



GEOProcessing 2025

Extra Virgin Olive Oil Price Prediction from Multi-source Variables and Machine Learning

Juan J. Cubillas
Dept. Information and
Communication
Technologies Applied to
Education
International University of
La Rioja

Ángel Calle
Dept. Computer Science
University of Jaen

M. Isabel Ramos
Dept. Cartographic Engineering,
Geodesy and Photogrammetry
University of Jaen (Spain)
miramos@ujaen.es

Ruth Córdoba
Dept. Cartographic Engineering,
Geodesy and Photogrammetry
University of Jaen (Spain)



Universidad de Jaén

unir
LA UNIVERSIDAD
EN INTERNET

ujaen.es

Motivation of the research

Price of Extra Virgen Olive Oil (€ /Kg)



Motivation of the research



State of the art

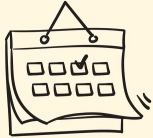
- Global significance.
- Use of Machine Learning in food pricing analysis.
- Predicting Extra Virging Olive Oil price.
- Challenges.





Objetive

2012-2024



Base Price EVOO



Diesel fuel price



Rainfall

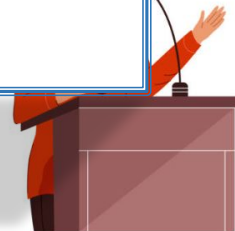
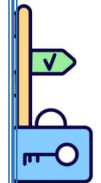


Predict the price Extra Virgin Olive Oil (EVOO) in SPAIN



Crop yield prediction^{a)}

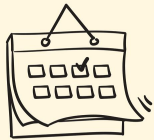
Oil production
in the world



Methodology

A. Comprehensive and Multi-source Data Acquisition and Integration.

2012-2024



Base Price EVOO



Diesel fuel price

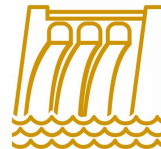


Rainfall



CONSEJO
OLEÍCOLA
INTERNACIONAL

World Oil Production



level in reservoirs



Fertiliser price



IPC



Crop yield prediction^{a)}



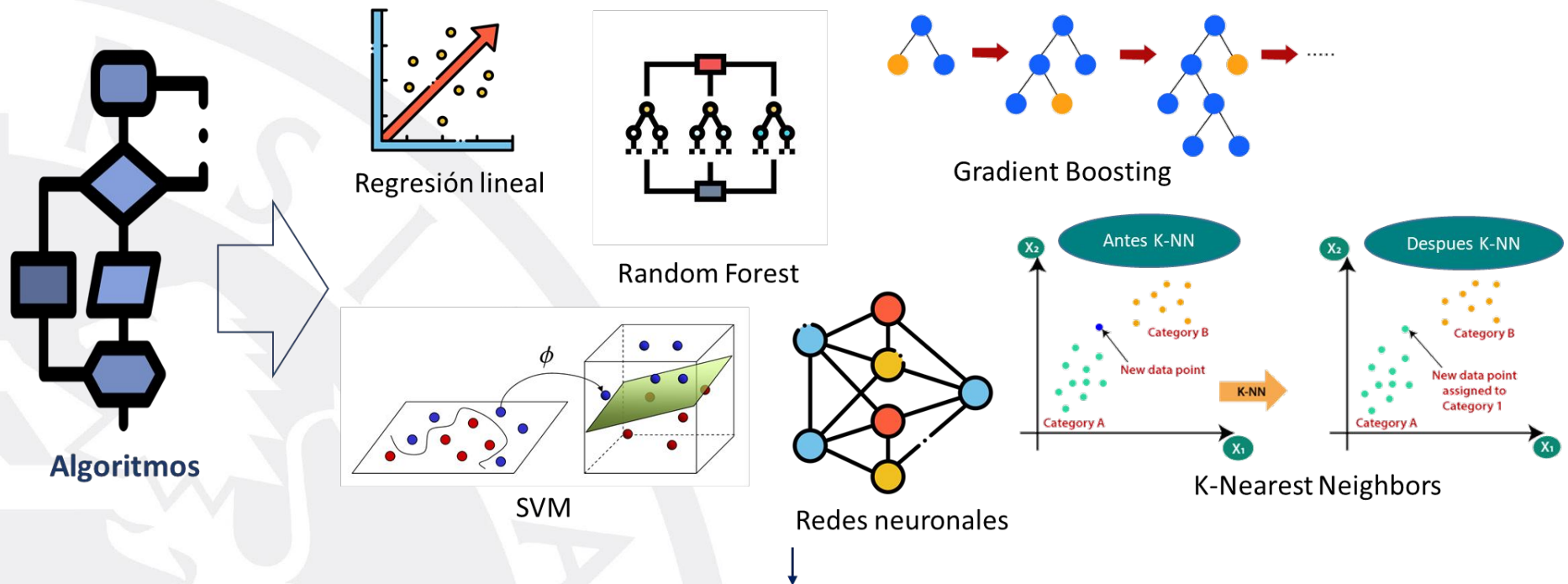
Methodology

B. Data preparation

- Data must be prepared to be compatible with ML algorithms, which involved **categorisation** and **normalisation**.
- This included **temporal aggregation** (monthly, daily, cumulative), spatial selection (province of Jaén as reference) and categorisation by month to capture seasonality.

Methodology

C. Application and Comparison of Machine Learning Algorithms

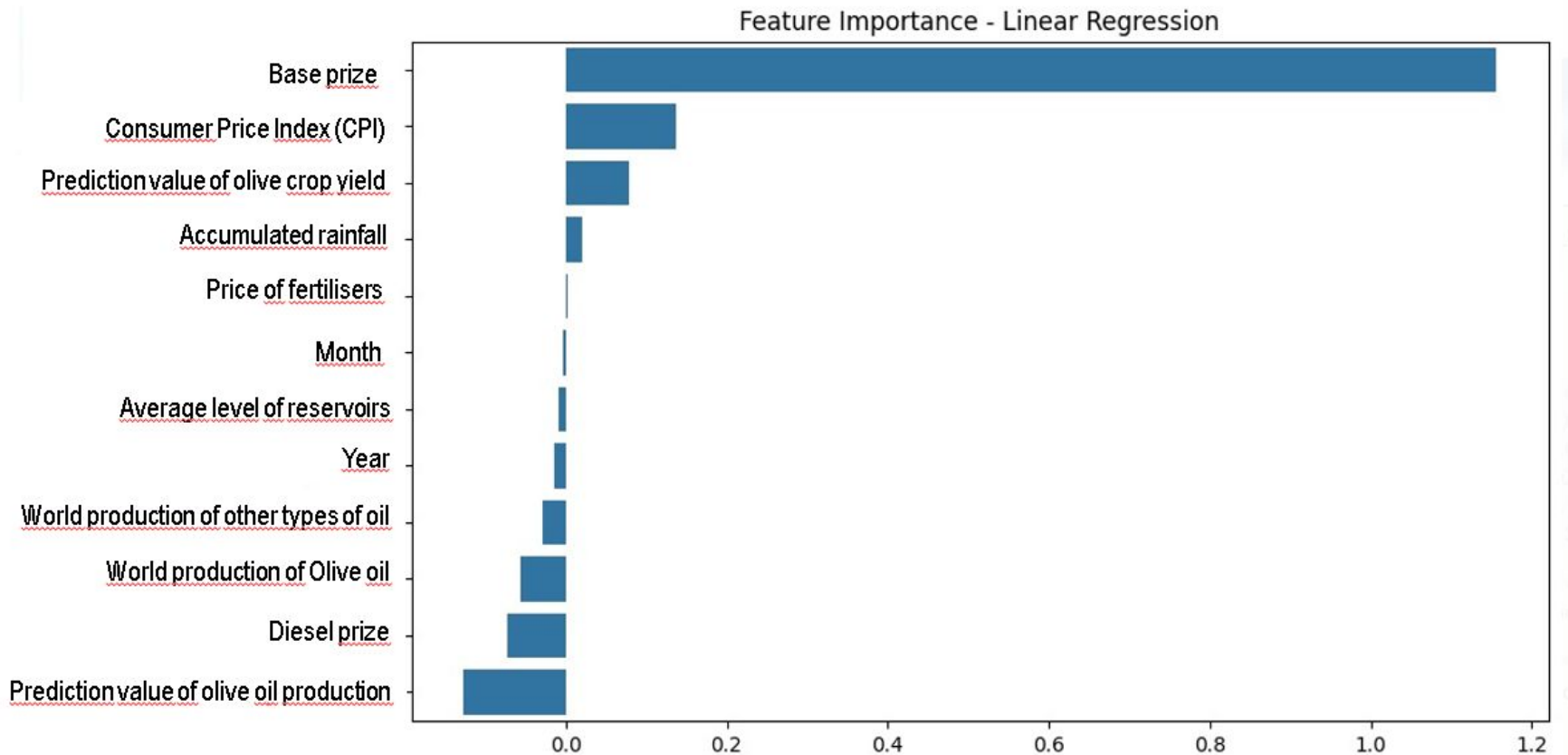


Methodology

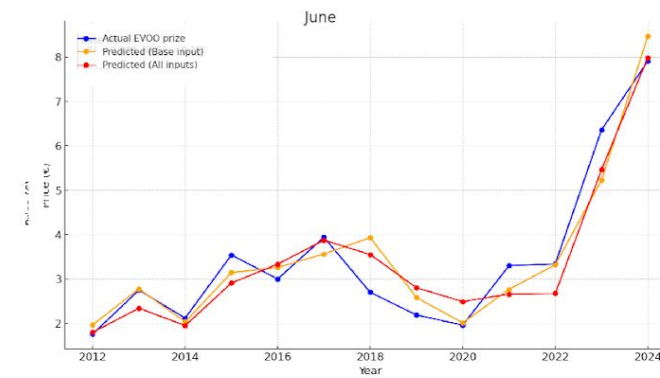
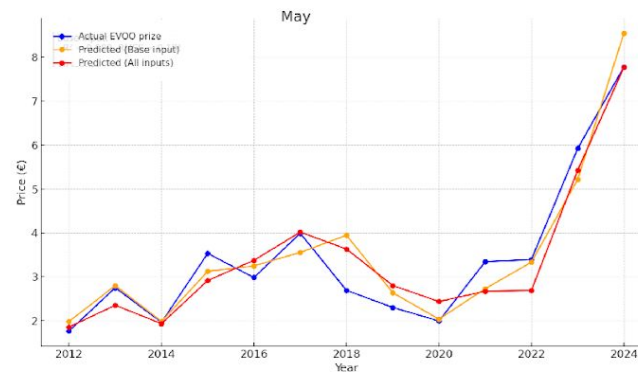
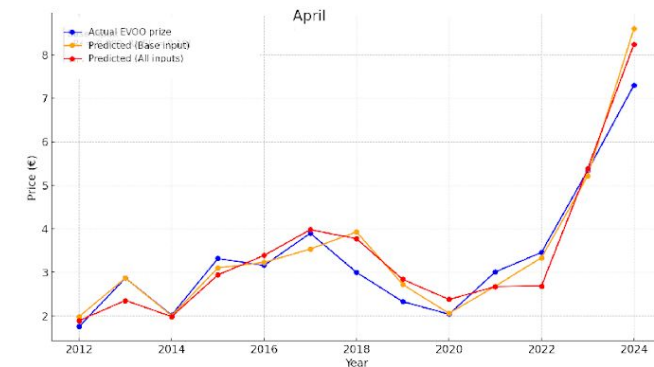
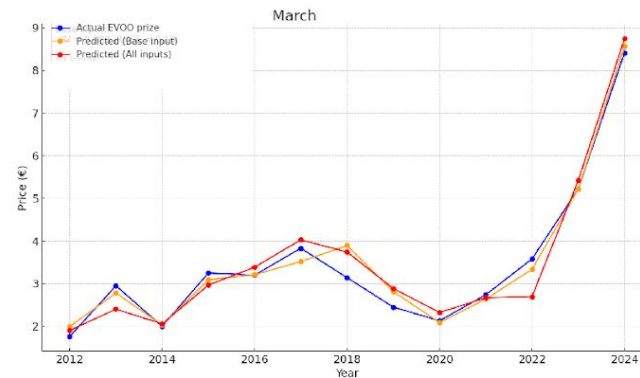
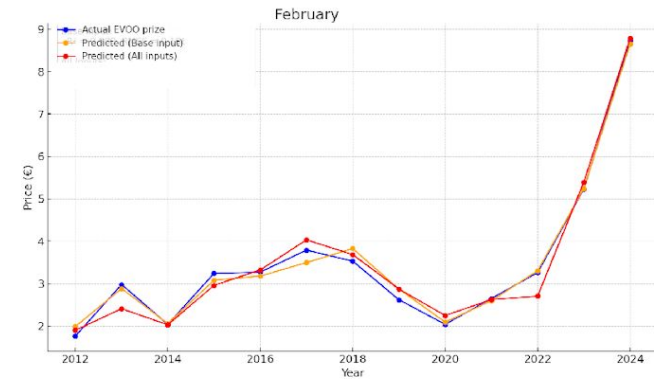
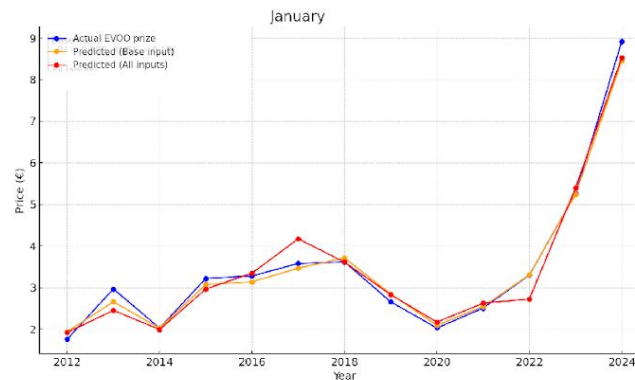
D. Model Evaluation and Validation

- The accuracy and fit of the models were analysed using error and fit metrics such as MSE, RMSE, MAE, MAPE and R^2 , and the Explained Variance Score.
- Cross-validation was performed to assess the quality and reliability of the models on data not seen during training .
- Validation was checked on a full-year and monthly basis to assess overall accuracy and to detect possible seasonal patterns in predictive ability.
- The performance of the model using all input variables was compared with a base model using only the historical price to confirm the added value of including the full set of factors.

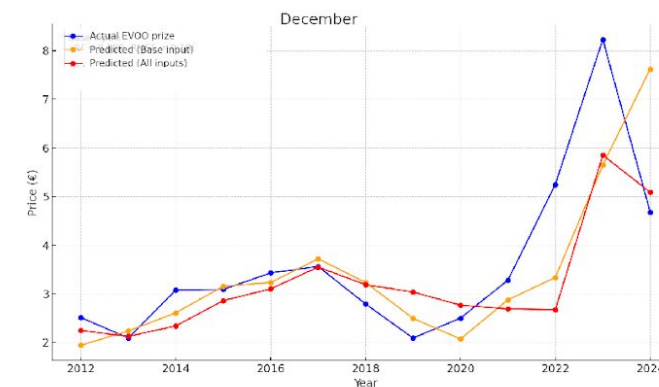
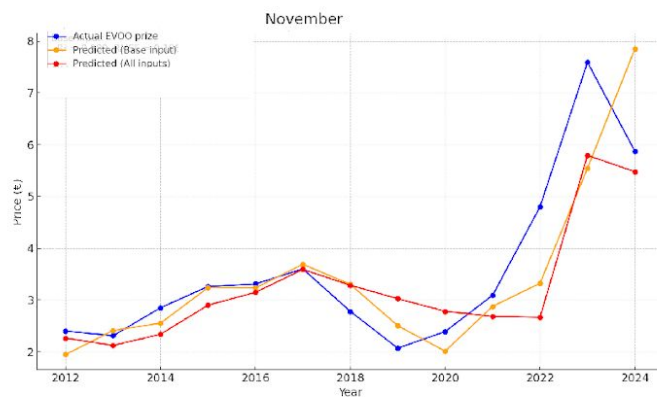
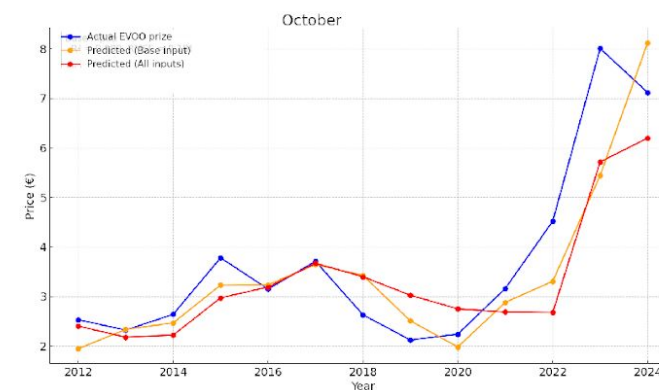
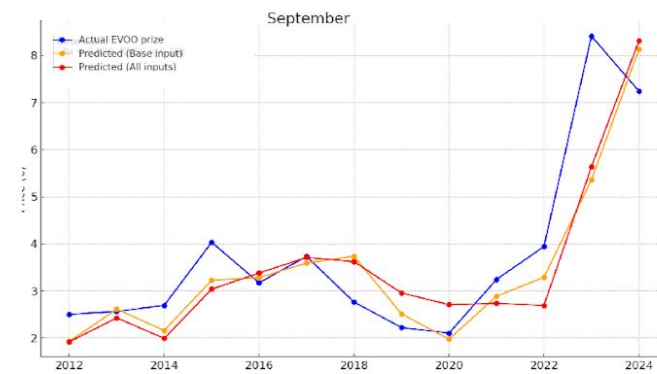
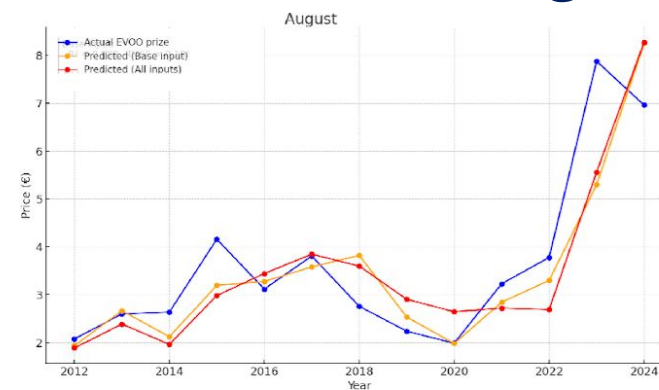
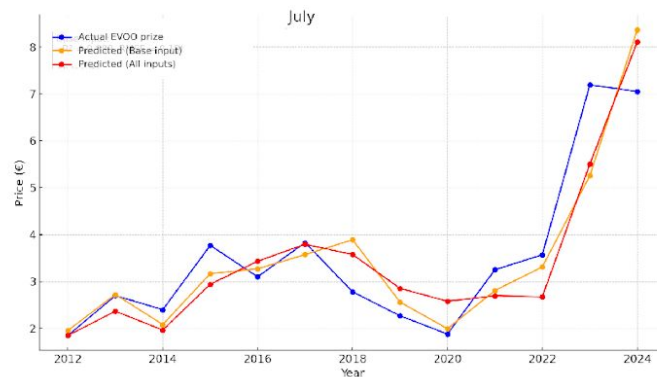
Results



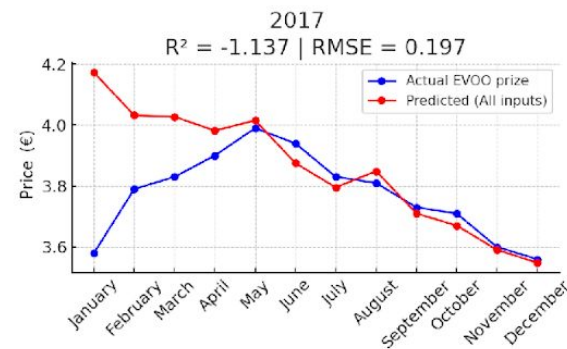
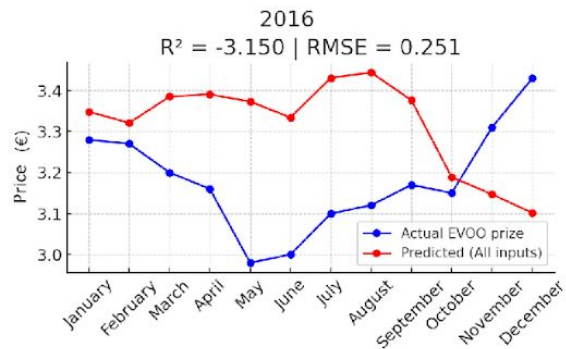
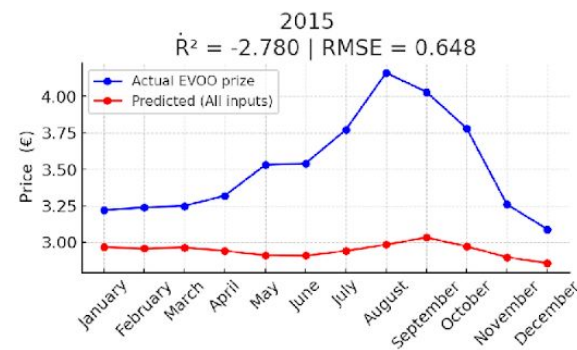
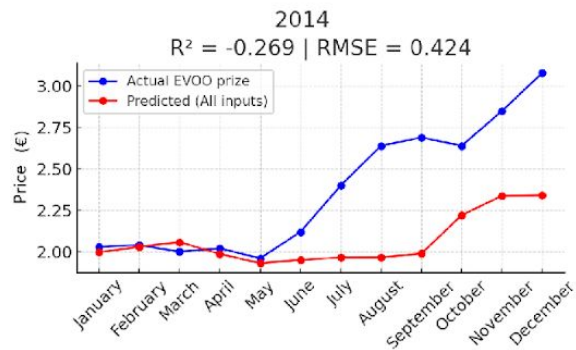
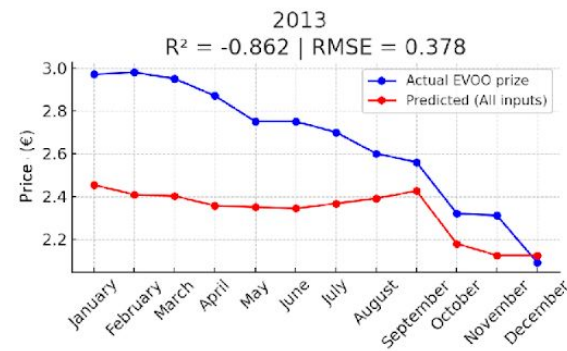
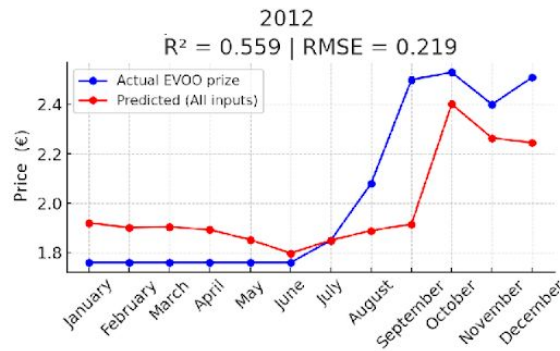
Results



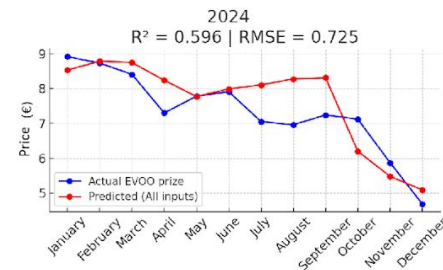
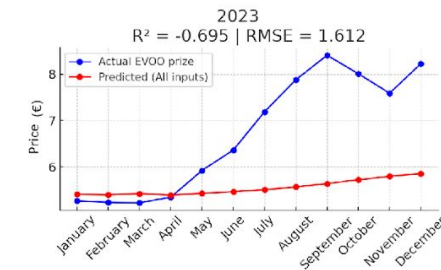
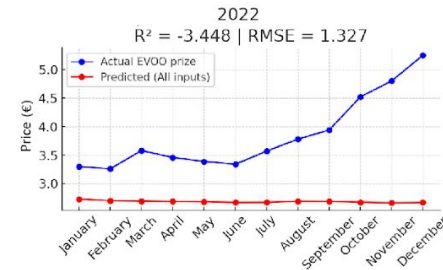
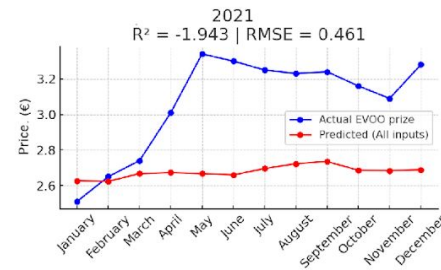
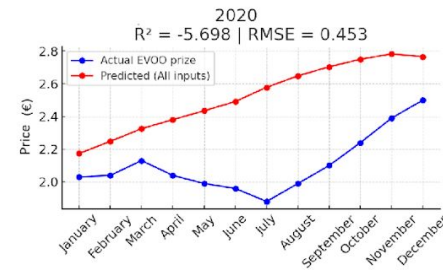
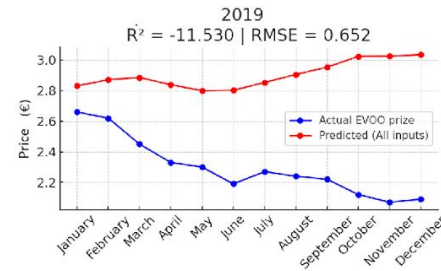
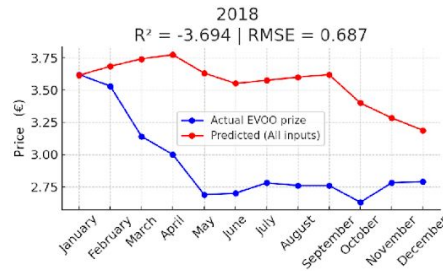
Results



Results



Results



Conclusions

1. **A robust methodology has been designed to achieve the objective of the work.** This is attributed to the integration of a **diverse and complete set of predictor variables** and the high predictive capacity demonstrated by certain algorithms used.
2. The **variables considered** reflect both **local** (weather, costs) **and global factors** (world production, inflation), providing a complete picture of the dynamics affecting the price of EVOO.
3. **Several machine learning algorithms were evaluated.** **Gradient Boosting and Random Forest** models proved to be the **most effective** at capturing the complex, non-linear relationships present in the market. This suggests that the EVOO market is influenced by multiple interconnected factors, and that non-linear models are more effective at capturing this complexity.
4. The **high accuracy of the models** obtained indicates that the selected input variables adequately **reflect the market dynamics in Spain**. It also suggests that the data sources and data processing were appropriate for building predictive models.
5. The promising results of this predictive approach have the potential to contribute to stabilising markets, optimising distribution and improving agricultural budgeting.



GEOProcessing 2025

Extra Virgin Olive Oil Price Prediction from Multi-source Variables and Machine Learning

Juan J. Cubillas
Dept. Information and
Communication
Technologies Applied to
Education
International University of
La Rioja

Ángel Calle
Dept. Computer Science
University of Jaen

M. Isabel Ramos
Dept. Cartographic Engineering,
Geodesy and Photogrammetry
University of Jaen (Spain)
miramos@ujaen.es

Ruth Córdoba
Dept. Cartographic Engineering,
Geodesy and Photogrammetry
University of Jaen (Spain)



Universidad de Jaén

unir
LA UNIVERSIDAD
EN INTERNET

ujaen.es