



Subject of thesis:

Big Data Analytic Platforms for Weather Analysis and Prediction: Mediterranean as a case study

CONTEXTE

Le changement du climat est un danger qui menace le globe terrestre en entier. Les experts et les météorologistes du monde essaient d'étudier son développement afin de limiter ses effets. Les éléments du temps sont affectés par ce changement et causent actuellement des phénomènes sévères causants des dégâts matérielles et humaines vastes.

En fait, l'Organisation Météorologique Mondiale (OMM), une institution spécialisée des Nations Unies, serve à collecter les données météorologiques provenant de stations distribuées dans un grand nombre de pays à travers des messages de format standard (SYNOP). Ces messages sont transmis par les stations d'une manière périodique, chaque trois heures, et contenant des informations de l'état actuel de tous les paramètres du temps (température, humidité, vent, pression, etc.). Ces informations sont sauvegardées et analysées pour produire les cartes météorologiques et faire les prévisions du temps. Elles sont aussi utilisées pour des buts climatiques pour obtenir les normes des éléments (température, humidité, précipitation, vent, etc.)

OBJECTIVE

Grace au nombre de stations et la collecte périodique des informations, une quantité massive des données ont été collecté au bout des années et décennies précédentes. Souvent, ces données constituent une source importante et riche aux experts qui peuvent la découvrir et l'exploiter afin de trouver des corrélations et prédire le changement du climat dans certains pays. Par conséquence, beaucoup de travaux ont été déjà fait dans des pays comme Chine, Etats-Unis, Allemagne, etc.

Le but de ce travail est d'étudier et d'analyser les données collectées à partir des messages SYNOP entre les stations météorologiques au bassin du méditerranée (comme Liban, France, Espagne, ...). Cependant, le travail vise à exploiter des nouvelles corrélations entre les données collectées au bout des derniers trente ans de pays de méditerranée, tout en utilisant des méthodes de recherche efficaces. Particulièrement, ce travail va concentrer sur la relation entre l'élévation de température, due au changement de climat, et la variation de la vitesse du vent et par suite les tempêtes. L'importance d'établir cette relation est de suivre le développement des orages et cyclones maritimes et aussi prévoir la vitesse du vent lors d'un changement de température pour prendre les précautions nécessaires et, par conséquent, sauver des vies. En fait, l'effet de serre provoque des régions chaudes où le vent est ascendant causant une zone de basse pression. Un vent provenant des régions de haute pression va le remplacer causant une circulation additionnelle du vent et qui s'ajoute aux circulations thermiques naturelles du vent. En plus, les surfaces maritimes et océaniques conservent l'énergie calorifique pour des périodes plus longues et nourrissent toutes les tempêtes passantes au-dessus en élevant ses puissances.

TRAVAUX ATTENDUS

La première partie du travail consiste à formuler un état de l'art des stations météorologiques et leur fonctionnement. Le but de cette partie est d'étudier l'architecture de ses stations, ses méthodes de collecte et transmission de données tout en concentrant sur le défi de collecte volumineuse de données.

Dans la deuxième partie, nous nous intéressons aux méthodes d'exploitation de données existantes comme intelligence artificielle, machine learning, data mining, etc. Elle consiste à étudier, intégrer ou proposer des nouvelles solutions au problème décrit ci-dessus.

Finalement, ce travail comprend également une partie pratique qui consiste à implémenter une plateforme basée sur les outils de Big Data (Hadoop et son écosystème) qui prend les données directement des stations, les sauvegarder, puis de faire exploiter les relations afin de pouvoir prédire les performances des algorithmes et des méthodes proposées.

MOTS CLÉS

Stations météorologiques ; Surveillance du climat ; Remote sensing; Méthodes d'exploitation de données; Analyse et prédiction de données ; Big data ; Hadoop écosystème.

REFERENCES

- Mwaniki W. Mercy. **“Assessment of the effects of climate change on land use and land cover using remote sensing: A case study from Kenya”**. DNC Working Paper, United Nations University Institute for Integrated Management of Material Fluxes and of Resources (UNU-FLORES), Vol. 3, 19 pages, 2017.
- S. N. Palve, P.D. Nemade, S.D. Ghude. **“The application of remote sensing techniques for air pollution analysis and climate change on Indian subcontinent”**. 8th IGRSM International Conference and Exhibition on Remote Sensing & GIS, 2016.
- Piero Lionello, Luca Scarascia. **“The relation between climate change in the Mediterranean region and global warming”**. Regional Environmental Change, Vol. 18, Iss. 5, pp 1481-1493, 2018.
- Yuriy Kuleshov. **“Use of Remote Sensing Data for Climate Monitoring in WMO Regions II and V (Asia and the South-West Pacific)”**. Australian Bureau of Meteorology. 2017.
- Anna Luganska and Raja Chakir. **“Climate-induced Land Use Change in France: Impacts of Agricultural Adaptation and Climate Change Mitigation”**. Université Paris-Saclay, France, pp.134-154, 2017.
- S. del Rioa, L. Herrero, C. Pinto-Gomes, A. **“Penas. Spatial analysis of mean temperature trends in Spain over the period 1961–2006”**. Global and Planetary Change, Vol. 78, pp.65-75, 2011.
- Ahmed Oussous, Fatima-Zahra Benelloun, Ayoub Ait Lahcen, and Samir Belfkih. **“Big Data technologies: A survey”**. Journal of King Saud University-Computer and Information Sciences, Vol. 30, Iss. 4, pages 431–448, 2018.
- Mehdi Mohammadi, Ala Al-Fuqaha, Sameh Sorour, and Mohsen Guizani. **“Deep learning for IoT big data and streaming analytics: A survey”**. IEEE Communications Surveys &

Tutorials, Vol. 20, No. 4, pages 2923–2960, 2018.

- Hongsong Chen, and Zhongchuan Fu. **“Hadoop-based healthcare information system design and wireless security communication implementation”**. Mobile Information Systems, Vol. 2015, 2015.
- Chen Prakash. **“Spark Streaming vs Flink vs Storm vs Kafka Streams vs Samza : Choose Your Stream Processing Framework”**. 21 March 2018. [Online]. Available: <https://www.linkedin.com/pulse/sparkstreaming-vs-flink-storm-kafka-streams-samza-choose-prakash/>.
-