

Thèse : Hybridation données/connaissances pour la prédiction de la consommation électrique à différentes échelles spatiales de la ville.

Mots clés : load forecasting, data science, urban building energy modeling, electrical grid

Résumé : Maintenir la stabilité du réseau électrique dans un contexte de forte pénétration de sources renouvelables est un challenge qui nécessite des modèles de prédiction. La prédiction de la demande électrique à une échelle locale est sans doute la prédiction la plus complexe à obtenir. Elle peut s'appuyer sur une approche basée sur les données mais les données sont peu nombreuses et ces méthodes sont peu efficaces sur des situations rares qui sont justement celles qui peuvent déséquilibrer le réseau. Cette thèse propose donc une approche hybride basée sur les données et la connaissance, dans un contexte où ces dernières sont variées mais peu nombreuses et hétérogènes.

Thématique :

Science des données. Prévion de la demande d'électricité.

Domaine :

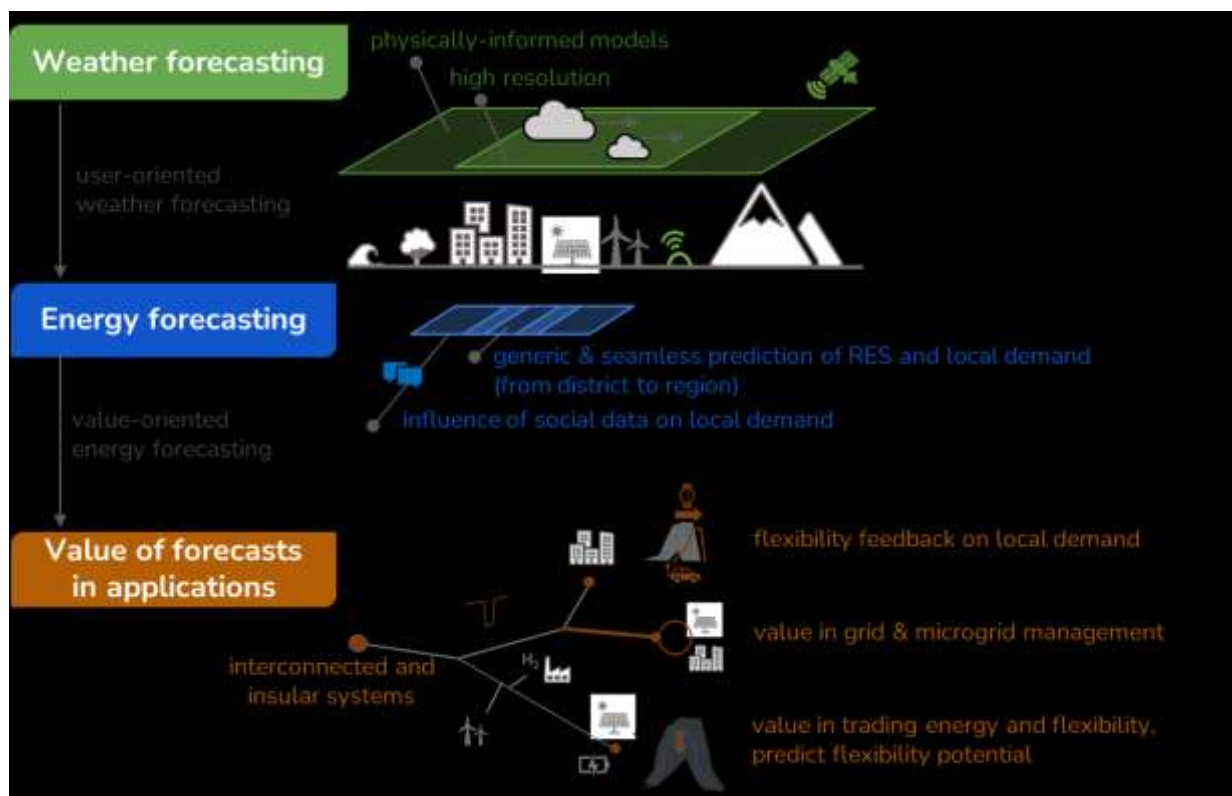
Ecole doctorale EEATS, spécialité Génie Electrique

Objectifs :

La transition énergétique nécessite l'intégration d'une quantité importante d'énergie renouvelable dans le réseau électrique. Cette énergie est intermittente et contribue à l'instabilité du réseau dont il est crucial de maintenir l'équilibre à chaque instant. Cet équilibre entre production et consommation est possible par le pilotage de la demande, dont il est nécessaire d'anticiper le comportement à différentes échelles spatiales et temporelles.

L'objectif de la thèse est donc de développer des nouvelles techniques de prédiction de la demande d'électricité en prenant en compte un grand nombre de données hétérogènes disponibles.

Ces données peuvent être des historiques de consommation, des prédictions météorologiques, ainsi que des données issues des réseaux sociaux, mais aussi des connaissances sur les caractéristiques des bâtiments et du réseau de distribution.



Contexte global du projet Fine4Cast

Contexte :

Les travaux de thèse seront réalisés dans le cadre du PEPR TASE dont l'objectif est le développement de réseaux d'énergie flexibles et résilients incluant une part significative d'énergies renouvelables. Ils se situent plus précisément dans l'action de recherche Fine4Cast qui vise à développer les nouvelles générations d'outils de prévision de demande d'énergie et de production renouvelable à des échelles spatiale et temporelle plus fines que celles qui sont couramment utilisées.

Les travaux se dérouleront à Grenoble au laboratoire G2Elab. Dans le consortium du projet le G2Elab / Grenoble INP – UGA ainsi que le Centre PERSEE des MINES PARIS – PSL sont spécialisés dans le domaine de l'énergie électrique. Les autres partenaires sont plus spécialisés sur les modèles météorologiques en lien avec Météo-France pour la prédiction des énergies renouvelables.

La ville de Grenoble sera choisie comme cas d'étude. Un partenariat avec le Gestionnaire du Réseau de Distribution (GRD) de la ville a été réalisé nous permettant d'avoir accès aux données topologiques du réseau ainsi qu'aux valeurs énergétiques réelles mesurées à un niveau agrégé.

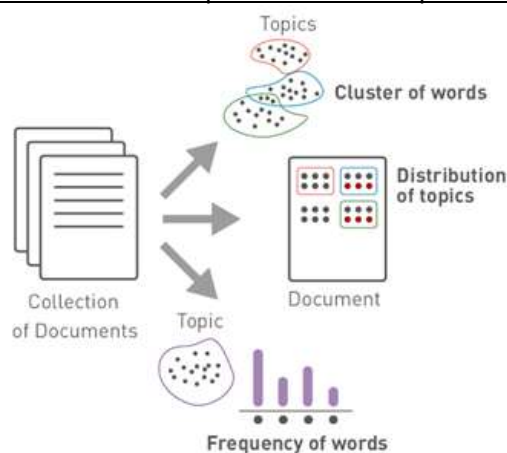
Méthode :

Le travail commencera par une revue de la littérature sur la modélisation de la demande électrique à de l'échelle bâtiment à l'échelle quartier/ville (UBEM : urban building energy modeling). Les différentes méthodologies existantes dont les approches TopDown – BottomUp [1], les archetypes [2], mais aussi les approches hybrides s'appuyant à la fois sur des connaissances descriptives associées à des modèles physiques simplifiés, et les techniques d'apprentissage automatique à partir des données historiques et autres variables explicatives [3].

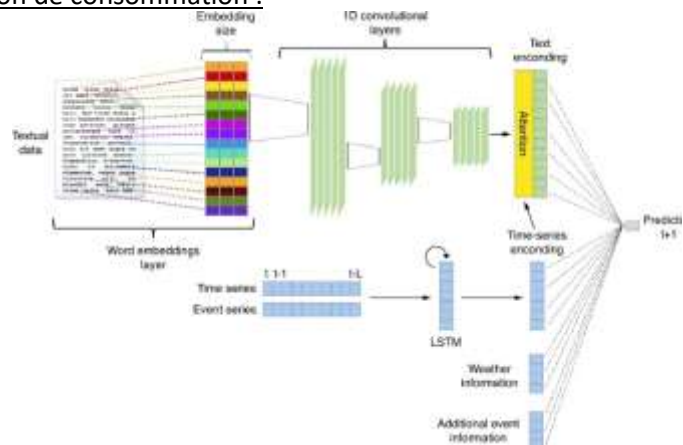
Les données disponibles pour le cas d'étude seront collectées puis analysées et une confrontation entre les méthodes de modélisation et les données disponibles conduira à définir la stratégie de modélisation. Afin de permettre une approche répliquable et de pallier au manque de données, celle-ci pourra s'appuyer sur l'apprentissage par transfert [4], l'adaptation de domaine [5], ou encore l'augmentation de données [6].

Cette stratégie devrait en particulier traiter de la fusion de données hétérogènes [7], et un focus particulier sera fait sur la fusion de séries temporelles à des données textuelles issue du web [8], afin d'étudier si les résultats que nous avons publié à l'échelle de France [9] peuvent s'appliquer à une échelle locale comme la ville de Grenoble.

Utilisation du NLP pour améliorer la prédiction de consommation :



topic extraction: cibler les bons documents



Plongement lexical (word embedding)

Mécanisme d'attention : Apprendre automatiquement des motifs ou des séquences de mots qui se rapportent aux données de séries temporelles observées

Résultats attendus :

L'indicateur principal est la justesse de la prédiction, mais il sera associé à la quantité / qualité des données ou informations nécessaires pour construire le modèle. De même, la facilité / accessibilité de mise en œuvre de la méthode sera considérée.

La méthodologie développée constituera le livrable principal. Elle devra être argumentée, testée et critiquée. Un article sera soumis dans une revue scientifique internationale avec comité de lecture. Un code python accompagné d'un tutorial de type notebook seront mis en accès ouvert.

Références bibliographiques :

- [1] YAKUT, M. Z., & Sinem, E. S. E. N. (2023). A comprehensive survey of the urban building energy modeling (UBEM) process and approaches. *International Journal of Energy Studies*, 8(1), 87-116.
- [2] Nayak, B. K., Sansaniwal, S. K., Mathur, J., Chandra, T., Garg, V., & Gupta, R. (2023). A review of residential building archetypes and their applications to study building energy consumption. *Architectural Science Review*, 66(3), 187-200.
- [3] Li, Z., Ma, J., Tan, Y., Guo, C., & Li, X. (2023). Combining physical approaches with deep learning techniques for urban building energy modeling: A comprehensive review and future research prospects. *Building and Environment*, 110960.
- [4] Luo, T., Tang, Z., Liu, J., & Zhou, B. (2023, April). A Review of Transfer Learning Approaches for Load, Solar and Wind Power Predictions. In *2023 Panda Forum on Power and Energy (PandaFPE)* (pp. 1580-1584). IEEE.
- [5] Teng, S. Y., van Nooten, C. C., van Doorn, J. M., Ottenbros, A., Huijbregts, M. A. J., & Jansen, J. J. (2023). Near real-time predictions of renewable electricity production at substation level via domain adaptation zero-shot learning in sequence. *Renewable and Sustainable Energy Reviews*, 186, 113662.
- [6] Iglesias, G., Talavera, E., González-Prieto, Á., Mozo, A., & Gómez-Canaval, S. (2023). Data augmentation techniques in time series domain: a survey and taxonomy. *Neural Computing and Applications*, 35(14), 10123-10145.
- [7] Fadhel, M. A., Duham, A. M., Saihood, A., Sewify, A., Al-Hamadani, M. N., Albahri, A. S., ... & Gu, Y. (2024). Comprehensive Systematic Review of Information Fusion Methods in Smart Cities and Urban Environments. *Information Fusion*, 102317.
- [8] Rodrigues, F., Markou, I., & Pereira, F. C. (2019). Combining time-series and textual data for taxi demand prediction in event areas: A deep learning approach. *Information Fusion*, 49, 120-129.
- [9] A. Alyafi, P. Cauchois, B. Delinchant and A. Berges, "Time series machine learning augmented with social network events to improve national electricity consumption profile estimation," 27th International Conference on Electricity Distribution (CIRED 2023), Rome, Italy, 2023, pp. 2283-2288, <https://doi.org/10.1049/icp.2023.1222>

Conditions scientifiques matérielles (conditions de sécurité spécifiques) et financières du projet de recherche

Salaire de 2 200 euros brut mensuel (augmentation à 2300€ brut au 1er janvier 2026). Accueil au laboratoire, tout matériel et logiciel fournis.

Objectifs de valorisation des travaux de recherche du doctorant : Participation à au moins un congrès international et valorisation des travaux en open-source.

Profil et compétences recherchées

Le ou la candidate doit être de formation scientifique disposant d'un master de recherche dans le domaine de la modélisation. Il ou elle doit connaître le domaine du génie électrique ou de la physique des systèmes énergétiques, la modélisation mathématique, statistique et numérique, et disposer de compétences solides en informatique, notamment en programmation python avec une connaissance des bibliothèques classiques de science des données. Il ou elle doit avoir le goût pour la recherche et l'innovation, l'investissement personnel dans un sujet qu'il ou elle portera.

Contact : benoit.delinchant@G2ELab.grenoble-inp.fr