

Prédiction par régression linéaire multiple : application au comportement thermique d'un bâtiment

Makram ABDELLATIF¹, Julien CHAMOIN¹, Jean-Marie NIANGA², Didier DEFER¹

¹ Yncréa, ULR 4515, Laboratoire de Génie Civil et géo-Environnement (LGCgE), Lille, F-59000, France

² Yncréa, ULR 7512, Unité Mécanique de Lille (UML), Lille, F-59000, France

RESUME Il existe différentes façons de prédire le comportement des bâtiments : la simulation thermique dynamique, les méthodes statistiques et les algorithmes d'apprentissage, et les approches hybrides. Cet article s'intéresse à la prédiction de la température intérieure d'une zone faisant partie d'un bâtiment à l'aide d'un modèle de régression linéaire multiple. Le modèle est entraîné avec les données d'un bâtiment mesurées in situ et les données météorologiques de son environnement. Un bâtiment simulé sous le logiciel TRNSYS a été étudié pour développer différents modèles et comparer leurs performances. Les résultats obtenus montrent que la régression linéaire multiple fournit des prévisions de bonne qualité avec un maximum d'erreur absolue moyenne en pourcentage inférieure à 3%. Le but de cette étude est de développer un modèle simple mais performant, qui sera intégré à une plateforme d'aide à la commande du pilotage des bâtiments permettant aux gestionnaires d'anticiper les problèmes d'inconfort thermique et de surconsommation énergétique.

Mots-clés Confort thermique, Algorithme d'apprentissage, Régression linéaire, Prédiction de la température intérieure, Contrôle de chauffage.

I. INTRODUCTION

En Europe, le secteur du bâtiment est responsable d'environ 40% de l'énergie finale totale consommée et 36% des émissions de CO₂, selon la directive 2010/31 de l'Union Européenne. Spécifiquement, les bâtiments tertiaires sont responsables d'environ 40% de toute l'énergie consommée par le secteur du bâtiment (Balaras et al., 2007). La consommation énergétique est répartie sur tout le cycle de vie du bâtiment. Durant la phase d'exploitation, une partie importante de l'énergie consommée par le bâtiment est absorbée par les systèmes de chauffage, ventilation et climatisation (CVC) permettant d'assurer à l'utilisateur un environnement confortable respectant les réglementations en vigueur. Durant les dernières décennies, plusieurs travaux de recherches ont traité l'amélioration des performances et de l'efficacité des systèmes CVC, permettant d'économiser une part considérable d'énergie. De plus, le pilotage optimal est devenu un sujet en vogue, créant une cohésion entre plusieurs disciplines (automatique, informatique, physique, etc.).

Au niveau réglementaire, l'Union Européenne (UE) pousse à l'installation de systèmes de régulation pièce par pièce ou zone par zone, dans le but de maintenir les performances des systèmes et des bâtiments dans le temps. Promettant des garanties de réduction de la consommation énergétique de 10 à 20%, l'utilisation de la régulation en tant que Gestion Technique Centralisée ou en tant que Gestion Technique des Bâtiments (GTB) se démocratise de plus en plus et tend à concerner la plupart des opérations de construction et de rénovation. La régulation vise à piloter

un (dans le cas d'une GTC) ou plusieurs lots (dans le cas d'une GTB), selon plusieurs paramètres. Le plus souvent le pilotage d'un lot (e.g. chauffage) est fait par rapport à l'occupation de la zone en question (via des plannings ou des capteurs de présence), pour réagir au système actuel et aux conditions environnementales.

Dans ce papier, une méthode d'aide à la commande du pilotage des bâtiments est proposée. Cette méthode est basée sur un modèle prédictif de la température intérieure. Ce résumé étendu introduit, dans un premier temps, les différents modèles de prédiction utilisés dans le domaine du bâtiment, puis présente la méthodologie proposée, en se terminant par un exemple de résultats énumérant les perspectives de ce travail.

II. MODELE DE PREDICTION

En général, avant chaque opération de construction ou de rénovation, le bâtiment est modélisé, afin d'évaluer ses performances et de le comparer aux exigences de la réglementation en vigueur. Dans le domaine de la recherche et de l'industrie, il existe différentes méthodes de modélisation des bâtiments permettant l'étude de leur efficacité énergétique. Ces méthodes peuvent être classées comme suit :

- Boîte blanche : étudie les phénomènes physiques et permet de les estimer à un moment et dans un espace donné (e.g. EnergyPlus, etc.) (Berthou, 2013),
- Boîte noire : un modèle basé sur des méthodes mathématiques et statistiques (e.g. apprentissage automatique, chaîne de Markov, etc.) (Berthou, 2013),
- Boîte grise : solution combinant les deux modèles précédents (Berthou, 2013) (e.g. RxCy).

Dans le cadre de ce travail, il a été convenu d'utiliser un modèle boîte noire de type régression linéaire multiple (MLR), pour sa simplicité de mise en place, sa rapidité d'exécution et son adaptabilité à tout type de bâtiment et toute saison de l'année.

III. METHODE

Dans le but de mettre en place un modèle de prédiction de la température intérieure (T_{in}), les données suivantes ont été prises en compte : la puissance de chauffage appelée (Q_{heat}) de la zone étudiée et les paramètres météo comme la température (T_{out}), l'humidité relative (RH), le rayonnement global horizontal (Rad), la pression atmosphérique ($Patm$), l'opacité du ciel ($Opac$) et la vitesse du vent ($WindS$). Le choix de ces données est basé sur les deux : la disponibilité des données et leurs sens physiques. Les coefficients du modèle MLR sont estimés à l'aide de la méthode des moindres carrés (Amiri et al., 2015).

A. Modèle développé

Pour le développement du modèle de prédiction, une zone d'un bâtiment simulé à l'aide du logiciel de STD TRNSYS (Mcdowell et al., 2017) a été étudié. Le pilotage du chauffage de cette zone est fait selon un scénario déterministe (période de non-chauffe : du 01 mai au 01 septembre, horaire de chauffe : entre 8 heure et 18 heure, hors weekend). A partir des résultats de la simulation sur une année (8760 heures) avec un pas de 1 heure, un jeu de données est construit. Celui-ci est divisé en deux parties : (I) Partie entraînement : un ensemble de données utilisées pour ajuster les paramètres du modèle, (II) Partie test : un ensemble de données pour tester les performances du modèle, en comparant les résultats obtenus avec les données initiales. En suivant ce principe, un ensemble de modèles a été mise en place en : variant la taille de la partie d'entraînement du jeu de données,

ajoutant au jeu de données les mesures de l'heure et/ou les heures passées (e.g. la T_{in} mesurée il y a 3 heures, etc.) et normalisant les données.

B. Comparaison

Afin de comparer les différents modèles MLR développés dans le cadre de cette étude, les indices de performances suivants ont été pris en considération :

- R^2 ajusté (R^2_{aj}) : le coefficient de détermination R^2 donne une idée sur la qualité de prédiction du modèle, en mesurant l'adéquation entre les données observées et les données obtenues. L'avantage du R^2 ajusté par rapport au R^2 c'est la prise en compte du nombre de prédicteurs,
- Carré moyen des erreurs (MSE pour Mean Square Error) : c'est la moyenne arithmétique des carrés des écarts entre prévisions du modèle et observations,
- Erreur quadratique moyenne (RMSE pour Root Mean Square Error) : racine carrée du MSE,
- Erreur absolue moyenne en pourcentage (MAPE pour Mean Absolute Percentage Error) : moyenne des écarts en valeur absolue par rapport aux valeurs observées. C'est donc un pourcentage et par conséquent un indicateur pratique de comparaison.

IV. RESULTATS

Initialement, toutes les données à disposition ont été utilisées pour entraîner le modèle, par la suite, un tri des prédicteurs a été fait, suivant la p-value qui donne une idée sur la significativité statistique de la variable en question. L'intervalle de confiance a été fixée à 5% et donc, toutes les variables ayant une p-value inférieure à 0.05 ont été écartées.

En suivant ce principe, les 3 modèles décrits dans le TABLEAU 1 ont été retenus. Chaque modèle a été entraîné avec un jeu de données constitué de 3600 observations, allant du 01 janvier au 12 juillet de la même année. Les MLR ont été testés par la suite sur une période allant du 25 septembre au 29 septembre. La FIGURE 1 présente les valeurs observées (simulation) et les valeurs estimées de la température intérieure de la zone étudiée et montre la performance des trois modèles.

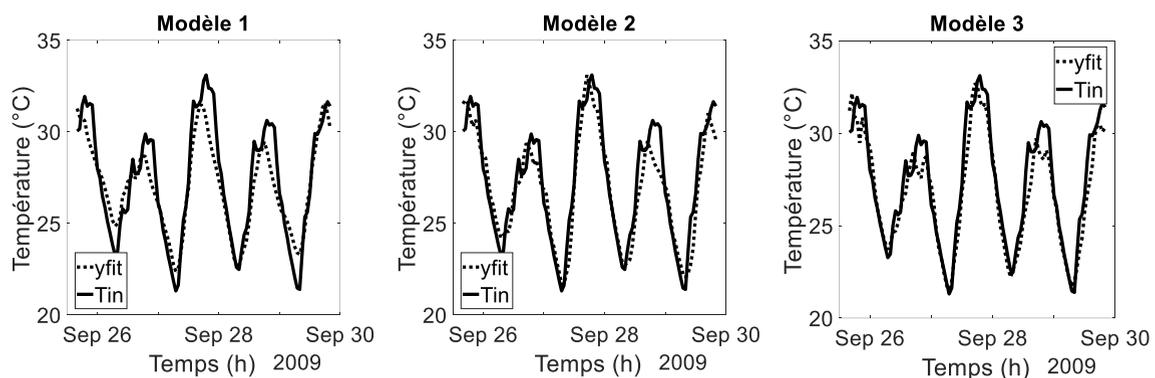


FIGURE 1. Comparaison entre la température intérieure observée (simulation) et estimée avec les modèles MLR 1, 2 et 3

Les résultats obtenus sont prometteurs, et le modèle 3 avec un MAPE de 2.88%, un MSE à 1.12, un RMSE à 1.06 et un R^2_{aj} à 0.96 reste le meilleur estimateur. Les résidus des trois modèles sont compris entre -2 et 2, avec quelques points singuliers sortant de cet intervalle. Il est à noter que les

résidus sont légèrement plus élevés quand la température intérieure est anormalement élevée (supérieure à 30°C).

TABLEAU 1. Description des trois modèles MLR

MLR	Formules	R^2_{aj}	RMSE	MSE	MAPE
1	Tin~Tout+RH+Qheat+Rad+Opac+Constante	0.89	1.29	1.66	3.65
2	Tin~Tout+Qheat+Rad+WindS+Opac+Tin1+Tout1+RH1+Qheat1+Rad1+WindS1+Opac1+Constante	0.94	1.11	1.24	3.33
3	Tin~Tout+RH+Qheat+Rad+Tin1+Tout1+RH1+Qheat1+Rad1+Opac1+Tout2+RH2+Qheat2+Rad2+Opac2+Constante	0.96	1.06	1.12	2.88

V. DISCUSSIONS ET CONCLUSION

Le but de cette étude est de développer un modèle simple permettant à un utilisateur non-expert (e.g. gestionnaire d'un bâtiment) de prédire la température intérieure future d'une zone (ou une pièce) donnée, sans passer par des modèles physiques complexes. Tel que démontré par les résultats ci-dessus, la régression linéaire multiple permet d'avoir des résultats tout à fait acceptables, avec un pourcentage moyen d'erreur absolue inférieur à 3%. Grâce à sa simplicité de mise en place et sa rapidité, le modèle pourra intégrer une plateforme de pilotage, dans le but d'anticiper les éventuels problèmes de confort thermique, tout en corrigeant les consignes de la GTB.

Dans le but d'améliorer la consommation énergétique et le confort thermique, dans les travaux à venir, une zone d'un bâtiment réel situé au nord de Lille (France) sera traitée.

VI. REMERCIEMENTS

Les travaux présentés dans ce papier s'inscrivent dans le cadre de projets de la chaire industrielle Smart Buildings as nodes of Smart Grids (SBnodesSG). Ce travail est financé par les entreprises VINCI Construction France, Projex Groupe Ingénierie et Cegelec.

REFERENCES

- Amiri, S.S., Mottahedi, M., Asadi, S., 2015. Development of multi-linear regression model to predict energy consumption in the early stages of building design. AEI 2015 Birth Life Integr. Build. - Proc. AEI Conf. 2015 85, 54–65. <https://doi.org/10.1061/9780784479070.006>
- Balaras, C.A., Gaglia, A.G., Georgopoulou, E., Mirasgedis, S., Sarafidis, Y., Lalas, D.P., 2007. European residential buildings and empirical assessment of the Hellenic building stock, energy consumption, emissions and potential energy savings. Build. Environ. 42, 1298–1314. <https://doi.org/10.1016/j.buildenv.2005.11.001>
- Berthou, T., 2013. Développement de modèles de bâtiment pour la prévision de charge de climatisation et l'élaboration de stratégies d'optimisation énergétique et d'effacement.
- Mcdowell, T.P., Bradley, D.E., Hiller, M., Lam, J., Merk, J., 2017. TRNSYS 18 : The Continued Evolution of the Software Thermal Energy System Specialists , LLC , Madison , WI , USA Transsolar Energietechnik GmbH , Stuttgart , Germany Centre Scientifique et Technique du Bâtiment , Sophia Antipolis , France Introduction 1922–1930.