Thesis: Data/knowledge hybridization for the prediction of electricity consumption at different spatial scales in the city.

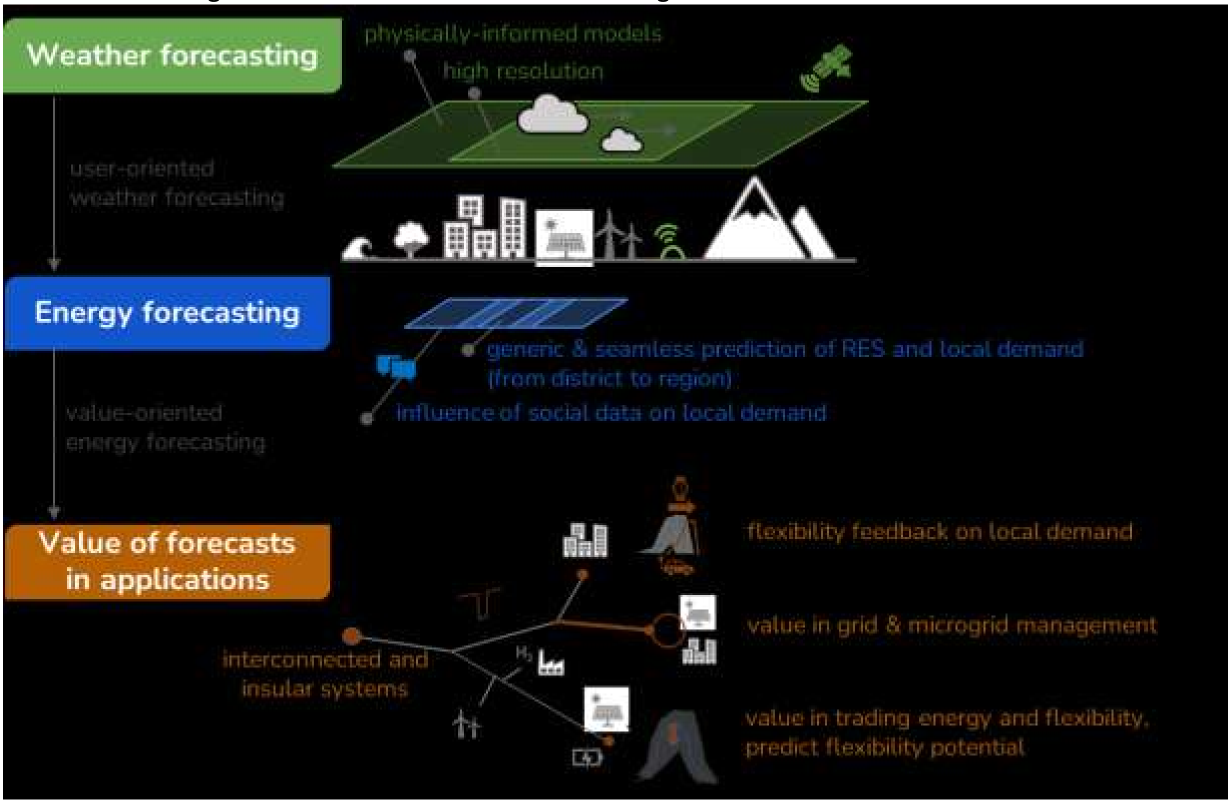
Keywords: load forecasting, data science, energy modelling of urban buildings, electricity grid.

Abstract: Maintaining the stability of the electricity network in a context of high penetration of renewable sources is a challenge that requires prediction models. Forecasting electricity demand on a local scale is undoubtedly the most complex to obtain. It can be based on a data-driven approach, but data is scarce and these methods are not very effective on rare situations, which are precisely those that can unbalance the network. This thesis therefore proposes a hybrid approach based on data and knowledge, in a context where the latter are varied but few and heterogeneous.

Theme: Data science. Electricity demand forecasting.

Field: EEATS doctoral school, specialising in Electrical Engineering

Objectives:
The energy transition requires the integration of a significant amount of renewable energy into the electricity grid. This energy is intermittent and contributes to the instability of the grid, the balance of which must be maintained at all times. This balance between production and consumption is achieved by controlling demand, the behaviour of which needs to be anticipated on different spatial and temporal scales.
The aim of this thesis is therefore to develop new techniques for predicting electricity demand by taking into account the large amount of heterogeneous data available.
These data can be consumption histories, weather predictions, as well as data from social networks, but also knowledge about the characteristics of buildings and the distribution network.



Contexte global du projet Fine4Cast

Context:

The thesis work will be carried out as part of the PEPR TASE programme, the aim of which is to develop flexible and resilient energy networks that include a significant proportion of renewable energies. More specifically, it is part of the Fine4Cast research action, which aims to develop new generations of tools for forecasting energy demand and renewable production on spatial and temporal scales that are finer than those currently used.

The work will be carried out at the G2Elab laboratory in Grenoble. In the project consortium, G2ELab / Grenoble INP - UGA and the PERSEE Centre of MINES PARIS - PSL specialise in the field of electrical energy. The other partners are more specialised in meteorological models in conjunction with Météo-France for the prediction of renewable energies.

The city of Grenoble will be chosen as a case study. A partnership with the city's Distribution System Operator (DSO) has been set up, giving us access to network topology data as well as actual energy values measured at an aggregate level.

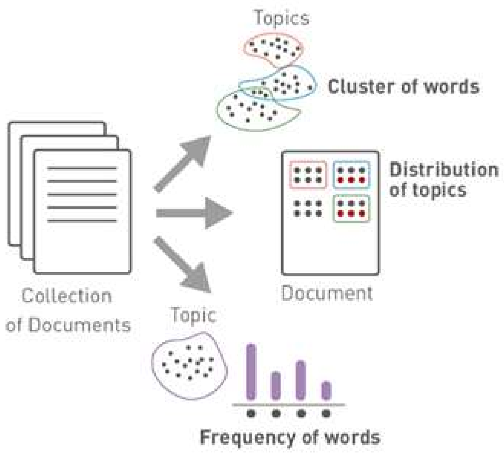
Method:

The work will begin with a review of the literature on the modelling of electrical demand at the building to district/city scale (UBEM: urban building energy modelling). The various existing methodologies include TopDown - BottomUp approaches [1], archetypes [2], as well as hybrid approaches based on both descriptive knowledge associated with simplified physical models, and machine learning techniques based on historical data and other explanatory variables [3].

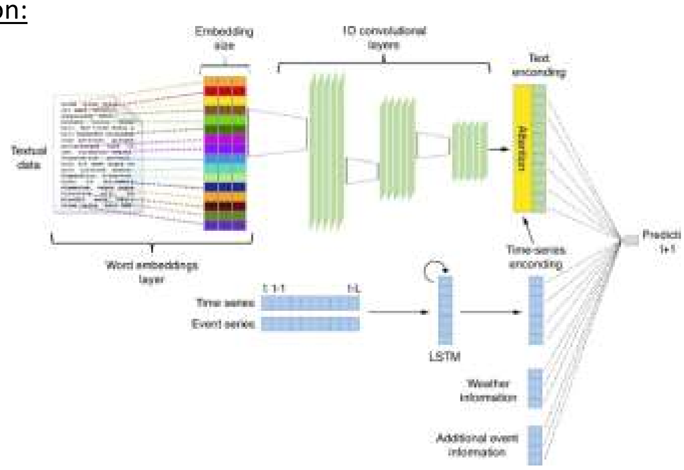
The data available for the case study will be collected and analysed, and a comparison between the modelling methods and the available data will lead to the definition of the modelling strategy. In order to enable a replicable approach and to compensate for the lack of data, this strategy could be based on transfer learning [4], domain adaptation [5], or data augmentation [6].

This strategy should deal in particular with the fusion of heterogeneous data [7], and a particular focus will be placed on the fusion of time series with textual data from the web [8], in order to study whether the results that we published at country scale for France [9] can be applied to a local scale such as the city of Grenoble.

<u>Using NLP to improve consumption prediction:</u>



topic extraction: target the right documents



attention mechanism to join NLP and time series

Expected results:

The main indicator is the accuracy of the prediction, but this will be associated with the quantity/quality of the data or information required to build the model. Similarly, the ease/accessibility of implementing the method will be considered.

The methodology developed will be the main deliverable. It will have to be argued, tested and criticized. At least one article will be submitted to a peer-reviewed international scientific journal. Python code and a notebook tutorial will be made available as open science.

Litterature references :

[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., Duhaim, 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

Scientific, material (specific safety conditions) and financial conditions of the research project :
Salary of €2,200 gross per month (rising to €2,300 gross on 1 January 2026). Office in the laboratory, all hardware and software provided.

Promoting the doctoral student's research work: PhD student will participate to at least one international conference. His developments will be published in open-source.

Profile and skills required : The candidate must have a scientific background and a research master's degree in modelling. He or she should be familiar with electrical engineering or energy systems physics, mathematical, statistical and numerical modelling, and have solid IT skills, particularly in Python programming with a knowledge of classic data science libraries. He or she must have a taste for research and innovation, and a personal investment in a subject that he or she will lead.

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