AgriGeo: Geoinformation solutions for agriculture based on Big Data analytics

Volpe, Fabio (1); Tricomi, Alessia (1); Pistillo, Pasquale (1); Biscardi, Mariano Alfonso (1); Chiara Francalanci (2); Angela Geronazzo (2); Paolo Giacomazzi (2); Paolo Ravanelli (2)

1: e-GEOS, Italy; 2: Politecnico di Milano, Italy
Progressive shift from simple mapping to complex monitoring applications.

High Level of Value Addition

Big data analysis

Information Products

Data

Low Level of Value Addition

Value Added Services
Turning data and images into information that can be browsed and analyzed

Solutions tailored on user requirements

Generation of analytics from the processing of huge time series of different data
Operational workflows

- Services provided by AgriGEO strongly rely on the usage of satellite time series, organised in multi-source Data Cubes.
- They are managed in a fully scalable environment, allowing a fast and efficient extraction of information for feeding **vertical workflow pipelines**, often requiring near-real-time delivery performances.
- One main driver for the provision of these services is the availability of big data analysis techniques enabling the extraction of information and analytics from huge amounts of data.
From mapping to information
Number of mowing events

the esa earth observation φ-week

ED Open Science and FutureEO
12–16 November 2018 | ESA–ESRIN | Frascati (Rome), Italy
The approach exploits multiple data sources and creates a composite model:

- Random Forest model based on optical data
- Bayesian model based on USDA historical series

The model has been developed on different test area, therefore considering differences in climate and growing crops.
The RF model

• Images can be seen as a multidimensional matrices where each value \((i,j)\) represents a set of spectral bands.
• USDA classification is used as ground truth to train ML models.
• An extensive tuning procedure was applied to fit the best set of hyperparameters at single test area level.
• Class labels were grouped to train a set of binary models, final prediction was assigned according to the highest probability label.
Soyben crop mask at May 30, 2018
Soybean crop mask at June 14, 2018
The Bayesian model considers 10 years of historical data and estimates periodic crop variation at test area level. The model comprises a pattern based component to account for longer sequences variation.
NEBRASKA – Bayesian model 2018
NEBRASKA – Situation at June 29, 2018
- Average difference between predicted and reference value:
  - For larger acreage: 3.2%
  - For smaller acreage: 11.5%
Summary

- Satellite time series are a powerful enabler for EO services dedicated to agriculture.
- By leveraging on these time series, e-Geos is adopting its Agrigeo platform for products generation and services provision, also carrying out research activities for enlarging and improving service portfolio.
- Together with Politecnico of Milano has been analized the computation of crop acreage over adminstrative units, with focus on its continuous update during the growing season.
- A promising solution is based on the integration of a bayesian model for the provision of crop acreage in the very early stage, with an RF model for later stages.