AquaSmart Solutions

Group 21

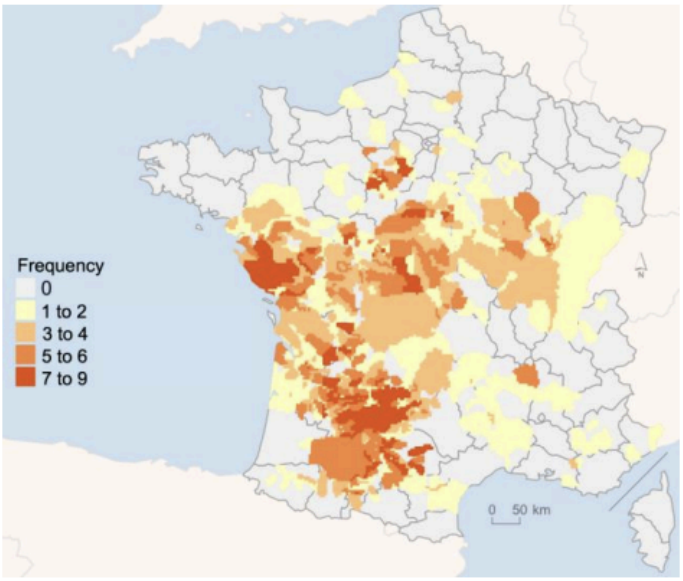
Overview

Water is the most valuable resource on Earth, necessary not only for drinking but also for many other uses such as power generation and agriculture. While water is extremely abundant, making up over 70% of the planet's surface area, most of this water is not fresh and difficult to use. Desalination to turn salty ocean water fresh is a costly and intensive endeavor that is often not the most efficient way to obtain fresh water. Hence, natural freshwater sources, such as groundwater are often used to supply the population. In France specifically, groundwater is consumed at approximately the same level as surface water, excluding canal and power station cooling usage.[1]

A key issue in these freshwater supplies emerges when looking at the supply throughout the seasons. Summer in particular is a strenuous time for freshwater supplies, when only 2% of the annual supply is available.[2] In France, this season is often filled with droughts and water restrictions, which is especially costly to the agricultural industry. Farmers, having just planted a fresh crop in the spring, are adversely affected by these restrictions; if they have planted too many crops in spring, they may not be able to provide them all the necessary water throughout the summer. Additional issues emerge when certain regions, both urban and rural, have high water demands but low groundwater reserves. When this occurs, water must be transported into these regions from regions not hit as hard. This transportation is often very costly, sometimes using trucking and shipping that can cost over €100,000.[3]

---

[1] https://www.statistiques.developpement-durable.gouv.fr/media/5639/download?inline
[2] https://www.statistiques.developpement-durable.gouv.fr/media/5639/download?inline
[3] https://www.dw.com/en/water-crisis-in-france-drought-hit-regions-rely-on-tankers/a-65649291



The map[4] shown above displays regions in France between 2012 and 2020 the number instances where water restrictions lasting at least a month.

Business Approach:

If a machine learning model could predict these droughts, hundreds of millions of euros could potentially be saved by the agricultural industry by adjusting their crop sizes to align with the summer's water availability since the agricultural industry spends about €3.4 billion on water annually.[5] Additionally, resources could be more appropriately allocated by water utilities suppliers to account for transportation costs, with the French water market taking in €85.1 billion yearly.[6] Such a model would be invaluable to both farmers and local utilities providers, making it potentially a highly profitable endeavor. A startup business developing such a model would be able to sell its predictions to these customers.

---

[4] https://www.statistiques.developpement-durable.gouv.fr/media/5639/download?inline
[5] https://www.french-property.com/guides/france/utilities/water
[6] https://www.marketresearch.com/MarketLine-v3883/Water-Utilities-France-38021008/

Scientific Approach:

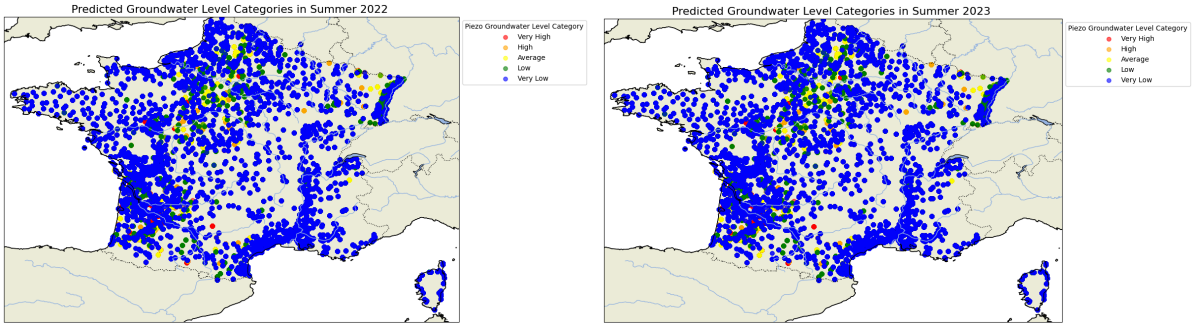
Our collected data contained a significant number of NaNs, so an important pre-processing phase was required. We thoroughly studied the missing data to try to complete it as best as possible. Sometimes it was possible to infer, such as with the weather, which can be estimated from the weather data provided by nearby stations. We sometimes used the median, grouping the closest data – particularly the date. From our clean dataset, we studied the variables (correlations, irrelevant data, classification by a Random Forest) to keep only the most relevant data. We then sorted the data according to whether they are numerical (processed by robust scaler) or categorical, separated into ordinal (processed by ordinal encoder) or non-ordinal (processed by one hot encoder). Our model relies on a Random Forest classifier that uses independently created decision trees to reduce overfitting. This model is particularly suited to our case as we seek to classify data.

Our approach is clearly in line with an environmentally friendly approach. Our model, being resource-efficient, can run on older servers, allowing us to recycle end-of-life servers and avoid resource waste. Additionally, we have agreements with companies supplying electricity in France to buy excess electricity from green power plants (nuclear, wind, solar). This not only allows us to operate on green energy and develop the sector but also minimizes their losses.

Our model has significant scalability because its predictions rely only on the previous 2 to 3 summers, which represents a small amount of data. This ensures continuity of performance and speed in the future.

Results and Future Potential:

The results of our model are conclusive: over 54% accuracy despite the small amount of data available to us. Our model shows enormous potential thanks to the development of precision and the increase in the number of measurements we need to make our model work.



We can then couple this model with knowledge of the water market in France to ensure optimal water purchases to resell at the best price to farmers, below market price, ensuring reduced costs for farmers and a significant profit margin for us.