Anomaly Detection. *IEEE Trans. Knowl. Data Eng.* 1–14. <https://doi.org/10.1109/TKDE.2023.3270293>.
- Yang, R., Wang, L., 2013. Development of multi-agent system for building energy and comfort management based on occupant behaviors. *Energy and Buildings* 56, 1–7. <https://doi.org/10.1016/j.enbuild.2012.10.025>.
- Yang, Y., Sautière, G., Ryu, J.J., Cohen, T.S., 2020. Feedback Recurrent AutoEncoder. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (2019): 3347–3351. <https://doi.org/10.48550/arXiv.1911.04018>.
- Yao, Y., Sharma, A., Golubchik, L., Govindan, R., 2010. Online anomaly detection for sensor systems: A simple and efficient approach. *Performance Evaluation* 67, 1059–1075. <https://doi.org/10.1016/j.peva.2010.08.018>.
- Ye, Y., Li, T., Adjero, D., Iyengar, S.S., 2018. A Survey on Malware Detection Using Data Mining Techniques. *ACM Comput. Surv.* 50, 1–40. <https://doi.org/10.1145/3073559>.
- Yu, Z., Fung, B.C.M., Haghighat, F., 2013. Extracting knowledge from building-related data — A data mining framework. *Build. Simul.* 6, 207–222. <https://doi.org/10.1007/s12273-013-0117-8>.

- Zhang, G., Li, Y., Deng, X., 2020. K-Means Clustering-Based Electrical Equipment Identification for Smart Building Application. *Information* 11, 27. <https://doi.org/10.3390/info11010027>.
- Zhang, J., 2013. Advancements of Outlier Detection: A Survey. *ICST Trans ScaInfoSys* 13, e2. <https://doi.org/10.4108/trans.sis.2013.01-03.e2>.
- Zhang, Z., Chen, Y., Wang, H., Fu, Q., Chen, J., Lu, Y., 2023. Anomaly detection method for building energy consumption in multivariate time series based on graph attention mechanism. *PLoS ONE* 18, e0286770. <https://doi.org/10.1371/journal.pone.0286770>.
- Zhou, X., Cheng, S., Zhu, M., Guo, C., Zhou, S., Xu, P., Xue, Z., Zhang, W., 2018. A state of the art survey of data mining-based fraud detection and credit scoring. *MATEC Web Conf.* 189, 03002. <https://doi.org/10.1051/mateconf/201818903002>.
- Zimek, A., Filzmoser, P., 2018. There and back again: Outlier detection between statistical reasoning and data mining algorithms. *WIREs Data Mining Knowl Discov* 8. <https://doi.org/10.1002/widm.1280>.
- Zou, H., Zhou, Y., Yang, J., Spanos, C.J., 2018. Towards occupant activity driven smart buildings via WiFi-enabled IoT devices and deep learning. *Energy and Buildings* 177, 12–22. <https://doi.org/10.1016/j.enbuild.2018.08.010>.

Annexes

Annexe A. Le cas des mesures d'humidité relative d'air intérieur

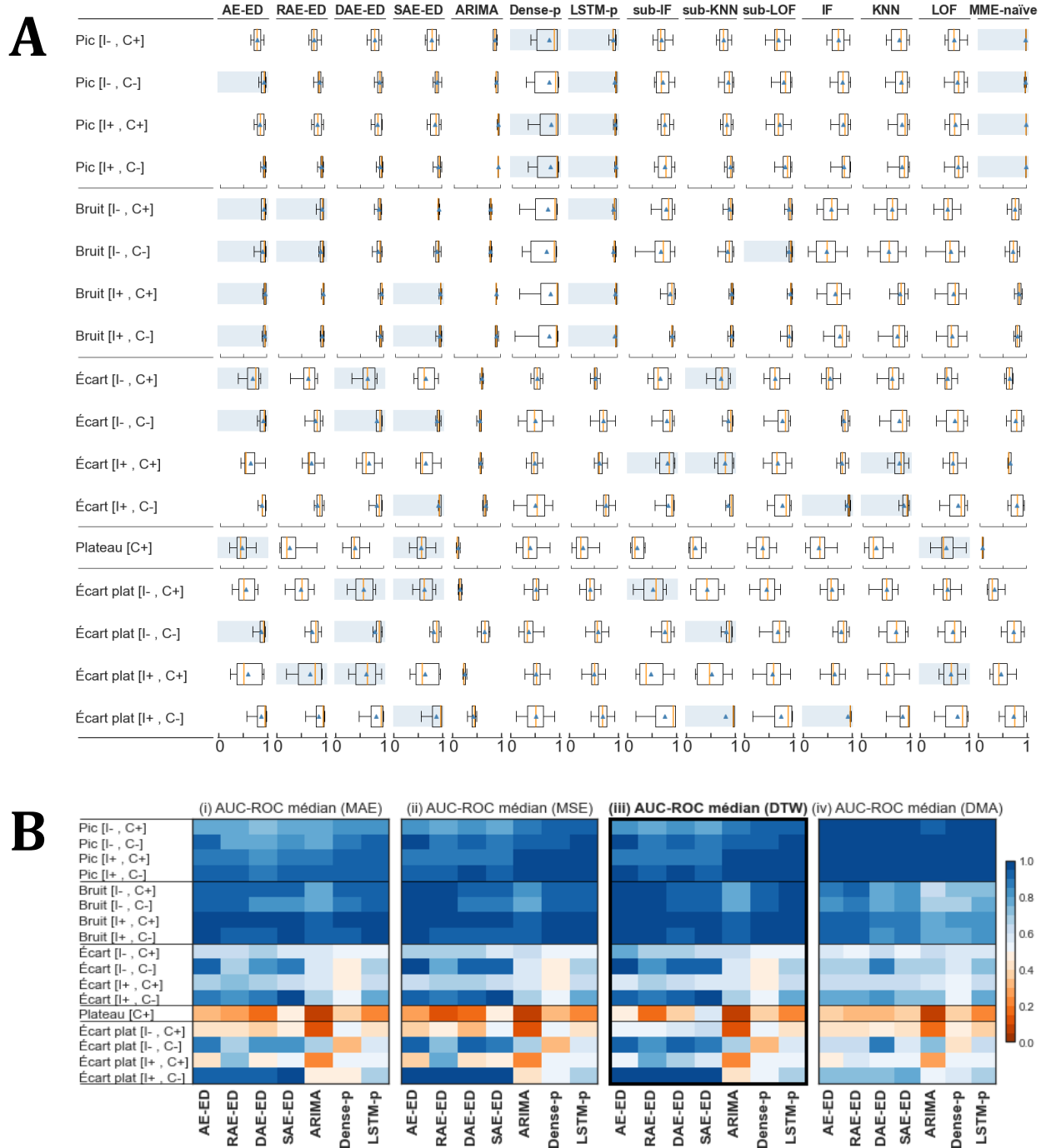


Figure 71 (A) Résultats AUC-ROC pour les mesures d'humidité relative d'air intérieur. Les intensités (I) et les taux de contamination (C) sont associés aux deux échelles d'intensité et de contamination, faibles (-) et fortes (+). La médiane des boîtes à moustaches est représentée par une ligne pleine orange, tandis que la moyenne est indiquée par un triangle bleu. (B) Scores d'anormalité des méthodes de reconstruction et de prédiction, calculés en employant (i) la MSE, (ii) la MAE, (iii) la DTW et (iv) la DMA.

Annexe B. Le cas des mesures de concentration de CO₂ intérieure

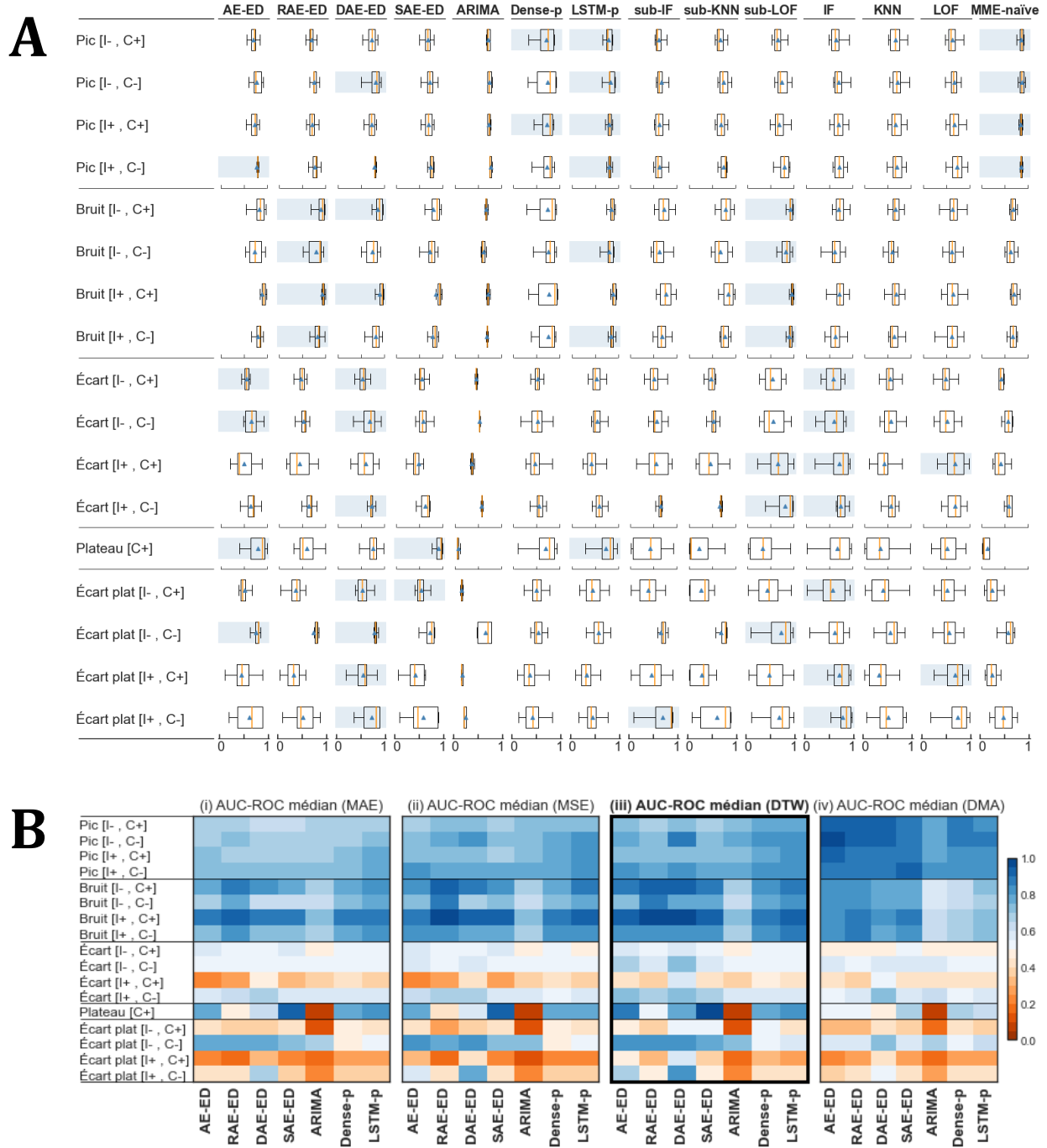


Figure 72 (A) Résultats AUC-ROC pour les mesures de concentration de CO₂ intérieure. Les intensités (I) et les taux de contamination (C) sont associés aux deux échelles d'intensité et de contamination, faibles (-) et fortes (+). La médiane des boîtes à moustaches est représentée par une ligne pleine orange, tandis que la moyenne est indiquée par un triangle bleu. (B) Scores d'anormalité des méthodes de reconstruction et de prédiction, calculés en employant (i) la MSE, (ii) la MAE, (iii) la DTW et (iv) la DMA.

Annexe C. Le cas des mesures de consommation électrique

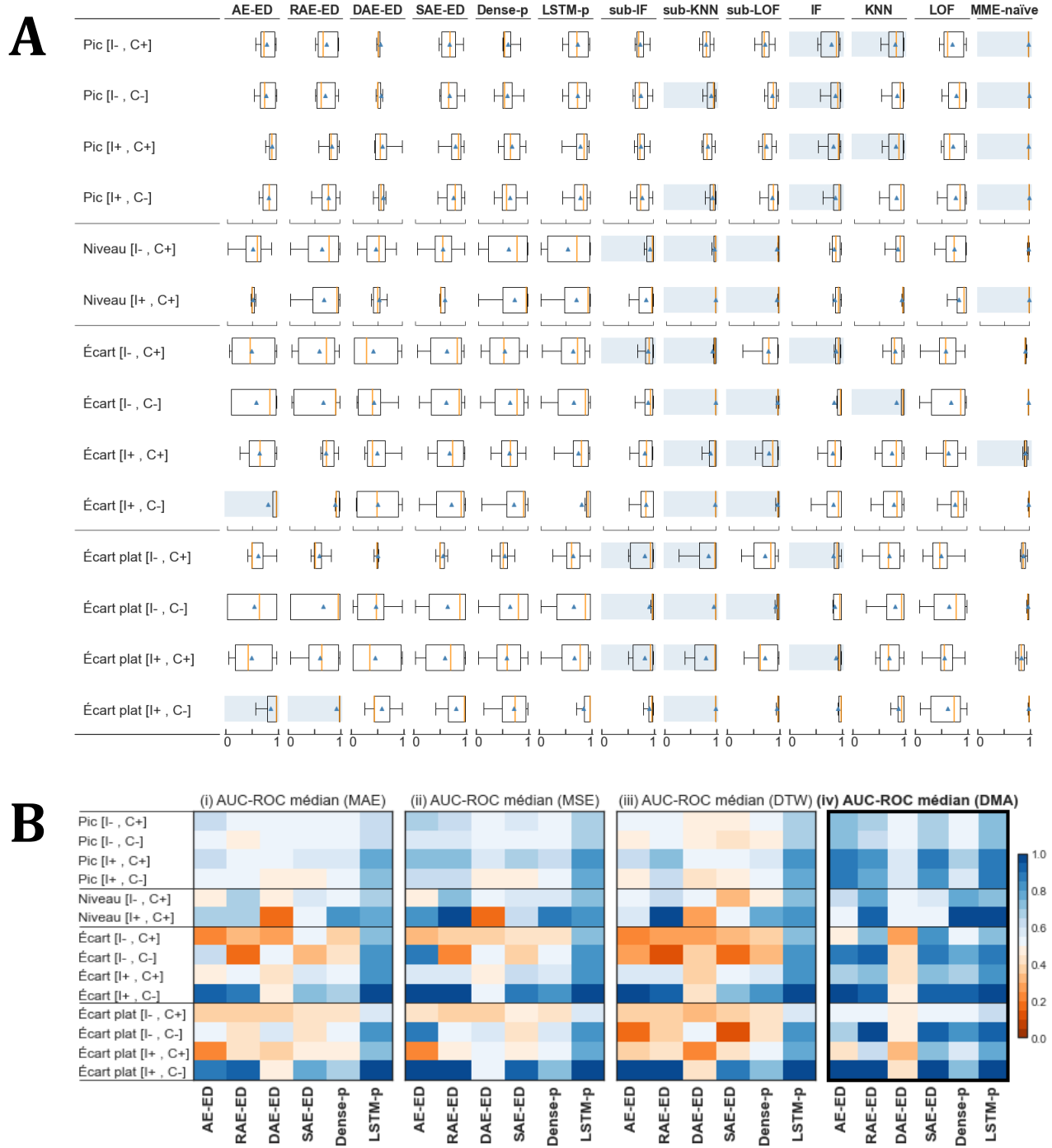


Figure 73 (A) Résultats AUC-ROC pour les mesures de consommation électrique. Les intensités (I) et les taux de contamination (C) sont associés aux deux échelles d'intensité et de contamination, faibles (-) et fortes (+). La médiane des boîtes à moustaches est représentée par une ligne pleine orange, tandis que la moyenne est indiquée par un triangle bleu. (B) Scores d'anomalie des méthodes de reconstruction et de prédiction, calculés en employant (i) la MSE, (ii) la MAE, (iii) la DTW et (iv) la DMA.

Annexe D. Influence des combinaisons, des associations et des transformations sur les performances de détection

A. Humidité relative

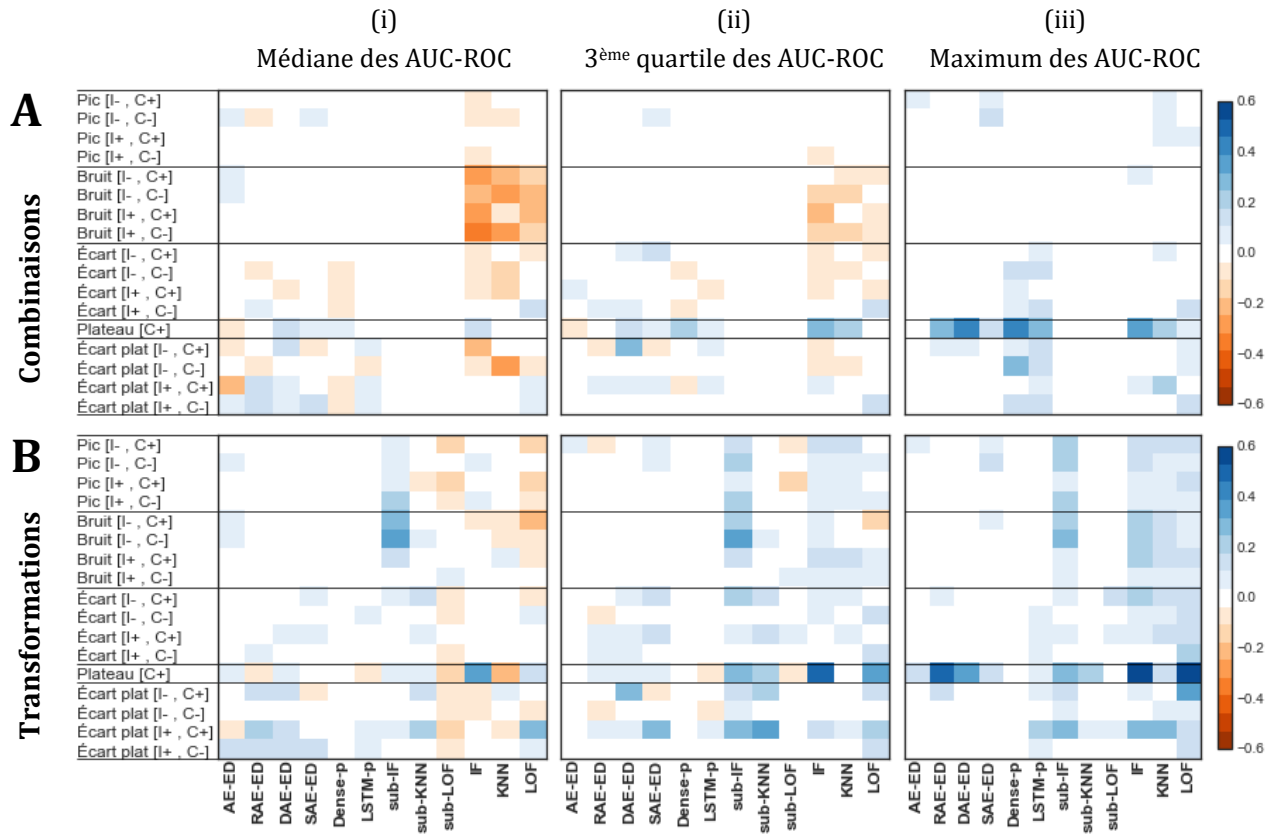


Figure 74 Carte thermique des variations des médianes (i), des troisièmes quartiles (ii) et des maximums (iii) des indices AUC-ROC de l'ensemble des modèles, classés par type d'anomalies et appliqués aux mesures d'humidité relative. Variations des indices AUC-ROC induites par l'apport de combinaisons (A) ou de transformations (B), par rapport aux performances obtenues en utilisant seulement l'humidité relative sans combinaison, associations ou transformations. Une valeur supérieure à 0 (resp. inférieur) présente une amélioration (resp. réduction).

B. Concentration de CO₂ (jeu anormal original)

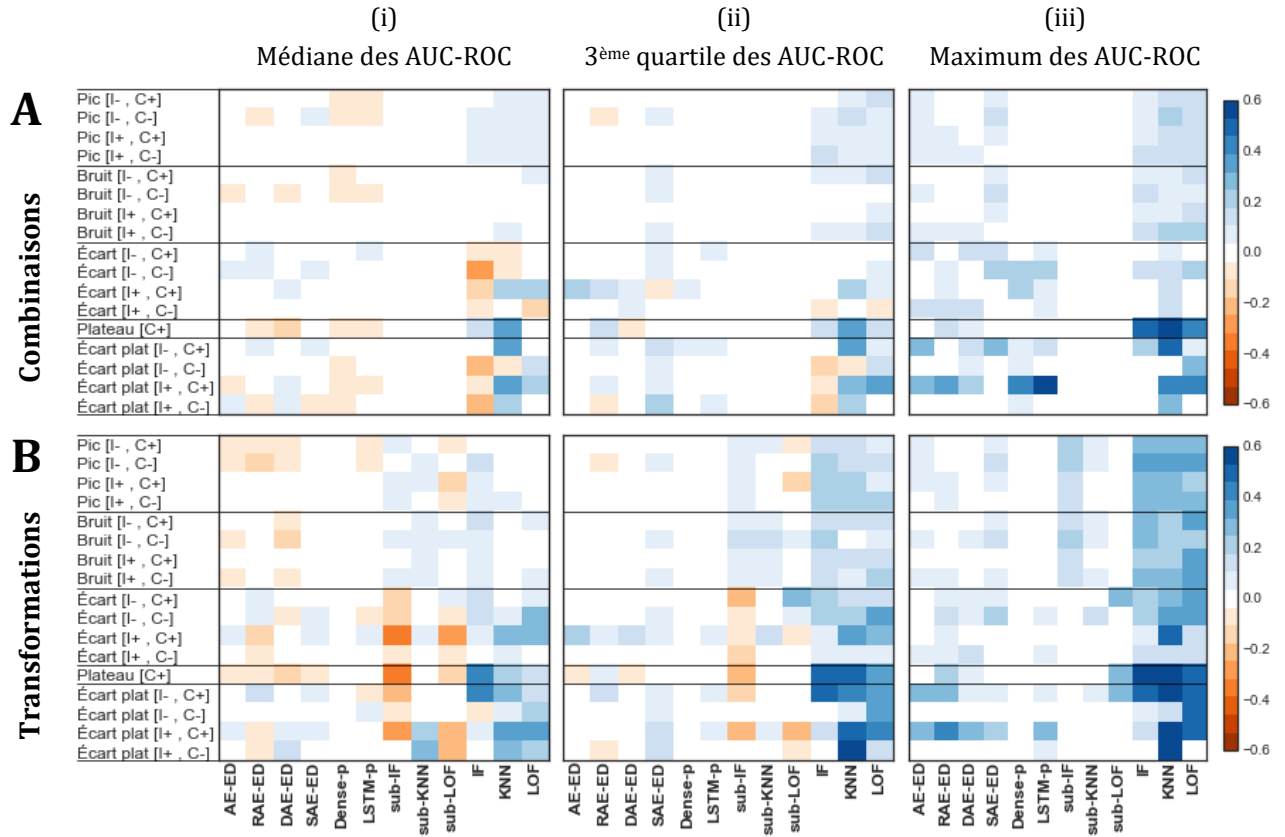


Figure 75 Carte thermique des variations des médianes (i), des troisièmes quartiles (ii) et des maximums (iii) des indices AUC-ROC de l'ensemble des modèles, classés par type d'anomalies et appliqués aux mesures de concentration de CO₂. Variations des indices AUC-ROC induites par l'apport de combinaisons (A) ou de transformations (B), par rapport aux performances obtenues en utilisant seulement la concentration de CO₂ sans combinaison, associations ou transformations. Une valeur supérieure à 0 (resp. inférieure) présente une amélioration (resp. réduction).

C. Consommation électrique (combinaison : temporalité)

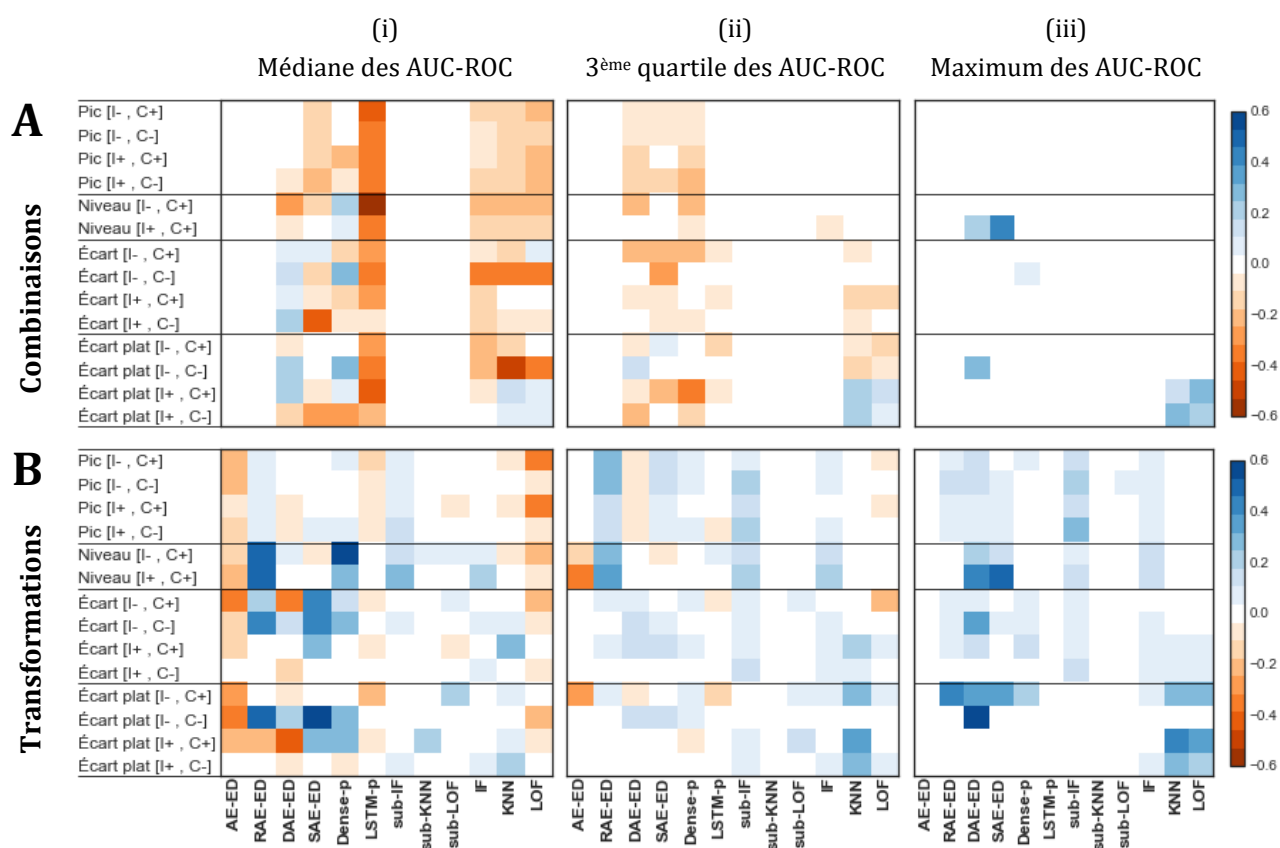


Figure 76 Carte thermique des variations des médianes (i), des troisièmes quartiles (ii) et des maximums (iii) des indices AUC-ROC de l'ensemble des modèles, classés par type d'anomalies et appliqués aux mesures de consommation électrique. Variations des indices AUC-ROC induites par l'apport de combinaisons (A) ou de transformations (B), par rapport aux performances obtenues en utilisant seulement la consommation électrique sans combinaison, associations ou transformations. Une valeur supérieure à 0 (resp. inférieur) présente une amélioration (resp. réduction).

Annexe E. Influence de la taille des fenêtres mobile sur les performances de détection

D. Humidité relative

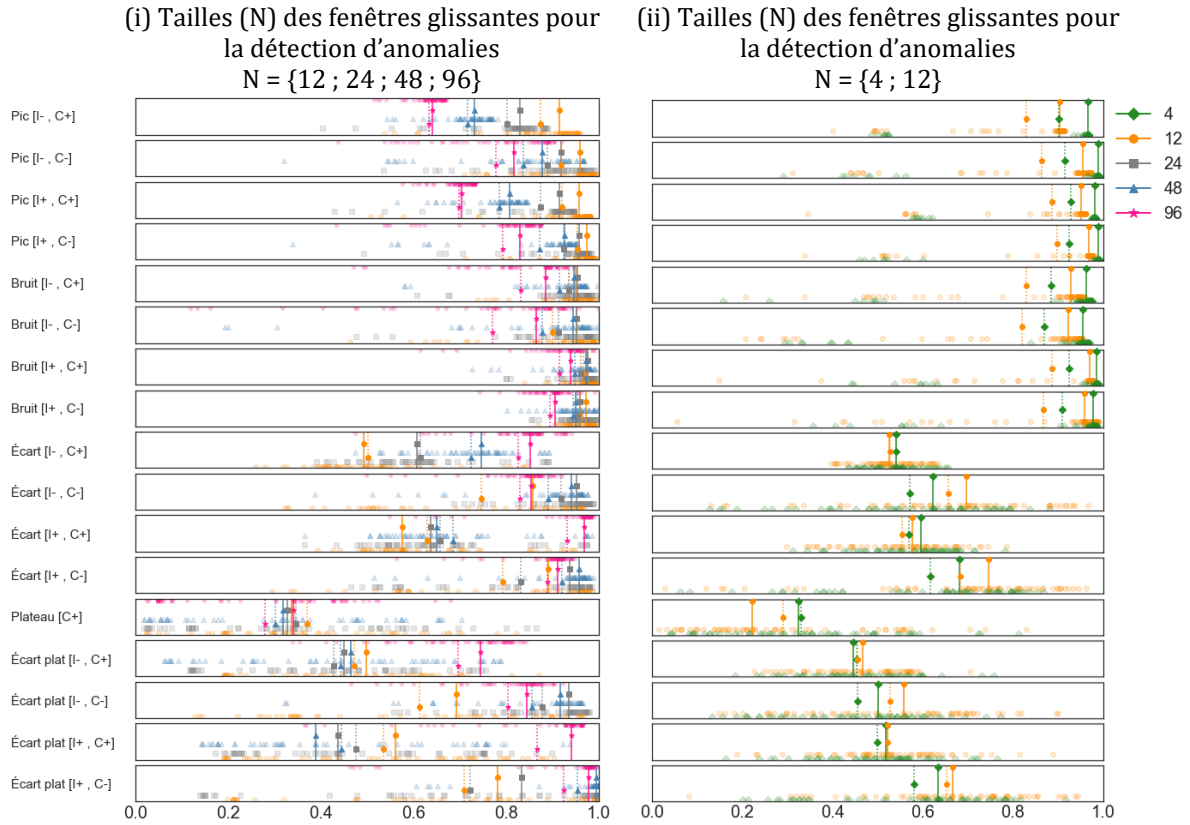


Figure 77 Influence de la taille des fenêtres glissantes de détection, prédiction et reconstruction, sur les performances de détection d'anomalies pour les mesures d'humidité relative. Les modèles de reconstruction et de voisinage séquentiel sont présentés en (i) et les modèles de prédiction en (ii). Les médianes sont représentées par les traits pleins dont les couleurs et les sigles respectifs sont annotés en légende. Les moyennes sont représentées par des traits en pointillés de même couleur et même sigle. La DTW a été employée comme mesure de similarité pour les modèles de prédiction et de reconstruction

E. Concentration de CO₂

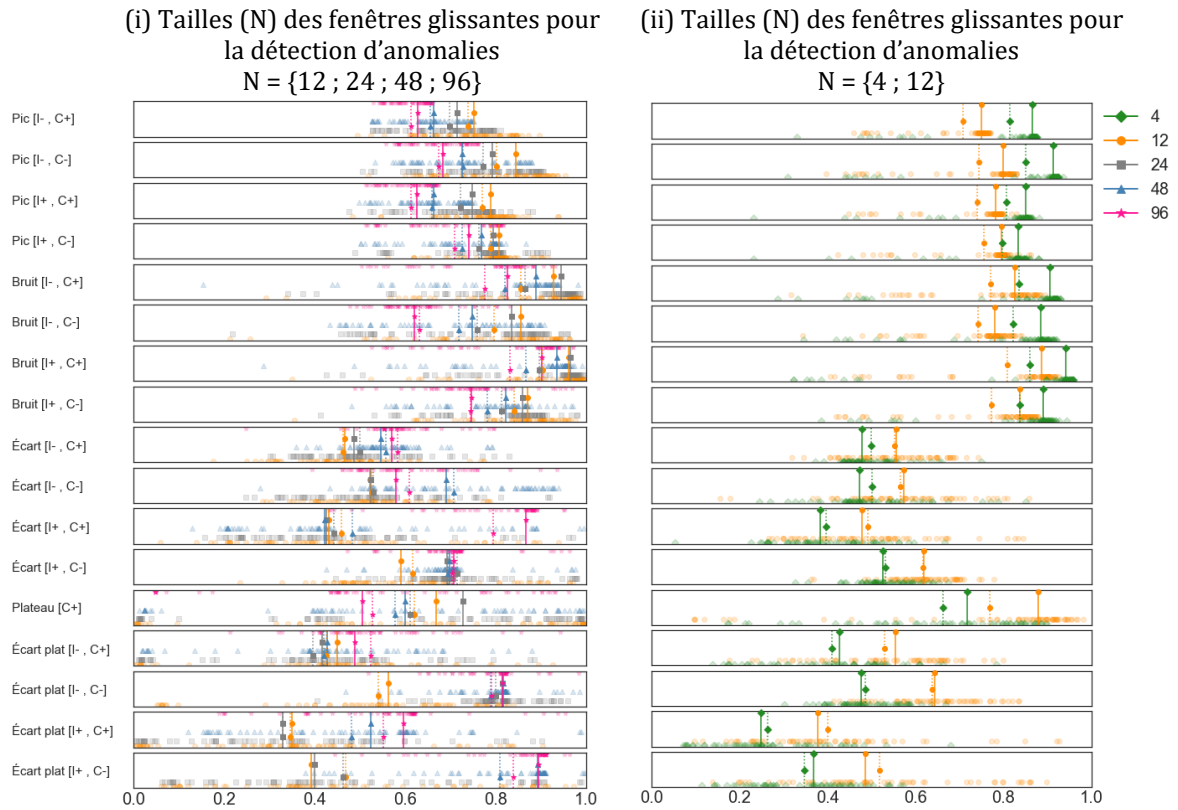


Figure 78 Influence de la taille des fenêtres glissantes de détection, prédiction et reconstruction, sur les performances de détection d'anomalies pour les mesures de concentration de CO₂. Les modèles de reconstruction et de voisinage séquentiel sont présentés en (i) et les modèles de prédiction en (ii). Les médianes sont représentées par les traits pleins dont les couleurs et les sigles respectifs sont annotés en légende. Les moyennes sont représentées par des traits en pointillés de même couleur et même sigle. La DTW a été employée comme mesure de similarité pour les modèles de prédiction et de reconstruction

F. Consommation électrique

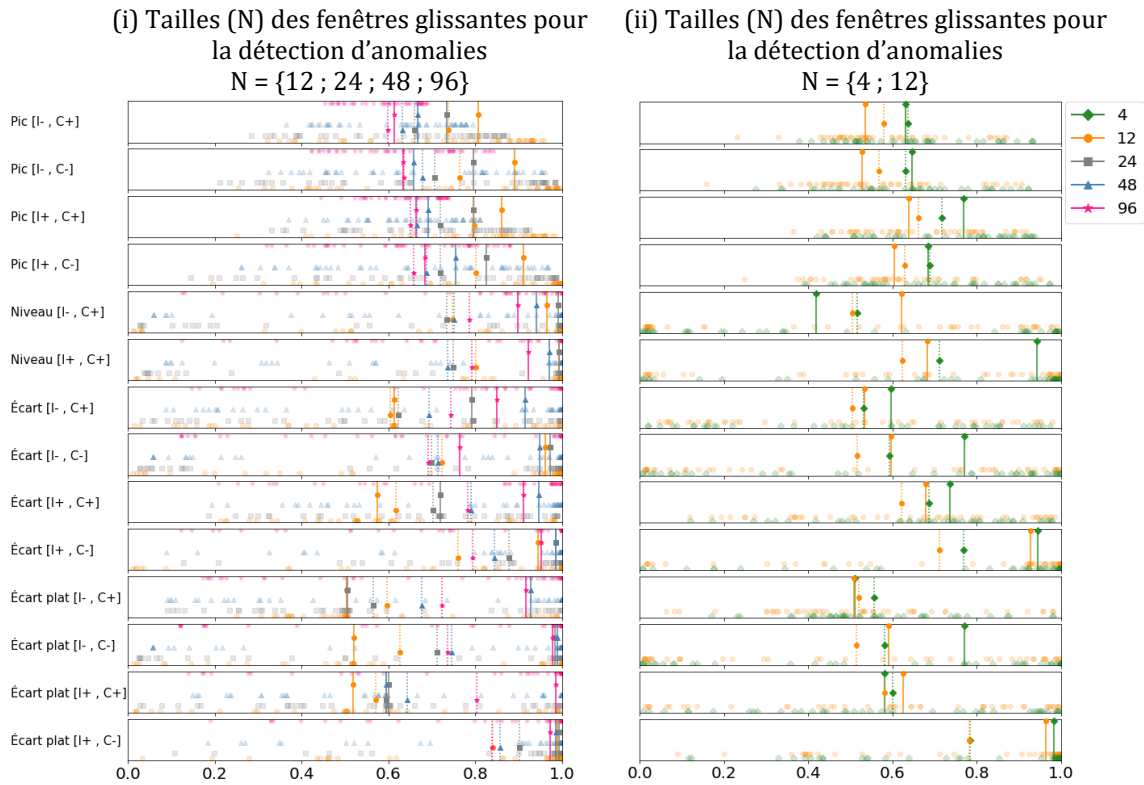


Figure 79 Influence de la taille des fenêtres glissantes de détection, prédiction et reconstruction, sur les performances de détection d'anomalies pour les mesures de consommation électrique. Les modèles de reconstruction et de voisinage séquentiel sont présentés en (i) et les modèles de prédiction en (ii). Les médianes sont représentées par les traits pleins dont les couleurs et les sigles respectifs sont annotés en légende. Les moyennes sont représentées par des traits en pointillés de même couleur et même sigle. La DMA a été employée comme mesure de similarité pour les modèles de prédiction et de reconstruction

Liste des figures

Figure 1 Gauche - Consommation finale énergétique par an et par secteur en France en TWh de 1990 à 2020 avec la répartition des consommations pour l'année 2020 (SDSE). Droite - Répartition des émissions de GES par secteur en France pour l'année 2020 (CITEPA).	1
Figure 2 Illustration de différents profils de bâtiment répartis en différents niveaux selon l'évolution de leur connectivité/intelligence. Inspirée des travaux de Buckman et al. (2014), Karimi et al. (2021) et Sovacool et Furszyfer Del Rio (2020).	13
Figure 3 Exemple de structure simplifiée en quatre couches, perception, réseau, traitement et applicative représentant l'écosystème des bâtiments connectés et des bâtiments intelligents.	15
Figure 4 Exemple de classement en six catégories des objets connectés constitutifs d'un bâtiment connecté ou intelligent. Issue de Nabil (2018).	17
Figure 5 Schématisation d'une chaîne de mesure. Chaîne constituée d'un capteur passif (i) ou d'un capteur actif (ii).	18
Figure 6 Exemple de séries temporelles univariée et multivariée.	19
Figure 7 Exemple d'anomalies provoquées par un phénomène parasite (★) et par un phénomène d'intérêt (■).	31
Figure 8 Exemples de représentation de courbes ROC. Une courbe ROC classique (bleu foncé) et l'aire sous la courbe (AUC-ROC) (zone en bleu clair) reflètent les performances d'un modèle pour une tâche de classification. La courbe en orange représente un modèle optimal avec une valeur de AUC-ROC de 1. La ligne rose représente une classification aléatoire. Inspirée de Tharwat (2021).	40
Figure 9 Les principales familles de techniques de détection d'anomalies.	41
Figure 10 Résultats visuels proposés par différentes méthodes de clustering, (i) la vérité terrain, (ii) les résultats de la méthode K-means et (iii) les résultats de la méthode DBSCAN. Chaque couleur, bleue, orange, verte et rouge représente un cluster ou des observations hors cluster. Issue de Bataineh (2022).	44
Figure 11 Répartition de l'utilisation des techniques de détection d'anomalies selon les domaines d'études, inspiré de Chalapathy and Chawla (2019) et Chandola et al. (2009). TDA : Technique de Détection d'Anomalies, AE : Autoencoder, CNN : Convolutional Neural Network, DBN : Deep Belief Network, DCA : Dilated Convolutional Autoencoder, DNN : Deep Neural Network, GAN : Generative Adversarial Network, GRU : Gated Recurrent Unit, LSTM : Long Short-Term Memory, RBM : Restricted Boltzmann Machine, RNN : Recurrent Neural Network, SPN : Sum Product Network, STN: Spatial Transformer Network, Hybride: modèle hybride composé de modèles différents appariés.	50
Figure 12 Exemple d'une anomalie (★) détectée par corrélation de variables.	60
Figure 13 Spectre des données normales aux données normales avec présence de bruit. Issue de Aggarwal (2017) et tiré de Foulon (2020).	60
Figure 14 Exemple de lissage par moyenne mobile exponentielle de mesure de concentration de CO ₂	61
Figure 15 Exemple de dynamiques de données.	62
Figure 16 Illustration de types d'anomalies. Pics (★), bruit (▲), valeurs constantes (■), écarts (●) et dérives (◆).	67
Figure 17 Représentation des types d'anomalies présentées. Issue de Chandola et al. (2009).	68
Figure 18 Représentation d'anomalies univariées et multivariées. Tirée de Blázquez-García et al. (2022).	69

Figure 19 Exemples d'anomalies indépendantes ponctuelles (i. ★), collectives (ii. ■) et globales (iii. ●)....	72
Figure 20 Exemples d'anomalies univariées ponctuelles (i. ★), collectives (ii. ■) et globales (iii. ●).....	73
Figure 21 Exemples d'anomalies multivariées ponctuelles (i. ★), collectives (ii. ■) et globales (iii. ●).....	74
Figure 22 Schématisation de la taxonomie des problématiques en trois dimensions proposée. L'exemple proposé est de dimensions (3, 3, 3) avec deux axes fixes de trois éléments chacun correspondant à la taille (i) et la nature (ii) des anomalies respectivement sur les abscisses (x) et ordonnées (y). Le troisième axe est de dimension variable mais est composé de trois éléments dans cet exemple. Il correspond aux mesures considérées et est positionné sur la cote (z). Les aires représentent visuellement les zones où les caractéristiques d'un axe x, y , ou z sont constantes.....	76
Figure 23 Exemple de la représentation des problématiques au travers une taxonomie de taille (3, 3, 3). Y sont représentés un problème sous forme d'un point dans l'espace de la taxonomie (i), un problème sous forme de vecteur (ii) et un problème sous forme d'un plan (iii).....	77
Figure 24 Exemple de détections d'ouvertures de fenêtres selon différentes combinaisons de variable. ..	78
Figure 25 Exemples de motifs récurrents anormaux liés aux ouvertures de fenêtres. À gauche (i) : deux jours sans ouverture dans une pièce avec des mesures de température, d'humidité relative et de concentration de CO ₂ intérieure. À droite (ii) : la même pièce avec plusieurs ouvertures sur une période de deux jours de mesure.....	84
Figure 26 Détection du comportement des occupants selon les données utilisées et les objectifs visés : (i) <i>a posteriori</i> , (ii) en temps réel et (ii) par prédiction.	86
Figure 27 Exemple de modélisation d'ouvertures de fenêtres pour deux modèles.	90
Figure 28 Exemples de vraies et de fausses ouvertures.	91
Figure 29 Exemples des scores de précision associés aux actions d'ouvertures et de fermetures de vraies ouvertures.....	92
Figure 30 Illustration des principes clés des modèles de séparateurs à vaste marge à travers trois exemples de classification dans un espace en deux dimensions (x, y). De gauche à droite : (i) montre l'infini de frontières possibles pour séparer deux classes ; (ii), modifiée d'après Cortes et Vapnik (1995), présente les vecteurs de support en orange plein, définissant la distance optimale pour séparer les deux classes ; (iii) modifiée d'après Theodoridis and Koutroumbas (2008), représente l'hyperplan optimal obtenu en utilisant un noyau RBF pour des données non linéairement séparables (ligne orange discontinue).	93
Figure 31 Exemple de réduction à une dimension par analyse discriminante linéaire pour un problème à deux classes en deux dimensions (x, y). Les données sont projetées sur un vecteur propre, choisi ici aléatoirement, pour maximiser le rapport entre la variabilité interclasse et la variance intraclasse.....	94
Figure 32 Exemples de représentation d'un arbre de décision pour des données non numériques (i) et des données numériques (ii), issus de Duda et al. (2001). Dans l'arbre (i), les nœuds correspondent aux questions, les branches aux réponses, et les feuilles portent les noms des fruits en fonction de leurs caractéristiques. Dans l'arbre (ii) deux régions de décision, $\mathcal{R}_1$ et $\mathcal{R}_2$, sont représentées.....	95
Figure 33 Représentation d'un réseau de neurones (i), d'une cellule neuronale classique (ii), d'une cellule neuronale récurrente (iii), et de la même cellule dépliée dans le temps (iv). Les couches d'entrée, cachée et de sortie à un moment donné t sont respectivement annotées x_t , h_t et y_t . Les poids associés aux connections des couches entrée-cachée, cachée-cachée et cachée-sortie sont respectivement annotés w_{xh} , w_{hh} et w_{hy} ...	97
Figure 34 Approche d'entraînement réalisée selon les modèles à deux étapes temporelles successives t et $t+1$	100
Figure 35 Positionnement des capteurs étudiés.	101
Figure 36 Période de test avec occupation et inoccupation (zone bleue centrale annotée).	102
Figure 37 Distribution des données pour les mesures de température et de concentration de CO ₂ intérieures (A) ainsi que pour leurs transformations et combinaisons respectives. Les transformations incluent le résidu (B), la DMM (C), et la dérivée (D). Les données sont représentées pour les périodes d'entraînement (en bleu), de test (en orange) et de test sans occupants (en rose).	106
Figure 38 Schéma du protocole d'expérimentation pour la combinaison et l'association des mesures. ..	107

Figure 39 F ₁ -scores pour chaque transformation, association, et combinaison de mesure d'ambiance pour un modèle donné	113
Figure 40 Métriques additionnelles pour les 20 meilleurs modèles des combinaisons de température (T _{int}), d'humidité (H _{int}) et de concentration de CO ₂ (CO _{2int}) intérieures.....	116
Figure 41 Carte thermique des ratios d'ouvertures détectées par les 20 meilleurs modèles, toutes combinaisons de mesures confondues. Les ouvertures sont numérotées de 1 à 18 et sont mentionnées en abscisse sous l'abréviation « O ». Par exemple, la cinquième ouverture est notée « O ₅ ». L'O ₈ n'a jamais été détectée et n'est donc pas affichée.....	118
Figure 42 Exemples d'anomalies générées artificiellement sur une mesure de concentration de CO ₂ . La variable originale est incluse pour évaluer les différences induites par la contamination. Toutes les anomalies non annotées correspondent à des pics d'intensité variable.	132
Figure 43 Exemples illustrant le fonctionnement de la méthode LOF. (i) Illustration du calcul de la distance d'accessibilité pour trois points différents et avec $K = 5$. Inspirée de Breunig et al. (2000). (ii) Présentation du score LOF pour des données en 2 dimensions. Les périmètres des cercles rouges représentent le score LOF, plus ils sont étendus, plus le score s'éloigne de 1. Tirée de Kriegel et al. (2009).	135
Figure 44 Exemple de fonctionnement de la méthode <i>Isolation Forest</i> . (i) Illustration d'un arbre composé de 12 nœuds qui isole une observation normale (x_1). Présentation d'un autre arbre composé de seulement trois nœuds permettant d'isoler une observation anormale (x_0). (iii) Montre la moyenne des distances pour chaque observation selon le nombre d'arbres de la forêt. Peu importe le nombre d'arbres, la distance moyenne de l'observation anormale (x_0) reste inférieure à celle de l'observation normale (x_1). Tirée de Liu et al. (2008).	136
Figure 45 Processus de détection d'anomalies avec des modèles de prédiction. (i) Exemple de rétro-prédiction : une observation, notée x_t , est prédite en utilisant une séquence (zone grise) de N observations antérieures. (ii) Exemple de comparaison entre valeurs prédites et valeurs réelles. Les figures (i) et (ii) sont tirées de Blázquez-García et al. (2022).	138
Figure 46 Comparaison de la structure d'un réseau de neurones « classique » (i) avec celle commune à l'Auto-Encodeur et à l'Encodeur-Décodeur (ii).	139
Figure 47 Représentation du fonctionnement d'un DAE. x représente la séquence d'observations d'origine des entrées, x' la séquence bruitée utilisée par le modèle (rouge pour les observations bruitées), z l'espace latent correspondant et y la sortie reconstruite par le modèle.....	140
Figure 48 Architecture d'un AE classique (i) et d'un Sparse AE (ii). Les neurones désactivés sont représentés en gris et ne sont plus reliés aux neurones actifs. Inspirée de Kokie-Akino et Wang (2020).	141
Figure 49 Représentation d'observations normales (bleues) et anormales (rouges) dans l'espace latent d'un VAE. Issue de Agiulli et al. (2022).	142
Figure 50 Exemple d'application de la DTW entre deux séquences de tailles M et N différentes (i). Adaptée de Mori et al. (2016). Exemple des chemins suivis pour calculer les similarités entre deux séquences à l'aide de la DTW (ii) et de la MSE (iii).	147
Figure 51 Exemple de détection d'anomalie en utilisant un score d'anormalité des observations. La ligne orange représente un seuil de décision (a) de 220 et la ligne rose un seuil (b) de 100. Les anomalies sont notées « A » et les observations sont notées « O ».	149
Figure 52 Exemple de détection d'anomalie en utilisant un score d'anormalité de séquences. La ligne orange représente un seuil de décision (a) de 200, tandis que la ligne rose correspond à un seuil (b) de 100. Les anomalies sont notées « A » et les observations sont notées « O ».	150
Figure 53 Position des capteurs de la maison.	152
Figure 54 Temps d'exécution moyen par observation pour chaque modèle appliqué sur une combinaison ou transformation de variable unique (échelle logarithmique). Les modèles semi-supervisés sont représentés par une couleur bleue et les modèles supervisés par de l'orange. L'écart type est présenté par une barre noire. Inspirée de Schmidl et al. (2022).	158
Figure 55 Évolution du poids moyen des modèles selon la taille des fenêtres de reconstruction et d'observations utilisées. L'impact du nombre de combinaisons et de transformations employées peut être estimé grâce à l'écart type.	160

Figure 56 Mesures de température d'air intérieure saine avant contamination.	161
Figure 57 (A) Résultats AUC-ROC pour les mesures de température intérieure d'air. Les intensités (I) et les taux de contamination (C) sont associés aux deux échelles d'intensité et de contamination, faibles (-) et fortes (+). La médiane des boîtes à moustaches est représentée par une ligne pleine orange, tandis que la moyenne est indiquée par un triangle bleu. (B) Scores d'anormalité des méthodes de reconstruction et de prédiction, calculés en employant (i) la MSE, (ii) la MAE, (iii) la DTW et (iv) la DMA.	162
Figure 58 Carte thermique des médianes (i), des troisièmes quartiles (ii) et des maximums (iii) des indices AUC-ROC de l'ensemble des modèles, classés par type d'anomalies et appliqués aux mesures de température d'air intérieure. L'encadré blanc met en évidence les performances des modèles prédictifs sur trois types d'anomalies spécifiques.	164
Figure 59 Mesures de l'humidité relative intérieure saine avant contamination.	165
Figure 60 Carte thermique des médianes (i), des troisièmes quartiles (ii) et des maximums (iii) des indices AUC-ROC de l'ensemble des modèles, classés par type d'anomalies et appliqués aux mesures d'humidité relative intérieure. La DTW a été utilisée comme mesure de similarité pour les modèles de reconstruction et de prédiction.	166
Figure 61 Mesures de concentration de CO ₂ intérieure saine avant contamination.	167
Figure 62 Carte thermique des médianes (i), des troisièmes quartiles (ii) et des maximums (iii) des indices AUC-ROC de l'ensemble des modèles, classés par type d'anomalies et appliqués aux mesures de concentration de CO ₂ . La DTW a été utilisée comme mesure de similarité pour les modèles de reconstruction et de prédiction.	168
Figure 63 Carte thermique des médianes (i), des troisièmes quartiles (ii) et des maximums (iii) des indices AUC-ROC de l'ensemble des modèles, classés par type d'anomalies et appliqués aux mesures de concentration de CO ₂ . (A) les neuf ensembles anormaux originaux, (B) les neuf ensembles anormaux avec uniquement la période d'occupation et (C) les neuf ensembles anormaux avec uniquement la période d'absence des occupants. La DTW a été utilisée comme mesure de similarité pour les modèles de reconstruction et de prédiction.	170
Figure 64 Mesure de consommation électrique cumulée saine avant contamination.	173
Figure 65 Carte thermique des médianes (i), des troisièmes quartiles (ii) et des maximums (iii) des indices AUC-ROC de l'ensemble des modèles, classés par type d'anomalies et appliqués aux mesures de consommation électrique. La DMA a été utilisée comme mesure de similarité pour les modèles de reconstruction et de prédiction.	174
Figure 66 Carte thermique des variations des médianes (i), des troisièmes quartiles (ii) et des maximums (iii) des indices AUC-ROC de l'ensemble des modèles, classés par type d'anomalies et appliqués aux mesures de température d'air intérieure. Variations des indices AUC-ROC induites par l'apport de combinaisons (A) ou de transformations (B), par rapport aux performances obtenues en utilisant seulement la température sans combinaison, associations ou transformations. Une valeur supérieure à 0 (resp. inférieure) présente une amélioration (resp. réduction).	177
Figure 67 Influence de la taille des fenêtres glissantes de détection, prédiction et reconstruction, sur les performances de détection d'anomalies. Les modèles de reconstruction et de voisinage séquentiel sont présentés en (i) et les modèles de prédiction en (ii). Les médianes sont représentées par les traits pleins dont les couleurs et les sigles respectifs sont annotés en légende. Les moyennes sont représentées par des traits en pointillés de même couleur et même sigle.	180
Figure 68 Carte thermique des médianes (i), des troisièmes quartiles (ii) et des maximums (iii) des indices AUC-ROC de l'ensemble des modèles, classés par type d'anomalies et appliqués aux mesures de consommation électrique. (A) les 9 jeux d'observations anormales originaux de la température et (B) les neuf jeux d'observations anormales composés uniquement des résultats des modèles de fenêtres d'une taille de 48 ou 96 observations.	182
Figure 69 Mesure de température d'air intérieure comportant des anomalies (séjour de la maison).	187
Figure 70 Exemple d'application de modèles de détection d'anomalies sur une mesure de température non étiquetée. Une couleur proche du rouge représente une observation ayant un score de suspicion proche de 1 tandis qu'une couleur proche du jaune indique un score proche de 0,5.	189

Figure 71 (A) Résultats AUC-ROC pour les mesures d'humidité relative d'air intérieur. Les intensités (I) et les taux de contamination (C) sont associés aux deux échelles d'intensité et de contamination, faibles (-) et fortes (+). La médiane des boîtes à moustaches est représentée par une ligne pleine orange, tandis que la moyenne est indiquée par un triangle bleu. (B) Scores d'anormalité des méthodes de reconstruction et de prédiction, calculés en employant (i) la MSE, (ii) la MAE, (iii) la DTW et (iv) la DMA..... 223

Figure 72 (A) Résultats AUC-ROC pour les mesures de concentration de CO₂ intérieure. Les intensités (I) et les taux de contamination (C) sont associés aux deux échelles d'intensité et de contamination, faibles (-) et fortes (+). La médiane des boîtes à moustaches est représentée par une ligne pleine orange, tandis que la moyenne est indiquée par un triangle bleu. (B) Scores d'anormalité des méthodes de reconstruction et de prédiction, calculés en employant (i) la MSE, (ii) la MAE, (iii) la DTW et (iv) la DMA..... 224

Figure 73 (A) Résultats AUC-ROC pour les mesures de consommation électrique. Les intensités (I) et les taux de contamination (C) sont associés aux deux échelles d'intensité et de contamination, faibles (-) et fortes (+). La médiane des boîtes à moustaches est représentée par une ligne pleine orange, tandis que la moyenne est indiquée par un triangle bleu. (B) Scores d'anormalité des méthodes de reconstruction et de prédiction, calculés en employant (i) la MSE, (ii) la MAE, (iii) la DTW et (iv) la DMA..... 225

Figure 74 Carte thermique des variations des médianes (i), des troisièmes quartiles (ii) et des maximums (iii) des indices AUC-ROC de l'ensemble des modèles, classés par type d'anomalies et appliqués aux mesures d'humidité relative. Variations des indices AUC-ROC induites par l'apport de combinaisons (A) ou de transformations (B), par rapport aux performances obtenues en utilisant seulement l'humidité relative sans combinaison, associations ou transformations. Une valeur supérieure à 0 (resp. inférieur) présente une amélioration (resp. réduction). 226

Figure 75 Carte thermique des variations des médianes (i), des troisièmes quartiles (ii) et des maximums (iii) des indices AUC-ROC de l'ensemble des modèles, classés par type d'anomalies et appliqués aux mesures de concentration de CO₂. Variations des indices AUC-ROC induites par l'apport de combinaisons (A) ou de transformations (B), par rapport aux performances obtenues en utilisant seulement la concentration de CO₂ sans combinaison, associations ou transformations. Une valeur supérieure à 0 (resp. inférieur) présente une amélioration (resp. réduction). 227

Figure 76 Carte thermique des variations des médianes (i), des troisièmes quartiles (ii) et des maximums (iii) des indices AUC-ROC de l'ensemble des modèles, classés par type d'anomalies et appliqués aux mesures de consommation électrique. Variations des indices AUC-ROC induites par l'apport de combinaisons (A) ou de transformations (B), par rapport aux performances obtenues en utilisant seulement la consommation électrique sans combinaison, associations ou transformations. Une valeur supérieure à 0 (resp. inférieur) présente une amélioration (resp. réduction). 228

Figure 77 Influence de la taille des fenêtres glissantes de détection, prédiction et reconstruction, sur les performances de détection d'anomalies pour les mesures d'humidité relative. Les modèles de reconstruction et de voisinage séquentiel sont présentés en (i) et les modèles de prédiction en (ii). Les médianes sont représentées par les traits pleins dont les couleurs et les sigles respectifs sont annotés en légende. Les moyennes sont représentées par des traits en pointillés de même couleur et même sigle. La DTW a été employée comme mesure de similarité pour les modèles de prédiction et de reconstruction 229

Figure 78 Influence de la taille des fenêtres glissantes de détection, prédiction et reconstruction, sur les performances de détection d'anomalies pour les mesures de concentration de CO₂. Les modèles de reconstruction et de voisinage séquentiel sont présentés en (i) et les modèles de prédiction en (ii). Les médianes sont représentées par les traits pleins dont les couleurs et les sigles respectifs sont annotés en légende. Les moyennes sont représentées par des traits en pointillés de même couleur et même sigle. La DTW a été employée comme mesure de similarité pour les modèles de prédiction et de reconstruction 230

Figure 79 Influence de la taille des fenêtres glissantes de détection, prédiction et reconstruction, sur les performances de détection d'anomalies pour les mesures de consommation électrique. Les modèles de reconstruction et de voisinage séquentiel sont présentés en (i) et les modèles de prédiction en (ii). Les médianes sont représentées par les traits pleins dont les couleurs et les sigles respectifs sont annotés en légende. Les moyennes sont représentées par des traits en pointillés de même couleur et même sigle. La DTW a été employée comme mesure de similarité pour les modèles de prédiction et de reconstruction 231

Liste des tableaux

Tableau 1 Thématiques des objectifs et bénéfices régulièrement associés aux bâtiments connectés et intelligents dans la littérature.....	21
Tableau 2 Définitions d' <i>outlier</i> et d' <i>anomalies</i> dans la littérature.	30
Tableau 3 Matrice de confusion pour un problème de détection d'anomalie.	38
Tableau 4 Exemples de données quantitatives et qualitatives dans le bâtiment.	59
Tableau 5 Types d'anomalies issues de mesures dans la littérature (capteurs ou séries temporelles). Les références suivies d'une * représentent des travaux qui recensent et introduisent des types issus de plusieurs publications. Inspiré de Kraeim (2021).	65
Tableau 6 Typologie d'anomalie adaptée aux anomalies et mesures issues du bâtiment.	75
Tableau 7 Liste et plage d'optimisation des paramètres et hyperparamètres de chaque modèles.	98
Tableau 8 Caractéristiques des données utilisées pour l'entraînement.	101
Tableau 9 Caractéristiques des données de test, avec (O) pour la période occupée, et (Ø) pour la période non occupée. Les valeurs qui diffèrent le plus des données d'entraînement sont en gras.	103
Tableau 10 F ₁ -score moyen et écart-type des 20 transformations et associations basées sur une mesure pour chaque modèle. Les meilleurs résultats par modèle sont en gras.	108
Tableau 11 F ₁ -score pour les cinq transformations de base appliquées aux mesures de température intérieure (T _{int}) et d'humidité intérieure (H _{int}). Les meilleurs résultats par modèle sont en gras.	109
Tableau 12 F-score moyen et écart-type des 20 meilleures combinaisons entre : la température d'air intérieure (T _{int}), l'humidité intérieure (H _{int}), la concentration de CO ₂ intérieure (C _{int}), la température d'air extérieure (T _{ext}), et l'humidité extérieure (H _{ext}). Les meilleurs résultats par modèle sont en gras.....	111
Tableau 13 Résumé des spécificités relatives aux entrées et sorties des modèles de voisinage, de prédiction et de reconstruction utilisés.	144
Tableau 14 Paramètres utilisés pour la contamination des jeux de test. M représente le nombre d'observations composant la variable contaminée.	154
Tableau 15 Présentation des transformations appliquées.	156
Tableau 16 Poids minimum, moyen et maximum de l'ensemble des modèles semi-supervisés entraînés, pour les familles de reconstruction (■) et de prédiction (▲). Les poids sont affichés en Mégaoctets (Mo).	159