

Méthodologie de comparaison des pilotages conventionnel et par apprentissage profond des sous-stations de réseaux de chaleur urbains

Comparison methodology of conventional and machine learning methods for the operation of district heating systems substations

Boris NEROT^{1*}, Julien RAMOUSSE¹, Faiza LOUKIL², Jaume FITÓ-DE-LA-CRUZ¹, Lamia BERRAH², Florine GIRAUD¹, Mickaël BETTINELLI², Alexandre BENOIT², David CORGIER³

¹ LabOratoire proCédés énergIe bâtimEnt

60 rue du lac Léman - 73370 Le Bourget-du-Lac

² Laboratoire d'Informatique, Systèmes, Traitement de l'Information et de la Connaissance

5 chemin de Bellevue - 74940 Annecy

³ MANASLU Ing.

23 Avenue du Lac Léman - 73370 Le Bourget-du-Lac

* (auteur correspondant : boris.nerot@univ-smb.fr)

Résumé - Cet article s'intéresse aux sous-stations de bâtiments résidentiels et tertiaires raccordés à un réseau de chaleur. Il présente une méthodologie de comparaison des pilotages conventionnels et par apprentissage profond. Une réflexion sur les critères de performance des réseaux de chaleur permet d'abord de définir les indicateurs sur lesquels le pilotage des sous-stations a une influence. L'exploitation de ces indicateurs par pilotage conventionnel et apprentissage profond est ensuite présentée. Cette comparaison repose sur 84 scénarios physiques introduits pour leur représentativité des cas réels rencontrés. Finalement, l'exploitation de données réelles en prédition appuie la viabilité d'une approche apprentissage profond appliquée au pilotage des sous-stations.

Abstract -

This article focuses on district heating substations of residential and tertiary buildings. It presents a methodology to compare the current operation method of substations with a machine learning approach. The intersection of performance criteria at the district heating level and potential actions in substations defines a multi objective performance indicator for substation operation. How this indicator is used in each operation method and the material for comparison (more than 80 simulation cases) are introduced. Finally, prediction results from machine learning applied on typical substations data confirm the viability of such an approach for substation operation.

Nomenclature

A	Actions de pilotage
C	Conditions extérieures
<i>P</i>	Puissance, <i>J/s</i>
<i>T</i>	Température, <i>K</i>
W	Paramètres du RNN

Symboles grecs

λ	Coefficient de descente de gradient
Σ	Modèle du système

Ω	Modèle RNN
----------	------------

Indices et exposants

<i>C, F</i>	Chaud, Froid
<i>d</i>	Demande thermique (ECS, chauffage)
<i>obs</i>	Grandeur observée, expérimentale
<i>pred</i>	Grandeur prédictive
<i>prim</i>	Primaire SST
<i>REF</i>	Grandeur référence, simulation

1. Introduction

La production de chauffage et d'eau chaude sanitaire (ECS) représente 75% des consommations énergétiques des logements en France [1]. Les réseaux de chaleur urbains (RCU) sont des solutions appropriées pour répondre à ces demandes dans un contexte de décarbonation.

Leur performance globale est dépendante de l'étape de livraison de la chaleur aux usagers en sous-stations (SST). Or le choix des débits circulant au niveau du bâtiment (secondaire SST), appelé pilotage, a aujourd'hui pour seul objectif la satisfaction de la demande, parfois au détriment de la performance de l'ensemble du système RCU.

Ce pilotage conventionnel pourrait notamment être amélioré par une approche prenant en compte les états récents du système SST. Parmi ces méthodes, l'apprentissage profond (AP) pourrait proposer des actions de pilotage précisément adaptées à chaque SST grâce à une évaluation de la sous-optimalité des actions précédentes.

Récemment, Trabert et al. [2] ont proposé un modèle prédictif permettant d'anticiper les besoins en chaleur d'un consommateur industriel et donc d'optimiser le pilotage du stockage sensible disponible au secondaire de la SST. La mise à profit du stockage au niveau consommateur permet de s'affranchir partiellement de l'utilisation de capacités centralisées de pointe, dont les sources énergétiques sont assez carbonnées. Bien que le pilotage lui-même ne soit pas réalisé par AP, les variables de décision requises à chaque instant par ce pilotage sont prédites par un réseau de neurones récurrent (RNN) à partir d'un historique récent du système. Similairement, l'usage de RNN en appui de modèle physiques traditionnels est mis en avant pour la prédiction des températures chez de gros consommateurs résidentiels. Lu et al. [3] associent par exemple un modèle physique au pas de temps journalier et un prédicteur statistique au pas de temps horaire pour déterminer la température à délivrer dans un réseau de chauffage secondaire couvrant plus de $10^5 m^2$ résidentiel. Zhang et al. [4] traitent une situation de demande similaire en affinant par AP la solution approchée par un modèle physique réduit. Les résultats issus du RNN sont couplés à une fonction objectif et une énumération totale des possibilités afin de sélectionner la juste prédiction parmi plusieurs. Buffa et al. [5] mettent aussi en avant les apports d'un RNN dans le pilotage des SST présentant un moyen de production (RCU de 5ème génération). Ils exploitent un pilotage optimisé par essaims particulaires (PSO) et basé sur une approximation RNN des composants de la SST. Ils soulignent les meilleures performances d'un tel pilotage en comparaison d'un pilotage conventionnel dans le cas typique de la recharge d'un stockage par une pompe à chaleur.

Ces travaux montrent notamment les apports potentiels de l'intégration de l'AP dans des tâches de prédiction au niveau consommateur des RCU. Pour autant, peu d'entre eux confrontent le pilotage entièrement par AP au pilotage conventionnel. Dans ce contexte, cet article présente une méthodologie de comparaison de ces deux modes de pilotage.

Dans un premier temps sont étudiés les indicateurs d'un pilotage performant en SST, du point de vue de l'ensemble du RCU. Puis différents scénarios caractérisant les SST et leur environnement sont définis. S'en suit une discussion de la complexité limitée des modèles physiques, en lien avec les contraintes spécifiques aux deux méthodes de pilotage. Finalement, le pilotage conventionnel d'un de ces modèles et la validation d'une méthode d'AP sur une des architectures étudiée permettent de discuter de la viabilité de l'étude complète à venir.

2. Méthodologie

Cette section explore les principaux aspects du problème du pilotage, en passant par les indicateurs de performance, les scénarios envisagés, la modélisation physique, ainsi que les approches de pilotage conventionnel et par apprentissage profond.

2.1. Indicateurs de performance

L'objectif principal de la livraison de chaleur en SST est la satisfaction de la demande réelle P_d^{REF} , où d peut être le chauffage ou l'ECS (Equation 1). Mais la performance globale du RCU est affectée par d'autres grandeurs. Ainsi, on mesure la différence entre la température retour primaire et une température de référence (Equation 2). Similairement, l'énergie livrée au primaire peut-être comparée à la somme des demandes (Equation 3). Enfin, les variations brutales d'appel de puissance sont dommageable au RCU. Une condition suffisante garantissant une absence de variations est l'obtention d'une monotone lisse à l'échelle annuelle. Tandis que le foisonnement à l'échelle de la ville contribue à atteindre cette situation, l'objectif doit aussi être recherché à l'échelle de la SST. Ainsi, on définit comme quatrième indicateur le ratio des centiles 90% (reflet des valeurs élevées) et 10% (valeurs faibles) par l'équation 4.

$$\epsilon_d(t) = \left(\frac{P_d(t)}{P_d^{REF}(t)} - 1 \right)^2 \quad (1)$$

$$\epsilon_T(t) = \left(\frac{T_{prim}^F(t)}{T_{prim}^{F,REF}} - 1 \right)^2 \quad (2)$$

$$\epsilon_E(t) = \left(\frac{\sum_{k=0}^t P_{prim}(k)}{\sum_{k=0}^t (P_{ECS}(k) + P_{CH}(k))} - 1 \right)^2 \quad (3)$$

$$\epsilon_P(t) = \frac{\text{centile}(\{P_{prim}(k), k \in [0, t]\}, 90\%)}{\text{centile}(\{P_{prim}(k), k \in [0, t]\}, 10\%)} \quad (4)$$

Notons que les équations 1 à 3 présentent une version normalisée des indicateurs.

Ces indicateurs sont évalués à chaque pas de temps t à partir des sorties du modèle physique de SST. Ils sont pondérés et sommés en une erreur totale L à minimiser, laquelle est transmise à l'algorithme de pilotage. Celui-ci en déduit des consignes de débit à l'instant $t + 1$ (Figure 1).

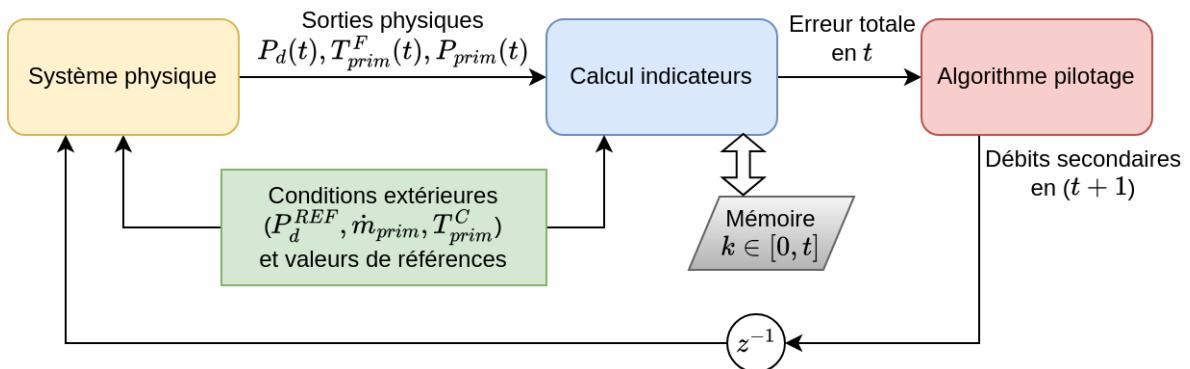


Figure 1 : Communication entre modèle physique et modèle de pilotage via la remontée d'indicateurs de performance

2.2. Scénarios

Dans cette section sont introduits les 84 scénarios envisagés comme support pour la comparaison des méthodes de pilotage. Chaque scénario comporte un aspect architecture (arrangement des composants chauffage et ECS en SST), un aspect dimensionnement (caractéristiques de ces composants) et des conditions extérieures supposées connues et indépendantes de l'opération du système.

2.2.1. Architectures

Sont considérées 4 architectures SST. L'architecture « chauffage seul » présente un échangeur de chaleur avec le RCU et le circuit de distribution de chauffage dans les logements, au secondaire. L'arrangement chauffage et ECS avec stockage aval décrit une seconde architecture qui couvre la demande ECS via un échangeur puis un ballon de stockage, le tout en parallèle du circuit chauffage (Figure 2a). Similairement, le positionnement du stockage ECS en amont de l'échangeur ECS dans l'arrangement parallèle définit une 3^{ème} variation. La dernière consiste en un montage cascade des échangeurs ECS (stockage aval) puis chauffage (Figure 2b).

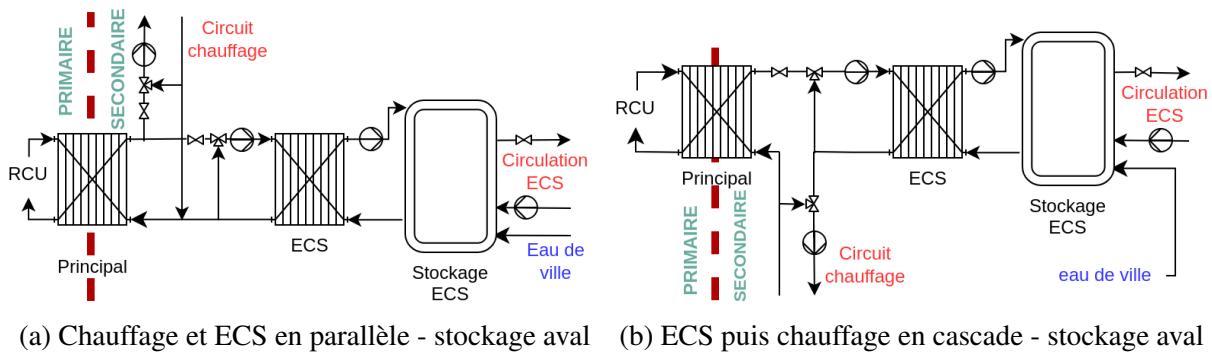


Figure 2 : Schémas hydrauliques de 2 des 4 architectures étudiées

2.2.2. Dimensionnements

Pour les 3 architectures assurant l'ECS, le volume du ballon peut prendre 2 valeurs. La première est le dimensionnement préconisé par l'ADEME [6], fonction notamment de la demande ECS. La seconde valeur vaut 130% de cette préconisation. Le dimensionnement de l'échangeur en puissance suit ces mêmes 2 cas (préconisation ADEME, 130% de la préconisation), sauf pour l'architecture avec stockage ECS en amont pour laquelle le dimensionnement est imposé par la demande ECS maximale.

2.2.3. Conditions extérieures

Les demandes thermiques sont caractérisées du point de vue de la puissance appelée et de la performance des émetteurs chauffage.

Concernant le premier point, on considère une taille unique de bâtiment (en terme de surface chauffée, $2,5 \times 10^3 \text{ m}^2$) mais 3 cas d'usage : résidentiel, tertiaire (sans ECS) et mixte.

La performance des émetteurs chauffage dépend de leur surface d'échange : émetteurs type plancher chauffant (surface élevée, faible température) ou radiateurs conventionnels (surface faible, température élevée) définissent 2 cas.

Enfin, la température entrante au primaire de l'échangeur RCU évolue linéairement avec la

température extérieure (loi d'eau) selon un coefficient faible ou élevé, soient 2 cas supplémentaires.

2.3. Modélisation physique

Le système physique thermique à l'étude comprend la sous-station et les réseaux de distribution de chauffage et d'ECS au secondaire du bâtiment.

A chaque pas de temps, l'algorithme de pilotage appelle au moins une fois la réalisation d'un pas de temps de la simulation thermique (plusieurs appels dans le cadre du pilotage conventionnel, voir Section 2.4., Equation 7). Par ailleurs, 84 simulations annuelles, autant que le nombre de scénarios, sont à réaliser. Enfin, un faible pas de temps (typiquement, $\Delta t = 5$ min) est requis pour saisir la dynamique du stockage ECS.

Pour ces raisons, la complexité algorithmique du modèle physique doit être maîtrisée. Du point de vue thermique, les effets inertiels dus au stockage sensible de la chaleur par les conduites, échangeurs et émetteurs sont ignorés. Du point de vue hydraulique, le pilotage est réalisé par débits imposés dans chaque branche, dans le respect des mélanges imposés par les vannes 3 voies mais sans prise en compte des pertes de charge. Également, la faible taille de bâtiment et le pas de temps au moins de l'ordre de la minute permettent de supposer immédiate la propagation du front de pression. Le stockage ECS est modélisé en une dimension par $n - 2$ zones intermédiaires de volume constant (thermocline) tandis que les deux volumes extrêmes sont de volume variable [8]. Les échangeurs eau-eau utilisent la méthode du NUT pour déterminer deux températures inconnues connaissant les débits, la surface et le coefficient d'échange.

2.4. Pilotage conventionnel

Dans un système réel, les demandes exactes en puissance à chaque pas de temps sont inconnues. En particulier pour la demande en chauffage, les SST sont pilotées hydrauliquement en ajustant vitesse de pompe et ouvertures de vannes. Cela permet l'adaptation aux variations de pression induites par le réglage des robinets de radiateurs par les usagers.

L'algorithme de pilotage conventionnel présenté ici suppose lui la demande connue car les aspects hydrauliques ne sont pas modélisés (voir section 2.3.).

Soit Σ la fonction décrivant le système physique et transformant les actions de pilotage \mathbf{A} en l'erreur totale L à chaque instant t (Equation 5). On appelle \mathbf{C} les conditions extérieures (demandes thermiques, température aller et débit au primaire de la SST) dont dépend également Σ .

$$L(t) = \Sigma(\mathbf{A}(t), \mathbf{C}(t)) \quad (5)$$

L'objectif des deux pilotages est de déterminer $\mathbf{A}(t + 1)$ qui minimise $L(t + 1)$ (Equation 6). Les conditions extérieures en $t + 1$ n'étant pas connues, on les suppose égales à celles en t .

$$\mathbf{A}(t + 1) = \operatorname{argmin}_{\mathbf{A}} (\Sigma |_{\mathbf{C}(t)}) \quad (6)$$

On cherche $\mathbf{A}(t + 1)$ par une descente de gradient (Equation 7), laquelle fait intervenir le coefficient λ .

$$\mathbf{A}(t + 1) = \mathbf{A}(t) - \lambda \cdot \nabla \Sigma |_{\mathbf{A}(t), \mathbf{C}(t)} \quad (7)$$

Le gradient en $\mathbf{A}(t)$ est approché par différenciation de Σ au voisinage de $\mathbf{A}(t)$. Contrairement à une résolution algébrique exacte, cette approche permet de traiter le cas général où le

modèle Σ peut contenir des équations implicites elle-mêmes résolues par un algorithme d'optimisation, comme c'est le cas dans la relation reliant la différence de température logarithmique moyenne et la puissance thermique échangée.

2.5. Pilotage AP

Le problème du pilotage fait intervenir en paramètre et en variables d'état du système des séries temporelles. Les modèles d'AP privilégiés pour saisir l'information portée par ces séries durant les instants passés sont des RNN. En particulier, les modèles LSTM [7] et XGBoost [9] sont envisagés et à implémenter via le framework TensorFlow.

Dans le cas général, le paramétrage $\mathbf{W}(t)$ du RNN $\Omega_{\mathbf{W}}$ est mis à jour à partir de L , également par un algorithme de descente du gradient. $\mathbf{W}(t)$ est une image du système Σ forgée sur les instants précédents et permet de fournir des actions $\mathbf{A}(t+1)$ à partir des j dernières conditions extérieures données en entrée (Equation 8). Cette relation suppose $\mathbf{C}(t+1)$ inconnue, comme pour le pilotage conventionnel.

$$\mathbf{A}(t+1) = \Omega_{\mathbf{W}(t)}(\{\mathbf{C}(t-i), i \in [0, j-1]\}) \quad (8)$$

3. Résultats

Le déploiement d'un algorithme de prédiction mène à questionner les données typiquement disponibles dans le contexte des RCU, ainsi que le sens physique qu'elles portent. Cette partie présente donc un cas d'étude de prédiction de la température intervenant dans l'indicateur $\epsilon_T(t)$ et introduit le potentiel des RNN dans une optique de pilotage futur.

3.1. Données

Les données utilisées sont issues de l'instrumentation d'une des sous-stations de l'ensemble résidentiel Arlequin, à Grenoble. Les données couvrent la période du 18 novembre 2016 au 28 juillet 2017. Bien qu'au pas minute, la majorité des valeurs sont constantes sur 2 à 5 minutes consécutives.

La SST en question est typiquement du type présenté en Figure 2a. Elle assure les besoins annuels en ECS (53 MWh, incluant pertes transfert, stockage et circulation) et chauffage du 18 novembre au 1er juin (354 MWh). Les données monitorées comprennent les débits, températures et puissances thermiques et électriques (circulations) sur les circuits chauffage et ECS (dont stockage), les données météorologiques (température, vent, irradiation), la température aller et le débit au primaire. Parmi les données chauffage et ECS manquantes, on note 2 périodes assez longues : du 6 décembre au 15 janvier et du 11 mars au 4 avril. Au total, 157 jours cumulés de mesure sont réellement exploitables.

L'objectif de l'apprentissage profond est de prédire la température retour au primaire de l'échangeur principal à partir des 43 autres variables explicatives. Ces variables sont transformées par Analyse en Composantes Principales (ACP) vers une nouvelle base de dimension 15 afin de réduire la complexité algorithmique associée à leur traitement tout en conservant leur potentiel d'explicabilité (95% de la variance est conservée).

3.2. Structure modèle et paramétrage

Le modèle utilisé est un réseau de neurones de mémoire à long terme (LSTM) [7], qui est un type de réseau de neurones récurrent conçu pour capturer et utiliser efficacement les

dépendances dans les données séquentielles.

La contrainte de continuité temporelle impose une période de test du modèle à la fin du jeu de données, i.e. du 14 juin au 28 juillet 2017 (20% du volume total), tandis que son entraînement est réalisé sur la période précédente (80% du volume total). Notons qu'aucune demande en chauffage n'est assurée durant la période test, contrairement à la période d'entraînement qui couvre chauffage et ECS.

Le modèle retourne, pour chacun des n pas de temps t , son estimation de la température retour $T_{prim}^{F,pred}(t)$. Cette valeur est comparée à celle observée ($T_{prim}^{F,obs}(t)$) selon les métriques *Root Mean Squared Error* (RMSE) (Equation 9) et *Mean Absolute Error* (MAE) (Equation 10).

$$RMSE = \sqrt{\frac{1}{n} \sum_t \left(T_{prim}^{F,pred}(t) - T_{prim}^{F,obs}(t) \right)^2} \quad (9)$$

$$MAE = \frac{1}{n} \sum_t \left| T_{prim}^{F,pred}(t) - T_{prim}^{F,obs}(t) \right| \quad (10)$$

3.3. Résultats

La confrontation des valeurs prédites et observées sur le jeu de données test est donnée par la Figure 3a (toute la période de test) et la Figure 3b (2 journées consécutives du mois de juin).

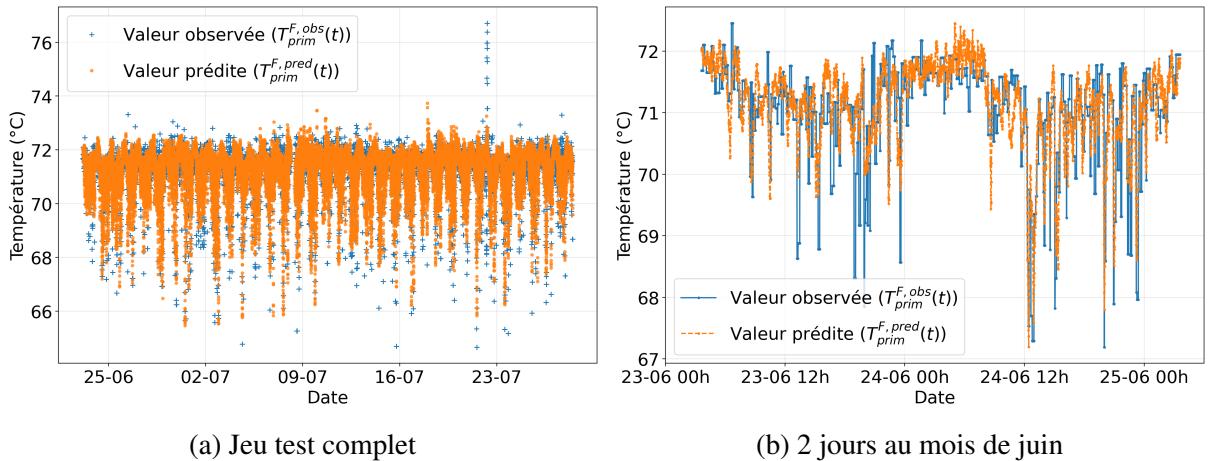


Figure 3 : Valeur prédite et mesurée de la température retour primaire

La Figure 3b montre que les variations de température quotidiennes sont bien saisies par le modèle, mais que celui-ci passe à côté de certains changements brutaux.

Les valeurs des métriques, $MAE = 0,512^\circ\text{C}$ et $RMSE = 0,724^\circ\text{C}$, traduisent une faible erreur moyenne entre les prédictions du modèle et les données réelles. Du point de vue physique, le fait que la température retour dépende de la température aller et que celle-ci soit assez stable en été simplifie le problème de prédiction. Mais la seule demande sur cette période est la demande ECS, dont la prévisibilité est faible et la satisfaction déphasée par l'usage du stockage, ce qui rend le problème difficile. Dans ce contexte, l'écart typique d'environ $0,5^\circ\text{C}$ (MAE) dénote donc une assez bonne performance.

4. Conclusion

Les méthodes conventionnelles de pilotage de SST ont actuellement pour seul objectif la satisfaction des demandes en chauffage et ECS. Cet article envisage une performance multi objectifs par pilotage de SST par AP. Une méthodologie exposant 84 cas d'étude physiques (architectures SST, dimensionnement, conditions extérieures) et le lien entre performance passée et actions de pilotage futures est présentée. La viabilité des RNN pour traiter en AP les problématiques propres aux RCU est ensuite confirmée par des résultats en prédition sur des données réelles.

La suite du travail portera sur la définition des modèles physiques et du paramétrage associés aux différents scénarios. Les premiers efforts de modélisation concernent le scénario le plus rencontré dans les systèmes réels, à savoir un montage chauffage/ECS en parallèle avec stockage ECS aval. En termes de difficulté à venir, les efforts de modélisation portent sur les échangeurs eau-eau, les émetteurs chauffage dans les logements et le stockage ECS. Le principe même du stockage ECS et de sa discréétisation en volume causent une rupture de la continuité du modèle physique et de l'erreur L , ce qui complique certains modes de pilotages conventionnels.

En termes d'outils de modélisation, les auteurs utilisent autant que possible des langages et bibliothèques *opensource* en prévision de la diffusion des résultats. La modélisation du système physique est réalisée avec Matlab Simulink (fermé, propriétaire) mais est exportée puis utilisée sans dépendance sous forme de code C. Les algorithmes de pilotage (conventionnels et AP) sont écrits en Python. La mise en oeuvre du pilotage par AP nécessitera des étapes d'évaluation et de comparaison des performances à partir des modèles entraînés. L'analyse comparative finale des modes de pilotage sera garantie par la distinction entre modèle physique et modèle de pilotage. L'exhaustivité dans les situations physiques et mode de pilotage couverts ainsi que le libre-accès des conclusions et du matériel requis pour reproduire l'étude contribuent à rendre fiable cette comparaison entre intelligence artificielle et méthode conventionnelle.

Références

- [1] Fedene, Enquête annuelle des réseaux de chaleur et de froid (2023)
- [2] U. Trabert, F. Pag, J. Orozaliev, U. Jordan, and K. Vajen, Peak shaving at system level with a large district heating substation using deep learning forecasting models, *Energy*, vol. 301 (2024)
- [3] Y. Lu, X. Peng, C. Li, Z. Tian, J. Niu, and C. Liang, A baseline model combining physics and data-driven approach for operation evaluation of district heating substation, *Energy and Buildings*, vol. 321 (2024)
- [4] Z. Zhang, X. Zhou, H. Du, and P. Cui, A new model predictive control approach integrating physical and data-driven modelling for improved energy performance of district heating substations, *Energy and Buildings*, vol. 301 (2023)
- [5] S. Buffa, A. Soppelsa, M. Pipiciello, G. Henze, and R. Fedrizzi, Fifth-Generation District Heating and Cooling Substations : Demand Response with Artificial Neural Network-Based Model Predictive Control, *Energies*, vol. 13 (2020)
- [6] ADEME, Le dimensionnement des systèmes de production d'eau chaude sanitaire en habitat individuel et collectif (2019)
- [7] S. Hochreiter and J. Schmidhuber, Long Short-Term Memory, *Neural Comput.*, vol. 9 (1997)
- [8] K. M. Powell and T. F. Edgar, An adaptive-grid model for dynamic simulation of thermocline thermal energy storage systems, *Energy Conversion and Management*, vol. 76 (2013)
- [9] T. Chen and C. Guestrin, XGBoost : A Scalable Tree Boosting System, *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016

Remerciements

Ce travail a été financé dans le cadre du projet SHINE associé au contrat ANR-22-EXES-0017.