



HAL
open science

Modèle de régression linéaire multiple pour prédire la température intérieure sur les bâtiments éducatifs en milieu Méditerranéen

Keovathana Run, Franck Cevaer, Jean-François Dubé

► **To cite this version:**

Keovathana Run, Franck Cevaer, Jean-François Dubé. Modèle de régression linéaire multiple pour prédire la température intérieure sur les bâtiments éducatifs en milieu Méditerranéen. Congrès Français du Génie Civil - CFGC 2023, May 2023, Saclay, France. 10.26168/ajce.41.1.22 . hal-04793599

HAL Id: hal-04793599

<https://hal.science/hal-04793599v1>

Submitted on 20 Nov 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.

Modèle de régression linéaire multiple pour prédire la température intérieure sur les bâtiments éducatifs en milieu Méditerranéen

Keovathana Run¹, Franck Cévaër¹, Jean-François Dubé¹

¹LMGC, Université de Montpellier, UMR CNRS 5508, cc048, Place E. Bataillon, 34095 Montpellier cedex 5, France

RESUME Pour les bâtiments existants, la rénovation énergétique est importante, de même que les stratégies opérationnelles et de contrôle qui nécessitent une surveillance continue pour économiser l'énergie. Dans le cadre de la surveillance continue et de la gestion de l'énergie, la prédiction joue un rôle important. La prédiction peut être plus précise avec le bon choix de la technique et des variables prédictives. Dans cette étude, les analyses sont effectuées sur des bâtiments universitaires (IUT de Nîmes) situés dans le sud de la France en période estivale. Le modèle de régression linéaire multiple est adopté pour identifier les variables les plus influentes sur l'évolution de la température intérieure et aussi pour explorer sa robustesse dans la prédiction. Le modèle proposé se compose de cinq prédicteurs tels que (1) la temporisation de trois heures de la température extérieure, (2) la température extérieure, (3) l'humidité extérieure, (4) l'humidité intérieure et (5) l'orientation de la salle. Les résultats obtenus sont satisfaisants puisque ce modèle prédit non seulement 76% des valeurs réelles avec un taux d'erreur moyen de 3% (ou 1,04 °C), mais affiche également une capacité de prédiction de 80% lors de l'étape de vérification. Ce modèle a cependant une tendance à surestimer les températures les plus élevées et minimiser les températures les plus basses.

Mots-clés Prédiction de la température intérieure, Modèle de régression linéaire multiple, Bâtiments éducatifs, Étude de cas de l'été méditerranéen.

I. INTRODUCTION

Le secteur du bâtiment représente environ 40% des émissions de dioxyde de carbone (CO_2) dans les pays développés, 37% de la consommation d'énergie et 40% de la production de déchets (Deshayes, 2012). Il a donc un impact important sur le changement climatique et sur le déficit commercial actuel dû aux importations d'énergie (Sidler, 2012). Pour économiser l'énergie de chauffage et de refroidissement, il est important de mettre en œuvre des stratégies opérationnelles et de contrôle appropriées qui nécessitent une prévision pour une surveillance continue (Deb et al., 2017). La prévision de la demande et de l'utilisation de l'énergie dépend fortement de la température de l'air, car il s'agit d'une caractéristique importante étroitement liée au confort thermique intérieur et qui affecte la consommation d'énergie des bâtiments (Deb et al., 2017 ; Tham and Ullah, 1993).

Plusieurs techniques ont été créées avec le temps pour anticiper l'estimation de la température intérieure. Elles relèvent de trois catégories de base : physique (modèle à boîte blanche), base de données (modèle à boîte noire) et hybride (modèle à boîte grise). Les modèles à boîte blanche sont

basés sur des équations de bilan énergétique et massique différentiel intégral, qui nécessitent une connaissance approfondie des caractéristiques d'un bâtiment, des gains d'éclairage, des charges d'équipement et de la densité réelle d'occupation (Attoue et al., 2018). Cette méthode a un coût de calcul élevé et il faut une compréhension détaillée de la structure globale du bâtiment, de ses composants et des systèmes énergétiques (Ramadan et al., 2021 ; Wang and Ma, 2008). En utilisant des méthodologies basées sur les données, la modélisation en boîte grise cherche à combler l'écart entre les bâtiments réels et les simulations en boîte blanche (Elmaz et al., 2021). Les modèles à boîte noire peuvent utiliser différentes méthodes telles que ANN (Artificial Neural Network), SVM (Support Vector Machine), MLR (Multiple Linear Regression) pour modéliser des données massives avec une grande précision de prédiction et des performances différentes par leur temps de calcul (Afram et Janabi-Sharifi, 2015 ; Afroz et al., 2018).

Cette recherche est effectuée sur les bâtiments éducatifs existants avec la disponibilité de données historiques suffisantes, le modèle à boîte noire est donc adopté. Moretti et al. ont utilisé le MLR pour prédire les tendances de la température dans les bâtiments historiques et ont constaté que 8 variables d'entrée permettent d'obtenir un coefficient de détermination R^2_{adj} de 0,92. Les données d'entrée comprennent la température de l'air extérieur, le rayonnement solaire, le système de chauffage et le type de saison (Moretti et al., 2021). Özbalta et al. ont utilisé le jour de l'année, la température extérieure, l'humidité relative extérieure et la vitesse du vent dans des modèles de régression et ont constaté que le meilleur modèle atteignait un R^2 de 0,87 et 0,81 en entraînement et en test, respectivement (Özbalta et al., 2012). Pérez-Lombard et al. ont comparé les modèles de régression linéaire multiple et d'arbre aléatoire en utilisant la température extérieure, l'humidité relative, la vitesse du vent, la visibilité, la pression, la température du point de rosée et la consommation totale d'énergie électrique. Les résultats montrent que le R^2 du modèle de régression varie entre 0,83 et 0,95 en entraînement sur différentes zones d'étude (Pérez-Lombard et al., 2008).

On constate que le climat extérieur est couramment utilisé comme données saisies. Cependant, peu d'études prennent en compte les caractéristiques du bâtiment (surface et orientation) dans le développement du modèle. L'objectif de cet article est d'étudier le taux de précision du modèle de prévision de la température intérieure par la technique de régression linéaire multiple en utilisant le climat extérieur, les caractéristiques du bâtiment et l'indice temporel.

II. MÉTHODOLOGIE

Le modèle de prédiction est basé sur la régression linéaire multiple, qui est l'analyse de la distribution conditionnelle $Y|x$ de la variable réponse Y donnée par le vecteur de prédicteurs $x = (x_1, \dots, x_p)^T$. Le modèle de MLR peut être décrit comme suit :

$$Y = \beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2} + \dots + \beta_p x_{i,p} \quad (1)$$

Les données utilisées pour l'analyse comprennent des variables qualitatives et quantitatives pour construire le modèle en utilisant un « *Training set* » de k échantillons aléatoires des n cas pour p prédicteurs afin de construire un modèle où $10p \leq n/2 \leq k \leq 0,9n$. Puis utilise un « *Testing set* » pour les autres $n - k$ cas afin de confirmer que le modèle construit avec *training set* est bon. Cette technique peut permettre de réduire les biais, mais nécessite $n \geq 20p$. Un modèle final est valide pour l'inférence s'il a été démontré qu'il était utile sur d'autres données obtenues après l'élaboration du

modèle (Olive, 2017). Il y a eu 25290 cas collectés dans l'étude de cas. Par conséquent, 70% des données collectées seront utilisées comme un *training set* et 30 % comme un *testing set* sélectionné au hasard sur une période de trois mois (du 21 juin au 21 septembre 2020). Un *validating set* est mis en œuvre du 01 au 31 mai 2020, recueilli sur une période différente pour vérifier le modèle.

A. ÉVALUATION DU MODÈLE

Les métriques conventionnelles sont incorporées comme indicateur primaire de l'évaluation des performances du modèle :

- R^2 (Coefficient de détermination) : le coefficient de l'adéquation des valeurs prédites par rapport aux valeurs réelles. La valeur de 0 à 1 est interprétée comme un pourcentage.

$$R^2 = \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y}_i)^2} \quad (2)$$

- MAE (Erreur absolue moyenne) : la moyenne de la différence absolue entre les valeurs réelles et les valeurs prédites de données. Elle mesure la moyenne des résidus de données.

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (3)$$

- MSE (Erreur quadratique moyenne) : la moyenne de la différence au carré entre les valeurs réelles et les valeurs prédites de données. Elle mesure la variance des résidus.

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (4)$$

- RMSE (Racine de l'erreur quadratique moyenne) : elle mesure l'écart-type des résidus.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (5)$$

- MAPE (Erreur absolue moyenne en pourcentage) : les erreurs en pourcentage sont calculées en termes d'erreurs absolues, sans tenir compte du signe.

$$MAPE = \frac{1}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (6)$$

Où \hat{y} et \bar{y} sont respectivement la valeur prédite et la valeur moyenne de la valeur mesurée y au $i^{\text{ème}}$ instant, et N représente le nombre de prédictions.

B. ÉTUDE DE CAS

Cette étude porte sur les bâtiments universitaires de l'IUT de Nîmes, situés dans le sud de la France à 43°49'N de longitude et 4°19'E de latitude. Cette situation géographique est soumise au climat méditerranéen qui est classé CSA selon Koeppen et Geiger (Peel et al., 2007). Selon Météo France - Courbessac, des températures supérieures à 34 °C sont fréquemment atteintes en juillet et août. Par conséquent, les bâtiments à ventilation naturelle de cette région ont la particularité d'être très

sensibles au confort d'été où le comportement de l'utilisateur et l'inertie du bâtiment ont un impact important.

Trois bâtiments d'enseignement et un bâtiment administratif du campus sont examinés pendant l'été du 21 juin au 21 septembre 2020. Les plans d'architectes des bâtiments d'enseignement ont la même disposition et l'un d'entre eux est représenté sur la *Figure 1*. Ils sont orientés vers le nord-ouest et le sud-est et abritent deux sections : l'enseignement et l'atelier. La section d'enseignement à deux étages est destinée aux salles de TD et aux bureaux. L'atelier se compose principalement de salles de TP et d'ateliers équipés de machines mécaniques. Le bâtiment administratif comporte trois étages et comprend principalement des bureaux, deux amphithéâtres et quelques salles de TD.

Chaque salle a une surface différente qui peut être classée en trois catégories : petite surface < 50 m², moyenne 50 m² < surface < 150 m² et grande surface > 150 m². Les quatre bâtiments sont ventilés naturellement, et le climat intérieur en été est contrôlé par l'ouverture et la fermeture manuelle des fenêtres. Ils sont occupés de 08h00 à 19h00 en semaine.



FIGURE 1. Plan du bâtiment d'enseignement

Deux types de capteurs sur les six installés sur le site sont utilisés. Elsys et Class'Air sont des capteurs climatiques qui mesurent la température de l'air intérieur (T_{in}), l'humidité relative (HR_{in}) et le niveau de dioxyde de carbone (CO_2). Les différences entre les deux capteurs sont qu'Elsys détecte également les mouvements et a été programmée pour envoyer des données toutes les dix minutes alors que Class'Air envoie des données toutes les heures. La température extérieure (T_{ext}), l'humidité extérieure (HR_{ext}) et le rayonnement solaire global proviennent de la station météorologique représentative la plus proche, Nîmes Courbessac de Météo France. À partir du rayonnement solaire global, nous calculons le rayonnement solaire vertical (RS) transmis par les vitrages en fonction de leur orientation. La *Figure 2* illustre des comparaisons de différents niveaux de chaque paramètre en fonction de T_{in} de 31 pièces sélectionnées. Le panel a) indique que les données de température intérieure sont normalement distribuées. Le panel d) montre que la

température intérieure moyenne est approximativement de 28 °C, et que l'intervalle de confiance relatif aux pièces de grande surface est plus large que celui des autres. Le panel g) est un résumé des influences conjointes des paramètres d'étage, d'orientation et de type sur la température intérieure. Les salles de TP ont la température la plus élevée dans la plupart des cas. Par conséquent, une salle TP dans chaque bâtiment d'enseignement est utilisée pour construire le modèle qui vise à prédire les températures intérieures horaires.

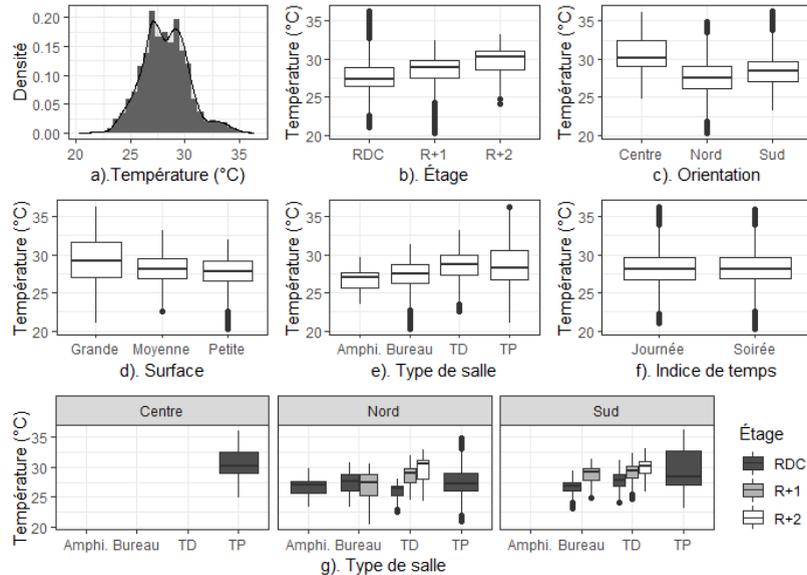


FIGURE 2. Analyse de la température intérieure en fonction de : a) la distribution des données, b) l'étage de la salle, c) l'orientation de la salle (Centre - pièce sans mur exposé à l'environnement extérieur, Nord - salle orientée vers le Nord-Est, Sud - salle orientée vers le Sud-Ouest), d) la surface de la salle, e) le type de salle, f) l'indice temporel (journée – 8h00h à 19h00, soirée - 19h00 à 8h00), g) combinaison des trois facteurs des salles de chaque étage par type et par orientation.

III. RÉSULTATS ET DISCUSSION

27 prédicteurs x au-dessous sont présélectionnés pour le traitement des données afin de calculer la variable de réponse Y .

- x_1 : Indice de jour (jour de la semaine, week-end et vacances)
- x_2 : Indice temporel (journée, soirée)
- x_3 : Orientation (nord, sud)
- x_4 : Surface (grande, petite)
- x_5 : HR_{ext} (%)
- x_6 : HR_{in} (%)
- $x_7, x_8 \rightarrow x_{13}$: $T_{ext}, T_{ext_2} \rightarrow T_{ext_7}$ (retard de 2 à 7 heures de T_{ext})
- $x_{14}, x_{15} \rightarrow x_{20}$: $CO_2, CO_{2_2} \rightarrow CO_{2_7}$ (retard de 2 à 7 heures de CO_2)
- $x_{21}, x_{22} \rightarrow x_{27}$: $RS, RS_2 \rightarrow RS_7$ (retard de 2 à 7 heures de RS)

En appliquant l'équation (1) à l'ensemble des données actuelles, le modèle initial peut être écrit comme suit :

$$Y = \beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2} + \dots + \beta_{27} x_{27} \tag{7}$$

L'analyse est effectuée à l'aide du paquet "Caret" dans Rstudio version 4.0.3. (2020-10-10). Une validation croisée 10-fold est utilisée sur la méthode ascendante pour contrôler le processus d'entraînement du modèle. Un modèle avec le RMSE le plus bas de 0,079 et le R² le plus élevé de 0,79 est systématiquement proposé. Ce modèle comprend 15 prédicteurs et est appelé « *Modèle 1* »:

$$Y = 0,29 + 0,63x_7 + 0,45x_{13} - 0,38x_6 + 0,33x_5 - 0,18x_3 - 0,09x_1 + 0,06x_{15} + 0,06x_2 - 0,06x_{27} - 0,06x_{20} + 0,05x_{14} - 0,05x_{12} - 0,05x_{17} + 0,02x_4 - 0,02x_{24} \tag{8}$$

Modèle 1 est appliqué à *training set* et les hypothèses de régression linéaire sont vérifiées. Le modèle est ensuite appliqué à *testing set* si les hypothèses ne sont pas violées. Si la valeur R² reste supérieure à 70%, le modèle est appliqué à *validating set*. La *Figure 3* est un graphique de réponse de la réponse Y (valeur mesurée) par rapport aux valeurs ajustées \hat{Y} (valeur prédite). Plus les points de données sont proches autour de la ligne bleu (ligne de régression), meilleur est l'ajustement des données. Dans les panneaux a) et b) de la *Figure 3*, la distance entre les points et la ligne de régression, augmente avec la diminution de la température. Cela démontre que lorsque la température mesurée est inférieure à 24 °C, le *Modèle 1* ne prédit pas de manière fiable le T_{in} . Le R² de *Modèle 1* est de 82%, 82%, 89% pour le *training set*, le *testing set* et le *validating set*, respectivement. Un deuxième modèle est composé de neuf prédicteurs, appelé par la suite "*Modèle 2*". Ce modèle peut être exprimé par l'équation ci-dessous.

$$Y = 0,3 + 0,65x_7 + 0,47x_{13} - 0,4x_6 + 0,34x_5 - 0,19x_3 - 0,08x_1 + 0,08x_{12} - 0,07x_{27} + 0,05x_2 \tag{9}$$

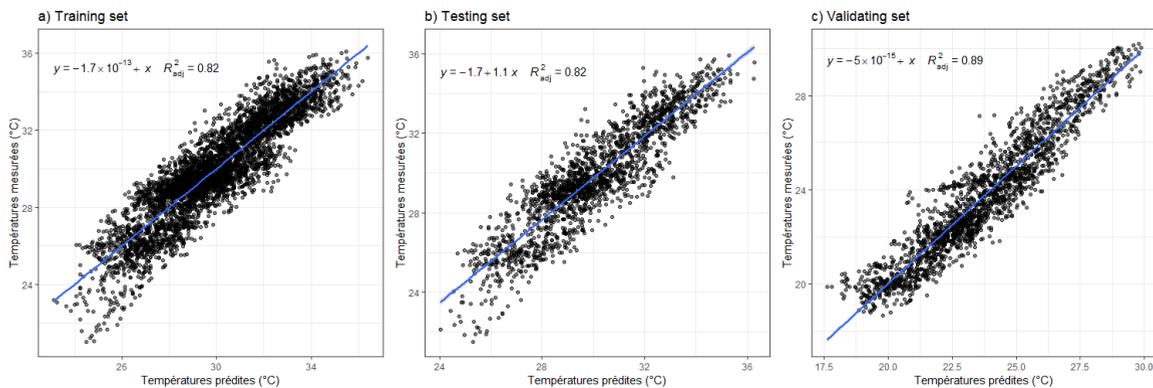


FIGURE 3. La corrélation entre la température intérieure mesurée et prédite à l'aide du *Modèle1* sur : a) *Training set*, b) *Testing set*, c) *Validating set*

La différence de performance entre le *Modèle 1* et le *Modèle 2* est minime, sauf pour le *validating set* où le R² passe de 89% à 81%. Un autre modèle supprime les variables prédicteurs dont les coefficients du *Modèle1* sont inférieurs à 0,1, ce qui donne une estimation de la température intérieure à partir de cinq variables, "*Modèle 3*" s'exprime alors par:

$$Y = 0,29 + 0,61x_7 - 0,42x_6 + 0,4x_5 + 0,34x_{13} - 0,18x_3 \tag{10}$$

La capacité du *Modèle 3* à prédire la valeur a diminué pour le *training set* et le *testing set*, le R^2 passant de 82% à 76%. Le *validating set*, au contraire, est resté plutôt stable à 80%. Étant donné que le taux de précision du modèle reste acceptable, car il est supérieur à 70% lors de la réduction des variables prédictes de 15 à 5, il est plausible que ce modèle soit relativement robuste dans sa capacité de prédiction.

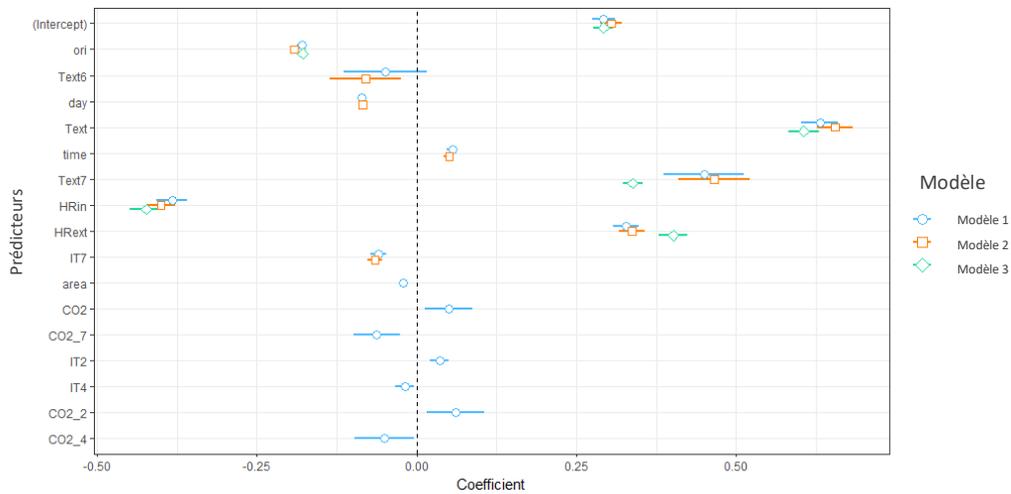


FIGURE 4. Comparaison entre les coefficients des variables prédictes entre le *Modèle 1*, le *Modèle 2* et le *Modèle 3*. La ligne de tiret définit la corrélation négative et positive des variables lorsqu'elles sont respectivement à gauche et à droite. Plus le point de données est éloigné de la ligne de tiret, plus la variable est influente dans le modèle

La *Figure 4* présente une comparaison des coefficients de variation entre les trois modèles. Plus la distance par rapport à la ligne de tiret est grande, plus le prédictes influe sur le modèle. En général, on constate que les variables basées sur le T_{ext} a une corrélation positive, tandis que celles basées sur le CO_2 et de le RS ont des corrélations négatives. Le niveau de CO_2 renseigne sur la concentration d'usagers dans la salle pour un même taux de renouvellement d'air et une même hauteur sous plafond. Il explique également que le coefficient du taux d'occupation dans les salles de TP est faible par rapport à celui de la taille, c'est pourquoi il n'a que peu ou pas d'influence sur le modèle. Le T_{ext} est le prédictes le plus important dans tous les modèles.

De plus, la combinaison de tous les coefficients du T_{ext} temporisée ($x_8 \rightarrow x_{13}$) avec le coefficient du T_{ext} exprime l'inertie thermique du bâtiment, qui est d'environ trois heures en arrière. Cela suggère qu'environ trois heures après une augmentation ou une diminution de la température extérieure, l'effet sur la température intérieure atteint son maximum. Le *panel b*) de la *Figure 5* confirme cette explication : le minimum de T_{ext} est à 05h00 et le minimum de T_{in} est vers 08h00.

Table 1. Erreurs et corrélations statistiques des modèles

Modèle	MAE (°C)	MSE (°C)	MAPE (%)	RMSE (°C)	R ² (%)
Modèle 1	0,89	1,26	0,032	1,12	0,82
Modèle 2	0,9	1,3	0,032	1,14	0,81
Modèle 3	1,04	1,65	0,035	1,28	0,76

Le *Tableau 1* ci-dessus présente quatre matrices d'erreurs et coefficients de détermination pour évaluer les trois modèles sur la période de collecte des données d'entraînement et de test. Comme prévu, les résultats obtenus avec le modèle à 5 entrées sont légèrement moins bons que les autres, lorsque les coefficients des variables inférieures à 0,1 sont également inclus. Néanmoins, les résultats acquis sont très satisfaisants. Mis à part le R² le plus faible, le pourcentage d'erreurs calculées par MAPE est essentiellement le même par rapport aux autres. Les grands avantages d'avoir un petit nombre de paramètres dans le modèle sont que les résultats peuvent être interprétés plus facilement, et que la corrélation des paramètres est moins complexe, rendant les inférences fiables.

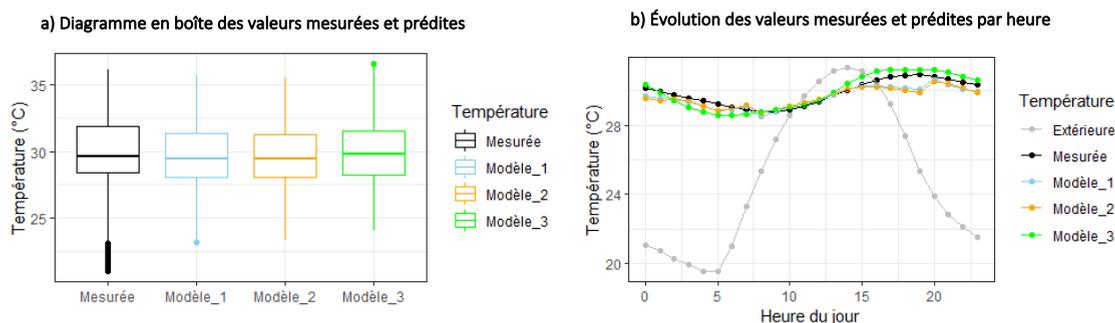


FIGURE 5. Comparaison des valeurs mesurées avec les valeurs prédites des trois modèles en a) Diagramme en boîtes, b) évolution de la température moyenne groupée par heure du jour

La *Figure 5* est une comparaison des températures mesurées avec les valeurs prédites des trois modèles. Le panel a) montre le faible écart (inférieur à 0,5°C) entre les valeurs moyennes de température mesurées et prédites des trois modèles. Néanmoins, les valeurs minimales des données mesurées sont inférieures de plus de 2 °C à celles des données prédites. Dans le panel b), chaque point de données représente la valeur horaire moyenne pour une période de trois mois. Il est évident que le *Modèle 1* et le *Modèle 2* présentent des tendances de progression identiques, le plus souvent abruptes, mais proches des valeurs réelles. En revanche, la tendance du *Modèle 3* est lisse par rapport aux données mesurées, bien que les valeurs minimales et maximales soient respectivement sous-estimées et surestimées. Cette figure montre également les écarts typiques entre les mesures de chaque modèle à chaque heure de la journée. L'imprécision moyenne maximale du *Modèle 3*, qui est inférieure à 1 °C, se produit à 05h00 du matin.

IV. CONCLUSION

L'objectif de cet article est d'identifier les variables prédictives influentes et d'explorer la robustesse d'un modèle de régression linéaire multiple qui vise à prévoir la température intérieure horaire dans les bâtiments scolaires existants. L'entraînement du modèle est basé sur 27 prédicteurs tels que l'indice de jour, l'indice temporel, l'orientation de la salle, la surface de la salle, l'humidité extérieure, l'humidité intérieure, la température extérieure, le CO_2 , le rayonnement solaire et leur décalage horaire de 2 à 7 heures. Trois modèles ont été conçus et comparés en utilisant uniquement les pièces présentant les conditions les plus similaires. Les résultats expliquent pourquoi il existe une différence statistiquement significative entre le *Modèle 1* et le *Modèle 3*, qui comportent respectivement 15 et 5 entrées. Malgré la suppression de 7 prédicteurs, la statistique entre le *Modèle 1* et le *Modèle 2* n'est pas significative. Le *Modèle 1* est globalement plus performant que le *Modèle 3* avec un R^2 de 82% contre 76%. Le modèle final est cependant proposé en tant que *Modèle 3* car ses performances sont satisfaisantes et le taux de précision est acceptable. De plus, ce modèle qui ne contient que 5 prédicteurs, est souhaité pour une structure simple, une interprétation correcte et une corrélation directe des paramètres.

Par rapport aux études précédentes comme Moretti et al. ($R^2_{adj} = 0,92$), Özbalta et al. ($R^2 = 0,87$) Pérez-Lombard et al. ($R^2 = 0,83$), le modèle proposé est apparemment le plus faible. Cependant, cette étude a révélé que l'orientation de la salle a une forte corrélation avec la variation de la température et que le rayonnement solaire n'est pas aussi important que les autres prédicteurs. Les variables les plus influentes sont la température extérieure, l'hygrométrie intérieure et extérieure, la température extérieure décalée de 7 heures et de l'orientation de la salle.

La température la plus basse et la plus haute du modèle proposé sont respectivement sous-estimées et surestimées d'environ $1^\circ C$ en moyenne. Le taux d'inexactitude est de 3%.

La possibilité d'utiliser ce modèle dans d'autres salles aux conditions variées est une autre restriction. La nature distincte de chaque salle rend difficile le développement d'un modèle pour les bâtiments scolaires réels. Dans cette étude de cas, chaque section de bâtiment différait en termes de hauteur, de nombre d'étages, de tailles, d'équipements et d'utilisation du bâtiment. Par conséquent, d'autres salles pourraient avoir des taux de précision de prédiction nettement inférieurs en utilisant le modèle proposé dans cet article. D'autres types de pièces (salles de cours et bureaux) ont en effet des fenêtres plus grandes que celles des salles TP, ce qui augmente le rayonnement solaire. Le taux d'occupation est également plus élevé que dans les salles TP, de sorte que, si les fenêtres restent fermées, le niveau de CO_2 peut alors devenir un prédicteur du modèle de température intérieure (du fait d'apports intérieurs élevés avec un renouvellement d'air faible).

La recherche démontre que pour les bâtiments universitaires ventilés naturellement, le modèle proposé offre une précision de prédiction de 76%. Par conséquent, la régression linéaire multiple est une méthode de prédiction simple, rapide et efficace à des fins de description, d'explication et d'inférence.

REFERENCES

- Afram, A., Janabi-Sharifi, F., 2015. Black-box modeling of residential HVAC system and comparison of gray-box and black-box modeling methods. *Energy Build.* 94, 121–149. <https://doi.org/10.1016/j.enbuild.2015.02.045>
- Afroz, Z., Shafiullah, G., Urmee, T., Higgins, G., 2018. Modeling techniques used in building HVAC control systems: A review. *Renew. Sustain. Energy Rev.* 83, 64–84. <https://doi.org/10.1016/j.rser.2017.10.044>
- Attoue, N., Shahrour, I., Younes, R., 2018. Smart Building: Use of the Artificial Neural Network Approach for Indoor Temperature Forecasting. *Energies* 11, 395. <https://doi.org/10.3390/en11020395>
- Deb, C., Zhang, F., Yang, J., Lee, S.E., Shah, K.W., 2017. A review on time series forecasting techniques for building energy consumption. *Renew. Sustain. Energy Rev.* 74, 902–924. <https://doi.org/10.1016/j.rser.2017.02.085>
- Deshayes, P., 2012. Le secteur du bâtiment face aux enjeux du développement durable : logiques d'innovation et/ou problématiques du changement, in: *Innovations*, n°37. De Boeck Supérieur, Louvain-la-Neuve, pp. 219–236.
- Elmaz, F., Eyckerman, R., Casteels, W., Latré, S., Hellinckx, P., 2021. CNN-LSTM architecture for predictive indoor temperature modeling. *Build. Environ.* 206, 108327. <https://doi.org/10.1016/j.buildenv.2021.108327>
- Moretti, E., Proietti, M.G., Stamponi, E., 2021. A multiple Linear Regression Model to predict indoor temperature trend in historic buildings for book conservation: the case study of “Sala del Dottorato” in Palazzo Murena, Italy. *J. Phys. Conf. Ser.* 2069, 012142. <https://doi.org/10.1088/1742-6596/2069/1/012142>
- Olive, D.J., 2017. Multiple Linear Regression, in: Olive, D.J. (Ed.), *Linear Regression*. Springer International Publishing, Cham, pp. 17–83. https://doi.org/10.1007/978-3-319-55252-1_2
- Özbalta, T.G., Sezer, A., Yıldız, Y., 2012. Models for Prediction of Daily Mean Indoor Temperature and Relative Humidity: Education Building in Izmir, Turkey. *Indoor Built Environ.* 21, 772–781. <https://doi.org/10.1177/1420326X11422163>
- Peel, M.C., Finlayson, B.L., McMahon, T.A., 2007. Updated world map of the Köppen-Geiger climate classification. *Hydrol. Earth Syst. Sci.* 11, 1633–1644. <https://doi.org/10.5194/hess-11-1633-2007>
- Pérez-Lombard, L., Ortiz, J., Pout, C., 2008. A review on buildings energy consumption information. *Energy Build.* 40, 394–398. <https://doi.org/10.1016/j.enbuild.2007.03.007>
- Ramadan, L., Shahrour, I., Mroueh, H., Chehade, F.H., 2021. Use of Machine Learning Methods for Indoor Temperature Forecasting. *Future Internet* 13, 242. <https://doi.org/10.3390/fi13100242>
- Sidler, O., 2012. *La rénovation thermique des bâtiments en France Enjeux et stratégie* (Company report). ENERTECH INGENIEURS CONSEILS.
- Tham, K.W., Ullah, M.B., 1993. Building energy performance and thermal comfort in Singapore, in: *Scopus*.
- Wang, S., Ma, Z., 2008. Supervisory and Optimal Control of Building HVAC Systems: A Review. *HVACR Res.* 14, 3–32. <https://doi.org/10.1080/10789669.2008.10390991>