

Titre de la thèse : Méthodes d'apprentissage automatique pour la prédiction du microclimat urbain

Mots clés : changement climatique, mécanique des fluides, machine learning.

Une version anglais est en deuxième partie du document / An english version is in the second part of the document

I-Résumé du projet de thèse

1. Contexte et objectifs

L'urbanisation accélérée et l'îlot de chaleur urbain augmentent le stress thermique des populations, un phénomène aggravé par le changement climatique (IPCC, 2022). En Europe du Sud, où le vieillissement démographique renforce les vulnérabilités, l'adaptation urbaine devient une priorité. D'ici 2070, les villes du sud de la France pourraient connaître plus de 30 jours de vagues de chaleur annuelles, avec des températures atteignant 50 °C.

La modélisation climatique urbaine est un levier fondamental pour anticiper ces impacts et concevoir des stratégies d'atténuation efficaces. Cependant, les approches existantes, basées sur la résolution numérique des équations modélisant la dynamique et la thermodynamique des écoulements atmosphériques, présentent des limites : si les modèles à méso-échelle permettent une simulation des processus urbains à grande échelle (Redon et al., 2020), ils simplifient excessivement la morphologie urbaine et l'effet des végétaux. À l'inverse, les modèles à micro-échelle, plus précis (Salim et al., 2015 ; Krc et al., 2021), restent limités par leur coût computationnel élevé, rendant difficile leur application opérationnelle à grande échelle. Dans cette perspective, ce projet s'inscrit dans la dynamique des recherches récentes visant à développer des modèles opérationnels, alliant précision et rapidité de calcul grâce à l'intégration de méthodes d'apprentissage automatique (Buizza et al., 2022).

L'objectif de ce projet de thèse est ainsi de renforcer les capacités de simulation du climat urbain à l'échelle de la rue et d'appliquer ces outils numériques à l'évaluation des scénarios de verdissement, en prenant Lyon comme terrain d'étude. Les autorités locales ont un besoin crucial d'éclairages quant au potentiel de rafraîchissement des stratégies d'adaptation basées sur la végétalisation. Le projet répond aux questions scientifiques liées à l'adaptation des territoires au changement climatique et d'outils de modélisation pour la gestion des risques climatiques en milieu urbain. Il visera à développer un modèle opérationnel optimisé, intégrant des algorithmes d'apprentissage automatique et des données de terrain, afin de simuler avec précision l'exposition thermique au niveau de l'agglomération et d'évaluer les leviers d'atténuation les plus efficaces.

2. Verrous scientifiques

Les simulations atmosphériques urbaines actuelles reposent sur diverses approches de modélisation basées sur la résolution des équations de dynamique des fluides (Maronga et al., 2020). Pour simuler avec précision les conditions microclimatiques, ces modèles doivent intégrer des modules multiphysiques prenant en compte les échanges radiatifs entre les surfaces urbaines ainsi que les échanges évapotranspiratoires et radiatifs des végétaux, ce qui constitue un défi majeur. La simulation des effets des arbres dans ces modèles représente notamment un défi particulier : la résistance au vent des feuilles doit être représentée par une force de traînée (Salim et al., 2015), tandis que les effets sur les échanges radiatifs sont modélisés via un facteur d'atténuation dans des algorithmes spécifiques (Krc et al., 2021).

En raison de leur coût computationnel, ces simulations numériques haute fidélité, sont limitées à de petits domaines et à des conditions météorologiques spécifiques. Elles ne sont donc pas adaptées aux simulations en temps réel ni à l'évaluation des scénarios de

verdissement à l'échelle d'une agglomération. Pour contourner cette limitation, des modèles opérationnels sont nécessaires. Conçus pour être accessibles aux non-experts, ces modèles doivent (i) nécessiter bien moins de ressources de calcul et (ii) fournir des valeurs horaires pour des variables telles que la température de l'air, l'humidité, la vitesse du vent et le rayonnement au niveau de la rue, en capturant les principales variations spatiales (Di Sabatino et al., 2013). Ces variables sont essentielles pour calculer des indicateurs de confort thermique, comme l'Universal Thermal Climate Index (UTCI) (Blazejczyk et al., 2012). Pour y parvenir, il est crucial de simplifier la géométrie urbaine et de paramétrer les dynamiques urbaines, en intégrant potentiellement des données de terrain via l'assimilation de données (Nguyen et Soulhac, 2021) ou l'apprentissage automatique (Buizza et al., 2022). La paramétrisation dépendra de l'utilisation du modèle—qu'il s'agisse de simulations en temps réel (gestion des risques sanitaires), d'analyses de scénarios (études épidémiologiques) ou d'optimisation de la planification urbaine. Si plusieurs modèles opérationnels existent pour la qualité de l'air (Soulhac et al., 2011) et le bruit (Gabillet, 1990), très peu sont dédiés à la simulation du climat urbain à l'échelle de la rue. Pour y parvenir, nous proposons dans ce projet de thèse d'intégrer au modèle opérationnel de nouveaux algorithmes d'Intelligence Artificielle (IA) pour la prédiction du stress thermique urbain, dont les verrous scientifiques à relever concernent notamment l'intégration de modalités de différents types et échelles ainsi que la nécessaire frugalité des données pour l'apprentissage des modèles.

II-Contexte de travail

Etablissement et laboratoires

Cette thèse se déroulera à l'Ecole Centrale de Lyon, un établissement public à caractère scientifique, culturel et professionnel (EPCSCP). Centrale Lyon forme des ingénieurs généralistes, des ingénieurs de spécialité, des étudiants en master et des docteurs. L'établissement accueille au total près de 3 000 étudiants, et dispose d'environ 500 personnels, dont 200 enseignants et enseignants-chercheurs. Il est caractérisé par une recherche reconnue à l'international, adossée à 6 laboratoires de recherche, tous Unités Mixtes de Recherche CNRS, mêlant activités fondamentales et appliquées, en particulier au travers de nombreux contrats industriels.

Du fait de son caractère pluridisciplinaire de la thèse (Mécanique des fluides / Machine Learning), elle se déroulera au sein de deux laboratoires :

- Laboratoire de Mécanique des Fluides et Acoustique (**LMFA**), UMR 5509
- Laboratoire d'InfoRmatique en Image et Systèmes d'information (**LIRIS**), UMR 5205

Le LMFA développe un continuum de recherches en mécanique des fluides et en acoustique, depuis la compréhension des phénomènes physiques et leur modélisation jusqu'aux recherches finalisées en partenariat avec les industriels et les organismes publics.

Quant aux recherches menées au sein du LIRIS, elles visent à relever les défis du monde numérique, notamment ceux posés par l'intelligence artificielle (IA), l'analyse de données volumineuses (Big Data), la vision par ordinateur, la cyber-sécurité, la transformation digitale ou l'apprentissage humain. Une partie des activités du LIRIS se situent aux interfaces des sciences humaines et sociales, de l'ingénierie, de la médecine, des sciences de la vie et des sciences de l'environnement.

La supervision de la thèse sera assurée par trois enseignants-chercheurs :

- **Pietro Salizzoni**, Professeur, Centrale Lyon, LMFA, pietro.salizzoni@ec-lyon.fr
- **Lionel Soulhac**, Professeur, INSA Lyon, LMFA, lionel.soulhac@insa-lyon.fr
- **Emmanuel Dellandrea**, Maître de Conférences, Centrale Lyon, LIRIS, emmanuel.dellandrea@ec-lyon.fr

Financement de la thèse : Bouquet de thèses 2025

Le sujet de thèse de doctorat décrit ci-après s'inscrit au sein d'un bouquet de thèses dont le but est de construire une approche scientifique pluridisciplinaire pour aborder l'enjeu sociétal « Société numérique responsable » et plus précisément, la thématique spécifique « Data et IA dans une démarche durable et responsable », identifiée comme un enjeu prioritaire par les 4 établissements du Collège d'Ingénierie Lyon Saint-Etienne (Centrale Lyon, ENTPE, INSA Lyon, Mines Saint-Étienne) et par l'Université Jean Monnet Saint-Étienne, qui soutiennent financièrement les thèses formant ce bouquet 2025.

Le bouquet de thèses 2025 regroupe 6 thèses qui couvrent différentes facettes de la science des données et de l'intelligence artificielle en abordant les questions suivantes :

- Monitoring des procédés de cristallisation par émission acoustique assistée par IA
- Conception assistée par IA de biopolymères biodégradables et/ou biosourcés pour la protection durable des cultures agricoles
- Méthodes d'apprentissage automatique pour la prédiction du microclimat urbain
- Identification non-linéaire de structures pilotée par des données
- Inférence et explicabilité en mode confidentiel : vers l'autodiagnostic via l'image
- Vers une certification de la surveillance vibratoire par IA explicative

Ces thèses rassemblent au total 16 encadrants rattachés à 11 laboratoires du site Lyon Saint-Etienne (Centre d'Innovation en Télécommunications et Intégration, Centre SPIN – Génie des Procédés, Ingénierie des Matériaux Polymères, Biologie Fonctionnelle, Insectes et Interactions, Institut Camille Jordan, Laboratoire de Mécanique des Fluides et d'Acoustique, Laboratoire de Tribologie et Dynamique des Systèmes, Laboratoire d'InfoRmatique en Image et Systèmes d'information, Laboratoire Hubert Curien, Laboratoire Vibrations Acoustique, Matériaux : Ingénierie & Science) dont les 5 établissements financeurs sont tutelles. Les 6 doctorants recrutés au titre de ce bouquet seront inscrits dans 3 Écoles Doctorales du site : MEGA, EDML, SIS.

Les équipes (doctorants et leurs encadrants) impliquées dans ces 6 thèses forment une communauté scientifique pluridisciplinaire : des échanges réguliers entre ces équipes se dérouleront tout au long des 3 années du parcours doctoral, notamment sous la forme de séminaires communs permettant de développer l'approche systémique pluridisciplinaire propre au bouquet et d'enrichir les compétences disciplinaires des équipes dans un esprit de partage et d'apprentissage. Les mémoires de thèses produits à l'issue du parcours doctoral reflèteront également le positionnement original des travaux de thèse au sein d'un bouquet en incluant un chapitre qui développera l'analyse de l'impact des travaux réalisés sur l'enjeu « Data et IA dans une démarche durable et responsable ».

III-Pré-requis du poste

Profil recherché

Tout.e candidat.e titulaire d'un diplôme d'ingénieur ou d'un Master, avec un excellent dossier académique, de cursus spécialisé en Informatique avec de fortes compétences en apprentissage automatique et une grande appétence pour la recherche pluridisciplinaire, notamment en lien avec la mécanique des fluides.

Processus de recrutement

Les candidat·es intéressé·e·s, ou souhaitant plus d'information, sont vivement invité·e·s à se manifester en envoyant un mail contenant un CV + un court message de présentation et de motivation à l'équipe d'encadrement (voir adresses ci-dessus).

Les candidatures pourront être faites jusqu'au **27 mai 2025**, sur le site <https://ecolecentraledelyon.recruitee.com/>

Les pièces à joindre au dossier sont le CV, les relevés de notes de Master et une lettre de motivation.

IV-Références bibliographiques

- Blazejczyk, K., et al., International Journal of Biometeorology, 2012, 56, p. 515-535.
Buizza, C. et al., Journal of Computational Science, Volume 58, 2022, 101525.
Di Sabatino, S., Buccolieri, R., Salizzoni, P., 2013, Int. J. of Env. Poll. 2013 52:3-4, 172-191
Ding, Y. et al., Neurocomputing, Volume 501, 2022, P. 246-257.
Dumas, G., 2021, PhD Thesis, <https://dumas.ccsd.cnrs.fr/OMP-TEL/tel-03700032>
Gabillet, Y. 1990, Cahier 2444. CSTB.
Grard, M., Dellandréa, E., Chen, L., Int. J. of Computer Vision, 2020, 128(5), pp. 1331–1359.
Krc, P., Geoscientific Model Development, 2021, 14, p. 3095-3120.
Maronga, B. et al., Geoscientific Model Development, 2020, 13, p. 1335-1372.
Masson, V., Boundary-Layer Meteorology, 2000, 94, p. 357-397.
Nguyen C.-V., Soulhac L., 2021, Atmos. Environ., 253, 118366.
Redon, E., et al., Geoscientific Model Development, 2020, 13, p. 385-399.
Ronneberger, et al., 2015, MICCAI, Springer, LNCS, Vol.9351: 234–241.
Salim, M.H. et al., Journal of Wind Engineering and Industrial Aerodynamics, 144, p. 84-95.
Schoetter, R., Masson, V. et al., Geoscientific Model Development, 2020, 13, p. 5609-5643.
Soulhac, L., Salizzoni, P. et al., 2011, Atmos. Environ., 45, 7379-7395.

ENGLISH Version

PhD thesis title: Machine learning methods for urban microclimate prediction

Keywords : climate change, fluid mechanics, mécanique des fluides, machine learning.

I-Summary of the PhD research project

1. Context and objectives

Accelerated urbanization and the urban heat island effect are increasing thermal stress on populations, a phenomenon exacerbated by climate change (IPCC, 2022). In Southern Europe, where demographic aging heightens vulnerabilities, urban adaptation is becoming a priority. By 2070, cities in southern France could experience more than 30 days of heatwaves annually, with temperatures reaching up to 50 °C.

Urban climate modeling is a fundamental tool for anticipating these impacts and designing effective mitigation strategies. However, existing approaches, which rely on the numerical resolution of equations modeling the dynamics and thermodynamics of atmospheric flows, have limitations: while mesoscale models enable simulation of urban processes on a large scale (Redon et al., 2020), they excessively simplify urban morphology and the effects of vegetation. Conversely, microscale models, which are more precise (Salim et al., 2015; Krc et al., 2021), remain limited by their high computational cost, making their operational application at large scale difficult. In this context, the project aligns with recent research efforts aimed at developing operational models that combine accuracy and computational speed through the integration of machine learning methods (Buizza et al., 2022).

The objective of this PhD project is thus to strengthen the capacity for urban climate simulation at the street scale and to apply these digital tools to the assessment of greening scenarios, using Lyon as a case study. Local authorities have a crucial need for insights into the cooling potential of adaptation strategies based on urban greening. The project addresses scientific questions related to the adaptation of territories to climate change and the development of modeling tools for managing climate risks in urban environments. It aims to develop an optimized operational model, integrating machine learning algorithms and field data, to accurately simulate thermal exposure at the metropolitan scale and to evaluate the most effective mitigation levers.

2. Scientific Challenges

Current urban atmospheric simulations rely on various modeling approaches based on the resolution of fluid dynamics equations (Maronga et al., 2020). To accurately simulate microclimatic conditions, these models must incorporate multiphysics modules that account for radiative exchanges between urban surfaces as well as the evapotranspirative and radiative exchanges of vegetation, which presents a major challenge. Simulating the effects of trees in these models is particularly complex: wind resistance of leaves must be represented by a drag force (Salim et al., 2015), while effects on radiative exchanges are modeled through an attenuation factor in specific algorithms (Krc et al., 2021).

Due to their computational cost, these high-fidelity numerical simulations are limited to small domains and specific meteorological conditions. They are therefore not suitable for real-time simulations or for evaluating greening scenarios at the metropolitan scale. To overcome this limitation, operational models are needed. Designed to be accessible to non-experts, these models must (i) require far fewer computational resources and (ii) provide hourly values for variables such as air temperature, humidity, wind speed, and radiation at street level, capturing the main spatial variations (Di Sabatino et al., 2013). These variables are essential for calculating thermal comfort indices, such as the Universal Thermal Climate Index (UTCI) (Blazejczyk et al., 2012). Achieving this requires simplifying urban geometry and parameterizing urban dynamics, potentially by integrating field data through data assimilation (Nguyen and Soulhac, 2021) or machine learning (Buizza et al., 2022). The parameterization will depend on the model's intended use—whether for real-time simulations (health risk management), scenario analysis (epidemiological studies), or urban planning optimization. While several operational models exist for air quality (Soulhac et al., 2011) and noise

(Gabillet, 1990), very few are dedicated to simulating urban climate at the street scale. To address this, we propose in this PhD project to integrate new Artificial Intelligence (AI) algorithms into the operational model for predicting urban thermal stress, with scientific challenges notably including the integration of modalities of different types and scales as well as the need for data efficiency in model training.

II-Work context

Institution and Laboratories

This PhD will take place at École Centrale de Lyon, a public institution of scientific, cultural, and professional character (EPCSCP). Centrale Lyon trains generalist engineers, specialized engineers, master's students, and PhD candidates. The institution hosts nearly 3,000 students and has around 500 staff members, including 200 faculty and research staff. It is distinguished by internationally recognized research, supported by six research laboratories, all of which are CNRS Joint Research Units, combining fundamental and applied activities, particularly through numerous industrial contracts.

Given the interdisciplinary nature of the thesis (Fluid Mechanics / Machine Learning), it will be carried out within two laboratories:

- Fluid Mechanics and Acoustics Laboratory (LMFA), UMR 5509
- Laboratory of Computer Science in Image and Information Systems (LIRIS), UMR 5205

The LMFA develops a continuum of research in fluid mechanics and acoustics, ranging from the understanding and modeling of physical phenomena to applied research in partnership with industry and public organizations.

Research at LIRIS aims to address the challenges of the digital world, particularly those posed by artificial intelligence (AI), big data analysis, computer vision, cybersecurity, digital transformation, and human learning. Part of LIRIS's activities are at the interface of social sciences, engineering, medicine, life sciences, and environmental sciences.

The supervision of the PhD will be provided by three faculty members :

- **Pietro Salizzoni**, Full Professor, Centrale Lyon, LMFA, pietro.salizzoni@ec-lyon.fr
- **Lionel Soulhac**, Full Professor, INSA Lyon, LMFA, lionel.soulhac@insa-lyon.fr
- **Emmanuel Dellandrea**, Associate Professor, Centrale Lyon, LIRIS, emmanuel.dellandrea@ec-lyon.fr

Founding of the PhD thesis : Bouquet de thèses 2025

The doctoral thesis described above is part of a series of theses designed to build a multidisciplinary scientific approach to the societal challenge of a "responsible digital society", and more specifically, the specific theme of "Data and AI in a sustainable and responsible approach", identified as a priority issue by the 4 institutions of the Lyon Saint-Etienne Engineering College (Centrale Lyon, ENTPE, INSA Lyon, Mines Saint-Étienne) and by the Université Jean Monnet Saint-Étienne, which are providing financial support for the theses making up this 2025 package.

The 2025 theses package includes 6 theses covering different facets of data science and artificial intelligence, addressing the following questions:

- Monitoring crystallization processes using AI-assisted acoustic emission
- AI-assisted design of biodegradable and/or biosourced biopolymers for the sustainable protection of agricultural crops
- Machine learning methods for urban microclimate prediction
- Data-driven non-linear structure identification
- Inference and explicability in confidential mode: towards self-diagnosis via images
- Towards certification of vibration monitoring with explanatory AI

These theses involve a total of 16 supervisors from 11 laboratories on the Lyon Saint-Etienne site (Centre d'Innovation en Télécommunications et Intégration, Centre SPIN - Génie des Procédés, Ingénierie des Matériaux Polymères, Biologie Fonctionnelle, Insectes et Interactions, Institut Camille Jordan, Laboratoire de Mécanique des Fluides et d'Acoustique, Laboratoire de Tribologie et Dynamique des Systèmes, Laboratoire d'InfoRmatique en Image et Systèmes d'information, Laboratoire Hubert Curien, Laboratoire Vibrations Acoustique, Matériaux : Ingénierie & Science) of which the 5 funding institutions are supervisors. The 6 PhD students recruited under this package will be enrolled in 3 Doctoral Schools on the site: MEGA, EDML, SIS.

The teams (doctoral students and their supervisors) involved in these 6 theses form a multi-disciplinary scientific community: regular exchanges between these teams will take place throughout the 3 years of the doctoral pathway, notably in the form of joint seminars to develop the multi-disciplinary systems approach specific to the bouquet and enrich the teams' disciplinary skills in a spirit of sharing and learning. Thesis papers produced at the end of the doctoral program will also reflect the original positioning of the thesis work within a bouquet, by including a chapter analyzing the impact of the work carried out on the "Data and AI in a sustainable and responsible approach" issue.

III-Position Requirements

Desired Profile

Any candidate holding an engineering degree or a Master's degree, with an excellent academic record, specialized in Computer Science with strong skills in machine learning and a strong interest in multidisciplinary research, particularly in connection with fluid mechanics.

Recruitment process

Interested candidates, or those seeking more information, are strongly encouraged to get in touch by sending an email containing a CV and a short cover letter (introducing themselves and outlining their motivation) to the supervisory team (see addresses above).

Applications can be submitted until May 27, 2025, on the website : <https://ecolecentraledelyon.recruitee.com/>

The documents to be included in the application are: a CV, Master's transcripts, and a motivation letter.

IV-References

- Blazejczyk, K., et al., International Journal of Biometeorology, 2012, 56, p. 515-535.
- Buizza, C. et al., Journal of Computational Science, Volume 58, 2022, 101525.
- Di Sabatino, S., Buccolieri, R., Salizzoni, P. , 2013, Int. J. of Env. Poll. 2013 52:3-4, 172-191
- Ding, Y. et al., Neurocomputing, Volume 501, 2022, P. 246-257.
- Dumas, G., 2021, PhD Thesis, <https://dumas.ccsd.cnrs.fr/OMP-TEL/tel-03700032>
- Gabillet., Y. 1990, Cahier 2444. CSTB.
- Grard, M., Dellandréa, E., Chen, L., Int. J. of Computer Vision, 2020, 128(5), pp. 1331–1359.
- Krc, P., Geoscientific Model Development, 2021, 14, p. 3095-3120.
- Maronga, B. et al., Geoscientific Model Development, 2020, 13, p. 1335-1372.
- Masson, V., Boundary-Layer Meteorology, 2000, 94, p. 357-397.
- Nguyen C.-V., Soulhac L., 2021, Atmos. Environ., 253, 118366.
- Redon, E., et al., Geoscientific Model Development, 2020, 13, p. 385-399.
- Ronneberger, et al., 2015, MICCAI, Springer, LNCS, Vol.9351: 234–241.
- Salim, M.H. et al., Journal of Wind Engineering and Industrial Aerodynamics, 144, p. 84-95.
- Schoetter, R., Masson, V. et al., Geoscientific Model Development, 2020, 13, p. 5609-5643.
- Soulhac, L., Salizzoni, P. et al., 2011, Atmos. Environ., 45, 7379-7395.