GEOProcessing 2024

Improved prediction of olive crop yield using satellite imagery

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

Lidia M. Ortega
Dept. Computer Science
University of Jaen

Ángel Calle
Researcher hired in project PREDIC I-GOPO-JA-20-0006
University of Jaen

Ruth M. Córdoba
Researcher hired in project PREDIC I-GOPO-JA-20-0006
University of Jaen

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

Universidad de Jaén

unir
LA UNIVERSIDAD EN INTERNET

ujaen.es

GEOProcessing 2024

Improved prediction of olive crop yield using satellite imagery

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

PhD in Geodesy and Cartography Engineering in 2008, University of Jaén. Teaching in this University since 1999. Senior Lecturer since 2019. In 2015, Extraordinary PhD Award. PhD thesis dealt with application of GNSS to olive grove environment. Additionally, her experience with the analysis of spatial data, geodatabase and high precision positioning techniques has allowed her to collaborate in multidisciplinary researches teams and she has lead several projects. Research interests focus on the integration of geomatics technologies in environmental applications, mainly in agriculture. Specially on analysis of geospatial data, remote sensing with UAV, GNSS and spatial data mining.

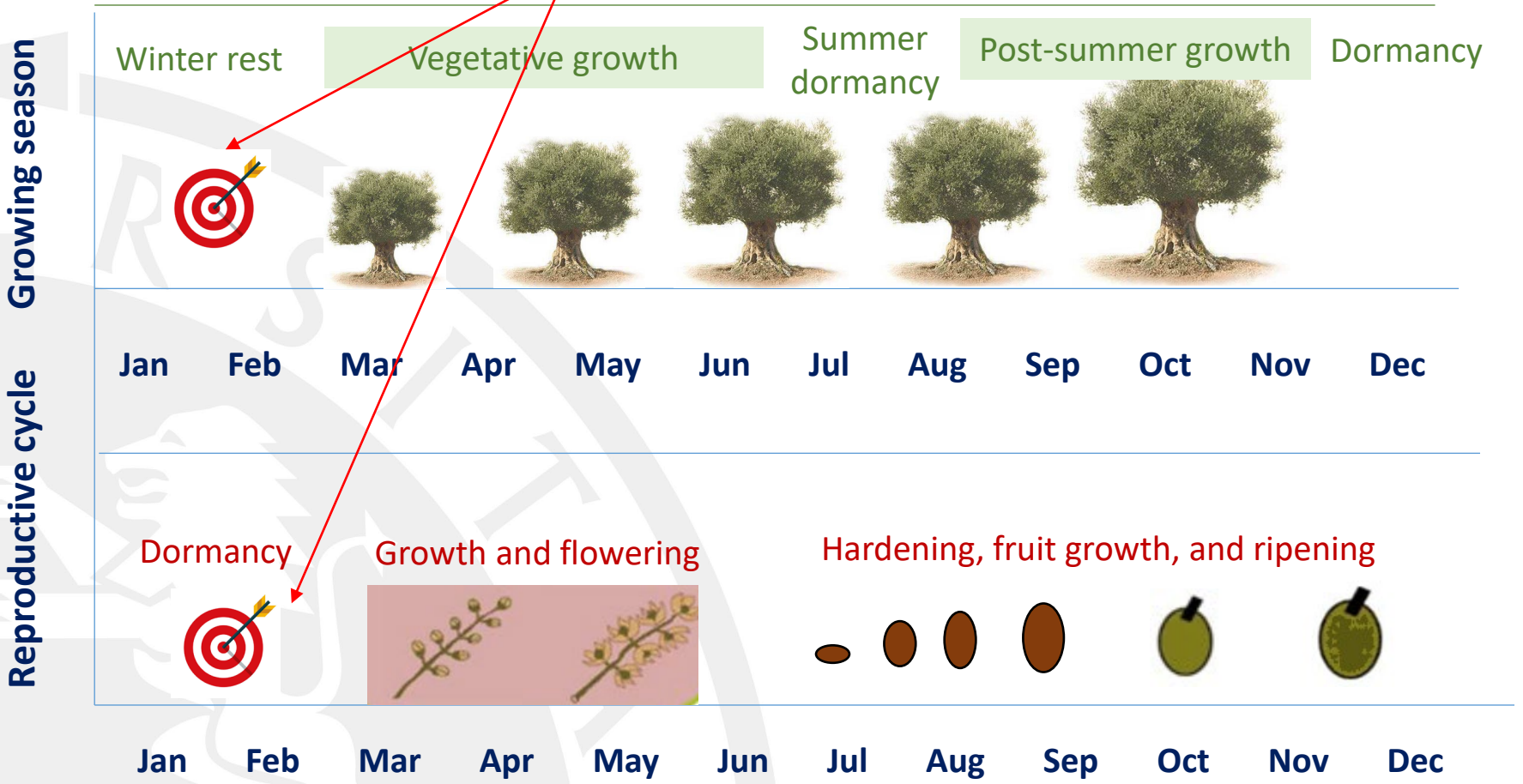
Motivation of the research

Agriculture is one of the strategic economic and social sectors in Spain

Universidad de Jaén

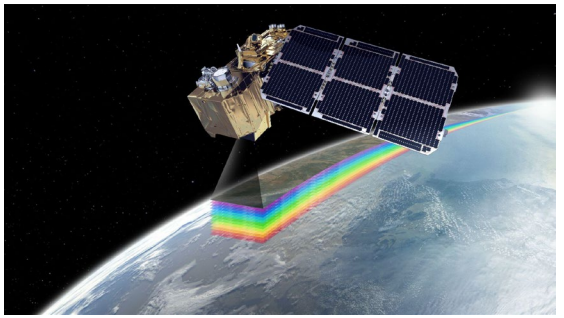
Objetive

Early prediction of crop yield



Objetive

Early prediction of crop yield



Data collection

- Historical olive harvest data
- Meteorological data

Data set

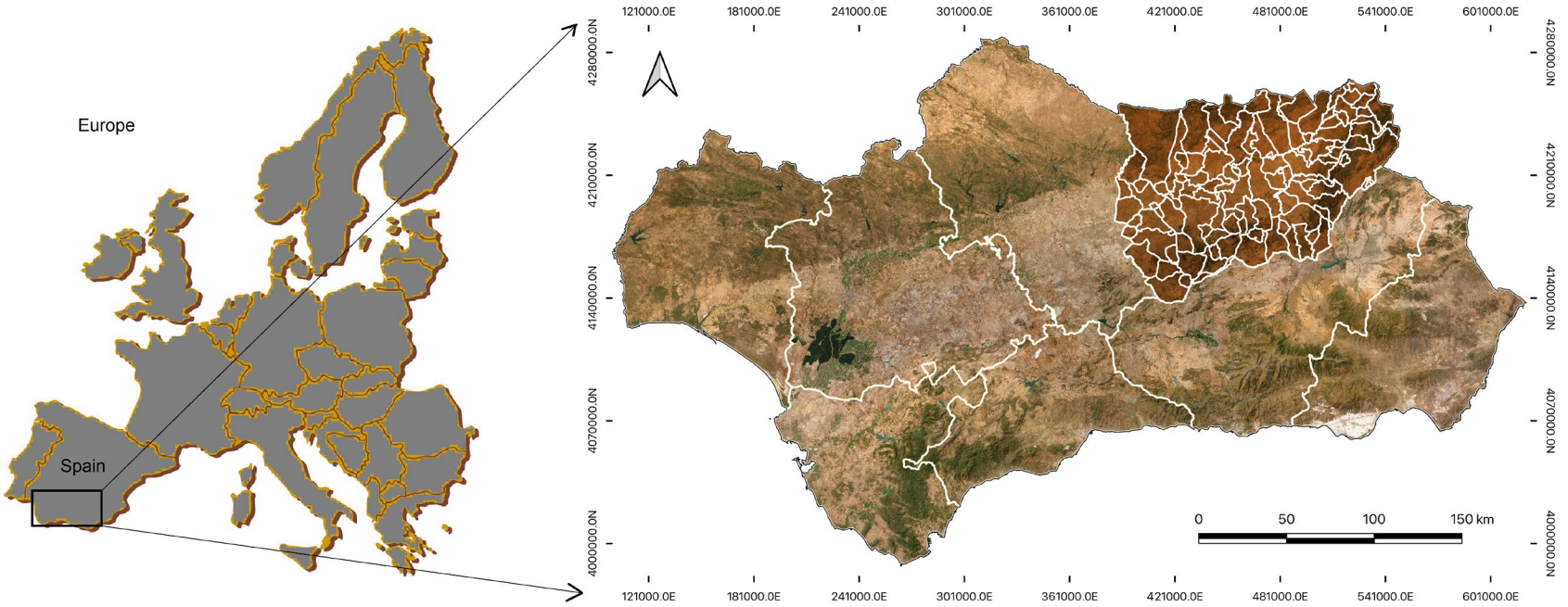
- Analysis of histograms
- Anomaly detection
- Transformation and scaling
- Grouping
  - MDL algorithm
- Influential data on the target variable

Models

- Train set
  - Algorithms:
    - GLM
    - SVM Gaussian
    - SVM Linear kernel
- Cross-validation
- Output

Methodology

Research Area



Methodology

Dataset

Data collection

- Historical olive harvest data
- Meteorological data

Data set

- Analysis of histograms
- Anomaly detection
- Transformation and scaling
- Grouping
  - MDL algorithm
- Influential data on the target variable

Municipios IGN

Acreditado

Esri España Contenidos

Contenidos Esri España

AEMet

Agencia Estatal de Meteorología

Nº Pesada	Fecha	Rendimiento	Kgs Pesada
197	26-dic	17,89	1.054
289	27-dic	17,94	1.323
368	28-dic	17,99	1.168
425	29-dic	14,77	1.343
515	30-dic	18,43	1.590
597	31-dic	17,21	2.241
1666	03-ene	18,39	1.208
666	04-ene	14,5	1.791
735	05-ene	15,92	1.322
833	06-ene	16,85	2.151

Methodology

Dataset

Data collection

- Historical olive harvest data
- Meteorological data

Data set

- Analysis of histograms
- Anomaly detection
- Transformation and scaling
- Grouping
  - MDL algorithm
- Influential data on the target variable

Municipios IGN

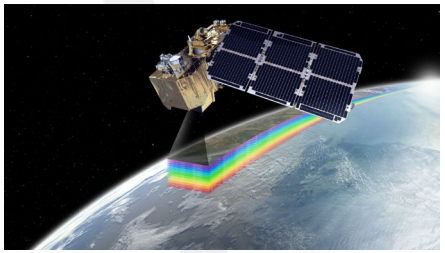
Acreditado

Esri España Contenidos

Contenidos Esri España

AEMet

Agencia Estatal de Meteorología



Nº Pesada	Fecha	Rendimiento	Kgs Pesada
197	26-dic	17,89	1.054
289	27-dic	17,94	1.323
368	28-dic	17,99	1.168
425	29-dic	14,77	1.343
515	30-dic	18,43	1.590
597	31-dic	17,21	2.241
1666	03-ene	18,39	1.208
666	04-ene	14,5	1.791
735	05-ene	15,92	1.322
833	06-ene	16,85	2.151

Methodology

Modelling process

The methodology followed consists of the following phases:

(1) The understanding of the data, identify the information it provides, the format and the meaning of the values;

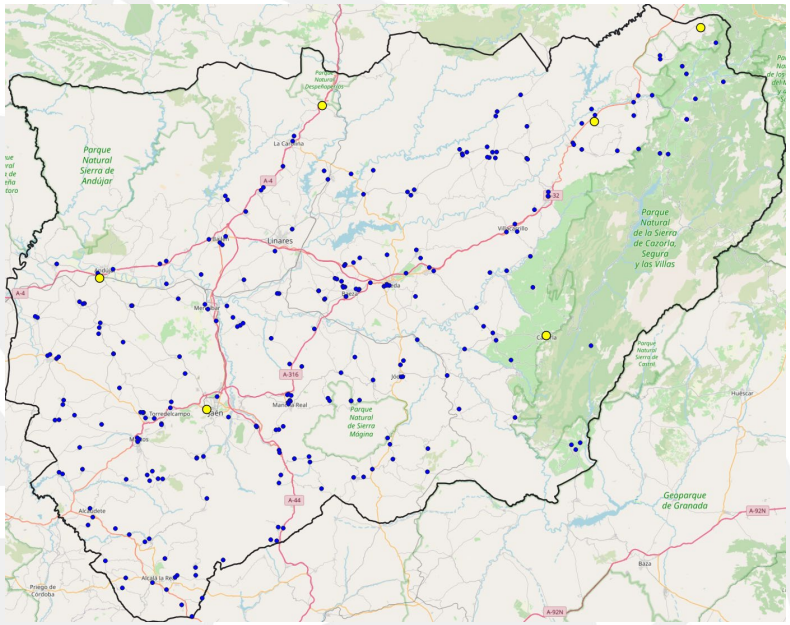
(2) Its acquisition and preparation process, Once captured, the data requires prior preparation before being inserted into ML models, some require categorisation, others labelling, detection of outliers, null values, etc.; (SVM and MDL algorithms).

(3) The generation of the predictive models using different algorithms (SVM, NN) and, finally,

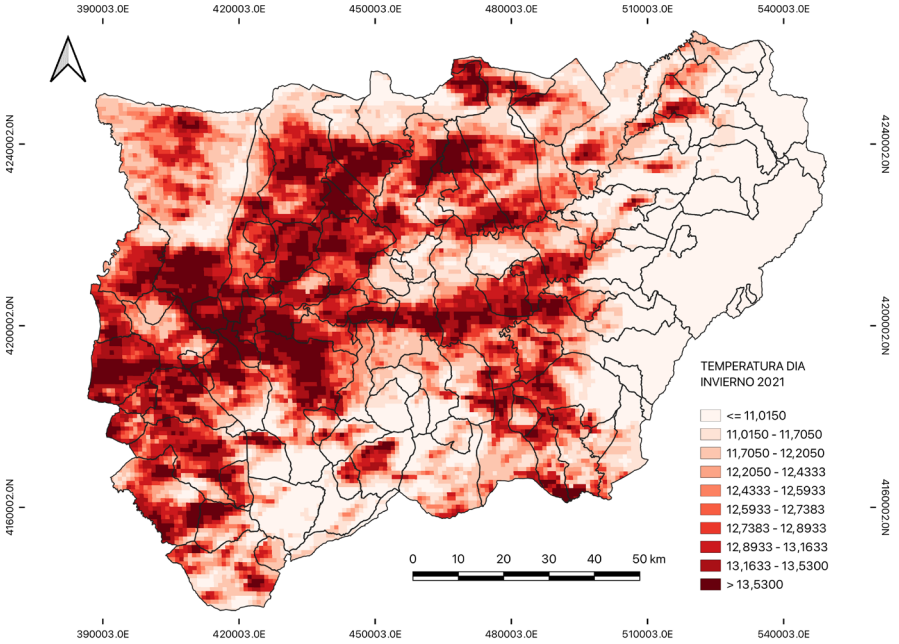
(4) The validation or analysis of the accuracy obtained, in this sense the k-fold cross validation technique was used to evaluate results in statistical analyses. The output data of the model is the amount of olive harvest (the number of Kg of olives). This information is collected for each campaign and for each one of all the province municipalities.

Results

A. Predictive models using only weather station data

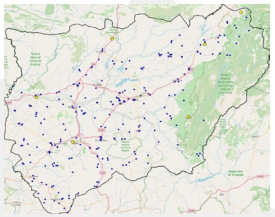


B. Predictive models using satellite imagery data

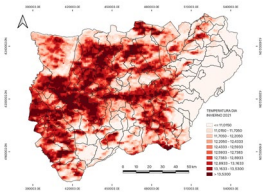


Results

A. Predictive models using only weather station data



B. Predictive models using satellite imagery data



MODIS

ERAS5

SENTINEL-2

Predictor variables	Weather stations			Satellite imagery		
Algorithms	SVM		NN	SVM		NN
	Gaussian	Lineal		Gaussian	Lineal	
Only Rainfall	23,56%	24,29%	25,67%	24,62%	25,70%	27,07%
Rainfall + Temperature	23,63%	24,20%	25,42%	24,58%	25,48%	26,26%

Conclusions

- This study presents a workflow methodology describing the steps followed in the analysis of the predictive calculation of olive crop yield at an early stage.
- The novelty of the work is how early, within the agricultural year, it makes the crop prediction. Just before the investment of resources begins, when there is still no visible or measurable sign of pollen or the beginning of fruit.
- The integration of satellite imagery into the model improves crop yield prediction. This is due to the fact that a better diagnosis of the state of the satellite imagery can overcome this drawback weather next to the area studied contributes to a good early prediction of its production.
- Satellite images have been fundamental in order to have sufficient temporality covering all the municipalities in the province of Jaen.
- The methodology developed is applicable to future works at different spatial scales, even at local or farm detail level.
- The short-term future development of this work includes the generation of an intelligent system in which the variables obtained from the satellite images are extracted automatically as well as the downloading of the meteorological values from web services. Thus, a non-expert user will be able to use the system by simply inserting the harvest values of the farm or the area under study.

Universidad de Jaén

GEOProcessing 2024

Improved prediction of olive crop yield using satellite imagery

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

Lidia M. Ortega
Dept. Computer Science
University of Jaen

Ángel Calle
Researcher hired in project
PREDIC I-GOPO-JA-20-0006
University of Jaen

Ruth M. Córdoba
Researcher hired in project
PREDIC I-GOPO-JA-20-0006
University of Jaen

Juan J.
Dept. I
Techno
Interna

ujaen.es