



ANNEXE 2

Présentation du projet de thèse

Apport de l'intelligence artificielle pour étudier le climat urbain à partir de données d'opportunité connectées.

Résumé pour un public non scientifique (saisie libre de 1000 caractères max. impérativement en français)

Ce paragraphe pourra être communiqué, sauf avis contraire du porteur de projet.

Le but de cette thèse est d'explorer les apports que peuvent induire l'utilisation des données de masse et d'opportunité dans la connaissance de la variabilité du climat urbain et dans la prévisibilité de celui-ci à très fine échelle.

Le premier objectif est de concevoir des algorithmes de fusion de données afin d'exploiter les données de plusieurs milliers de stations météorologiques connectées sur chaque agglomération et les données des systèmes embarqués de plus de 400000 voitures connectées.

Le deuxième objectif est d'explorer les méthodes de deep learning pour améliorer les prévisions des modèles météorologiques en ville, en combinant ces données de masse de différentes natures (stations, voitures, modèles).

A terme, ceci permettra de développer des services climatiques urbains pertinents pour les utilisateurs, comme des services de prévision météorologique et d'état de la chaussée aux usagers des véhicules connectés en ville (comme ce qui est initié dans le cadre de la collaboration avec Continental sur la place toulousaine).

Présentation scientifique du projet de recherche (objectifs, méthode, ...) (saisie libre de 5000 caractères max.) :

Les villes influencent fortement la météorologie locale, créant un climat urbain spécifique. Il se traduit par une importante variabilité temporelle mais aussi spatiale au sein des villes, en lien avec les hétérogénéités paysagères, morphologiques, architecturales, ainsi que l'usage des bâtiments. Pour une meilleure prévisibilité du climat urbain, et pour répondre aux attentes des collectivités de plus en plus tournées vers l'atténuation de l'îlot de chaleur urbain et un aménagement plus durable, une meilleure connaissance des effets à fine échelle du climat urbain et de leur représentation dans les modèles atmosphériques s'impose.

Toutefois, très peu de mesures conventionnelles existent pour décrire la

météorologie spécifique des villes. Les stations météorologiques sont placées en général sur les aéroports à l'extérieur des villes.

Cependant, la forte concentration démographique et l'évolution technologique, offrent en ville des opportunités uniques d'accès à des données météorologiques de masse, mise en œuvre par les citoyens et les divers vecteurs avec des systèmes embarqués. Grâce à un partenariat avec l'entreprise Continental Automotive, basée dans la région à Toulouse, le CNRM travaille actuellement à la qualification des données issues de véhicules connectés, à partir d'une flotte totale de 400 000 véhicules, en utilisant notamment des données d'un réseau de mesures météorologiques déployé par Toulouse Métropole dans le cadre de la smart city. Cette utilisation des données issues des systèmes embarqués connectés des voitures pour étudier la météorologie et le climat urbain est actuellement à notre connaissance une première mondiale.

Toutefois, dans le cadre de la météorologie, ces données d'opportunité ont été étudiées séparément, et n'ont jamais été envisagées conjointement. Franchir ce pas ouvre de nouveaux enjeux scientifiques, mais aussi des verrous, liés aux données de masse et à l'apprentissage machine.

Objectifs scientifiques

Les objectifs scientifiques de la thèse visent à répondre à ces enjeux liés aux données d'opportunité pour la prévision du climat urbain. Ils sont au nombre de deux, complémentaires, et seront abordés successivement au cours de la thèse.

1) Fusion de données

Il s'agit de lever les verrous scientifiques liés à la fusion de données d'opportunité de masse, de nature différente (stations personnelles, données des collectivités, systèmes embarqués des voitures connectées), pour améliorer la connaissance de la variabilité du climat urbain.

2) Utilisation de l'apprentissage machine pour prévoir la variabilité fine-échelle du climat urbain à partir de données urbaines, d'opportunité et de simulation numérique.

Il s'agit ici de choisir et éventuellement adapter une méthode d'apprentissage statistique, et potentiellement du « deep learning », pour extraire de l'information supplémentaire des données d'opportunité multi-source et de la modélisation atmosphérique pour diagnostiquer puis prévoir la variabilité du climat urbain à une échelle plus fine que celle simulée par le modèle de prévision du temps.

Méthodologie

1) Fusion de données

Cette approche se base sur plusieurs sources de données. Les quantités physiques utilisées seront essentiellement des variables météorologiques, principalement la température de l'air. Ceci permet de se focaliser sur la

problématique de fusion multi-sources, sans se préoccuper de la transformation physique de la donnée. L'utilisation de ces données pose toutefois des problèmes de qualité, inconnue a priori, et est fortement dépendante de la façon dont les capteurs sont utilisés (station météo sur un balcon, voiture dans un embouteillage,...).

Ces données sont :

- Les stations météorologiques connectées personnelles Netatmo
- Les données des systèmes embarqués des véhicules connectés
- Des données décrivant le tissu urbain, disponibles à l'échelle du pâté de maison.
- Des images satellite de température de surface.

Elles seront complétées par des données qui pourront permettre une validation croisée avec des données indépendantes.

- Des stations météorologiques mises en œuvre dans le cadre de la smart city par Toulouse Métropole.

Le travail sera de trouver des méthodes pour fusionner ces données de façon à en extraire de l'information pour mieux décrire, puis analyser, le climat urbain et sa variabilité spatiale et temporelle.

- 2) Utilisation de l'apprentissage machine pour prévoir la variabilité fin-échelle du climat urbain à partir de données urbaines, d'opportunité et de simulation numérique.

Il s'agira là d'explorer comment améliorer la description du climat urbain simulé par les modèles atmosphériques. Les modèles qui seront utilisés décrivent déjà les interactions entre ville et micro-climat, mais à l'échelle de la maille du modèle (1.5km de côté).

Les prévisions du modèle AROME de Météo-France depuis 2008 permettent d'avoir une masse de données simulées importante comme les températures de l'air ou des routes en ville. Elles recouvrent les périodes pour lesquelles sont disponibles les données d'opportunité (depuis 2017).

Il conviendra dans cette seconde moitié de la thèse d'explorer des méthodes d'apprentissages machines novatrices permettant d'améliorer a posteriori la description du climat urbain simulé à l'échelle de la maille du modèle jusqu'à une échelle plus fine (des quartiers et même jusqu'aux routes) en prenant en compte les observations et le tissu urbain.

Liens du sujet de thèse avec les activités de l'unité de recherche d'accueil (saisie libre de 1000 caractères max.) :

L'équipe VILLE du CNRM est dédié aux recherches en climat urbain, qui ont débuté il y a 20 ans au laboratoire. Les travaux du laboratoire sont reconnus mondialement, et le CNRM a ainsi organisé en 2015 la 9th International Conference on Urban Climate, qui a regroupé 550 chercheurs de plus de 60 pays.

Les recherches s'axent selon deux directions principales : (1) l'amélioration des modèles numériques de prévision du temps et du climat en ville, notamment via le développement du modèle de canopée urbaine TEB, utilisé par une vingtaine de services météorologiques dans le monde (2) l'amélioration de la connaissance du climat urbain, par l'observation et l'étude de processus. La thèse proposée concerne le 2^{ème} axe. Elle se basera sur les collaborations en cours avec Toulouse Métropole, Continental Automotive et les travaux récents, comme celui mené avec l'université technologique de Berlin pour exploiter des stations météorologiques connectées Netatmo pour étudier l'îlot de chaleur urbain sur Toulouse, Paris et Berlin.

La Directrice-Adjointe
UMR 3589-CNRM


Claire DOUBREMELLE