AI4FOOD
A DATA-FUSION AND TIME SERIES FRAMEWORK AS A SERVICE

Fuse TS: A Cloud-based Toolbox for S1 and S2 Time Series (TS) Fusion, Gap-filling and Cropland Phenology Analytics

Jochem Verrelst[1]; Bram Janssen[2]; Matic Lubej[3]; Jeroen Dries[2]; Darius Couchard[2]; Kristof Van Ticht[2]; Nejc Vesel[3]; Grega Milcinski[3]; Matias Salinero Delgado[1]; Eatidal Amin[1]; Patrick Griffiths[4]

[1] Image Processing Laboratory (IPL), University of Valencia, Spain; [2] VITO, Belgium; [3] Sinergise, Slovenia; [4] ESA-ESRIN, Italy

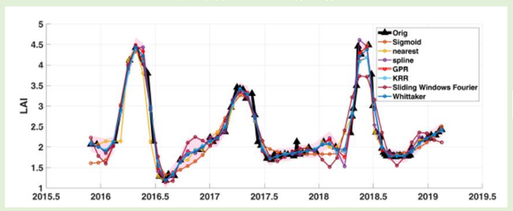
Introduction

Satellites capture an extensive amount of data daily, resulting in an ever-growing collection of Earth Observation (EO) data. Despite the availability of this data, there are still challenges when it comes to extracting relevant information from long time series (TS) data streams. The recently finished ESA's AI4FOOD project aimed to address these challenges, particularly in data fusion and advanced time series analytics within cropland monitoring applications. AI4FOOD focused on advanced Machine Learning (ML) techniques to develop new algorithms for the creation of continuous optical and radar data streams. Specifically, the project focused on the fusion of Sentinel-2 and Sentinel-1 TS data and evaluating aspects such as the predictability of time series in dynamic land environments. This collaborative effort was achieved with partners from VITO, Sinergise, and the University of Valencia. **The undertaken activities led to the development of the FuseTS toolbox – an open-source toolbox supporting users in complex data fusion and time series analytics tasks.** The goal of FuseTS is to give users the necessary tools and features to efficiently monitor cropland dynamics.

Materials & Methods

State-of-the-art algorithms for gap-filling TS data streams:

Some time series (TS) gap-filling algorithms:

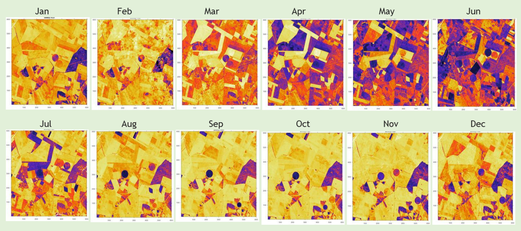


Year 2017 area (21291 pixels). Statistics calculated against images removed window (13/03-02/05)

Method	RMSE	R2	RRMSE	Processing Time	
				Total (min)	Per pixel (sec)
Next	0.46	0.77	14.93	0.26	0.0007
Polyfit	1.38	0.04	44.90	0.37	0.0010
Previous	0.82	0.52	26.70	0.26	0.0007
Spline	0.47	0.76	15.36	0.26	0.0007
Fourier1	0.82	0.26	26.65	1.27	0.0036
Fourier2	0.78	0.31	25.25	1.26	0.0036
Fourier3	1.43	0.05	46.36	2.78	0.0078
Sigmoid	0.60	0.78	19.54	758.05	2.1363
Whittaker	0.35	0.89	11.48	0.58	0.0016
ARES	0.95	0.43	30.71	31.51	0.0888
BAGTREE	0.85	0.63	27.75	148.97	0.4198
GPR	0.36	0.88	11.68	32.78	0.0924
KNNR	0.78	0.50	25.25	3.20	0.0090
KRR	0.50	0.76	16.09	5.60	0.0158
NN	0.78	0.56	25.26	1232.74	3.4740
RF1	0.50	0.80	16.11	46.91	0.1322
RF2	0.46	0.77	14.93	74.62	0.2103

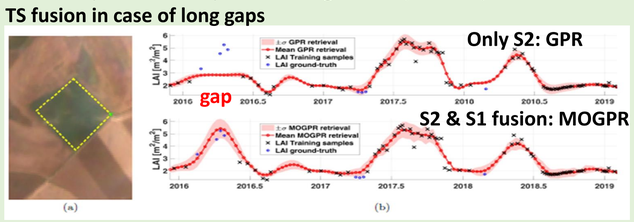
- The **Whittaker** fitting method and the ML regression algorithm **GPR** are evaluated as most accurately reconstructing time series data.
- Interpolation methods fastest but suboptimal
- Considering trade-off accuracy and runtime, then **Whittaker** is clearly the most attractive method: it runs 56 times faster than GPR in fitting data.
- **GPR**, as being designed in a Bayesian context, provides associated **uncertainties** together with the predictions.

Whittaker fast a flexible gap-filling algorithm:

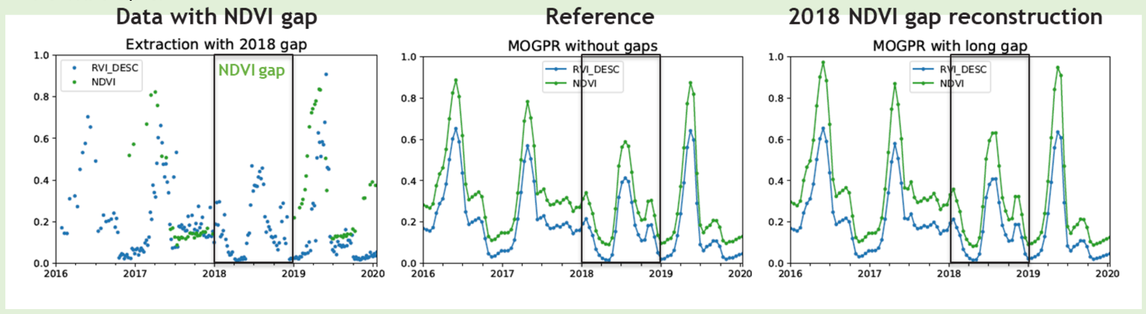


- **3 years NDVI area (600x600) in 14 min for gap-filling 218 images (<4sec/subset).**
- **Here: 2018 (mid-month)**
- **Whittaker ideal for small gaps in time. In case of persistent cloud cover, fusion with S1 data is required.**

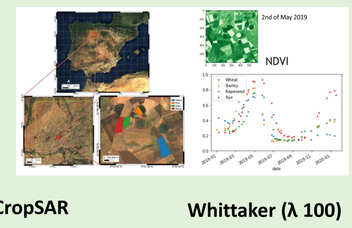
Multi-output Gaussian processes regression (MOGPR) for S2 and S1 TS fusion in case of long gaps



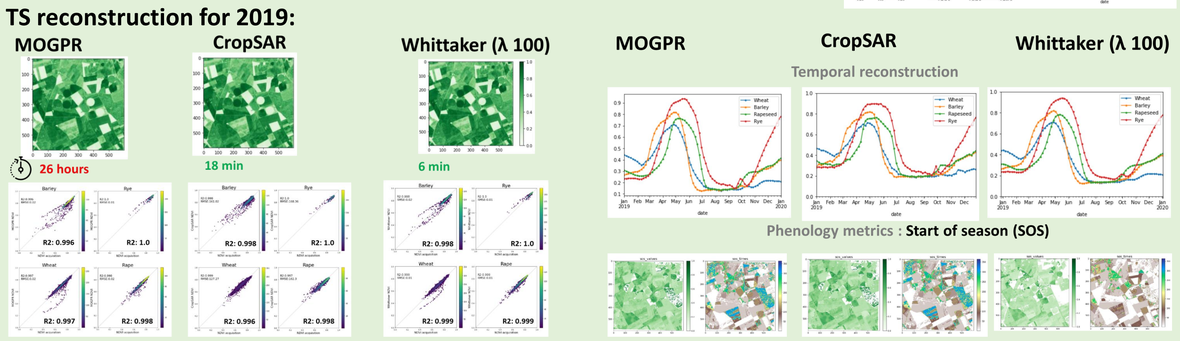
Gap in one time series is filled taking into account the common information with other sensors modeled by common GP weighted by the coefficients within the coregionalization matrices…



- MOGPR fusion able to reconstruct long gap (entire 2018) thanks to guidance of S1 (RVI_DESC) data
- Growing cycle accurately reconstructed (would not be possible with single-source fitting)
- Note that accurate reconstruction is only possible when similar temporal pattern radar data

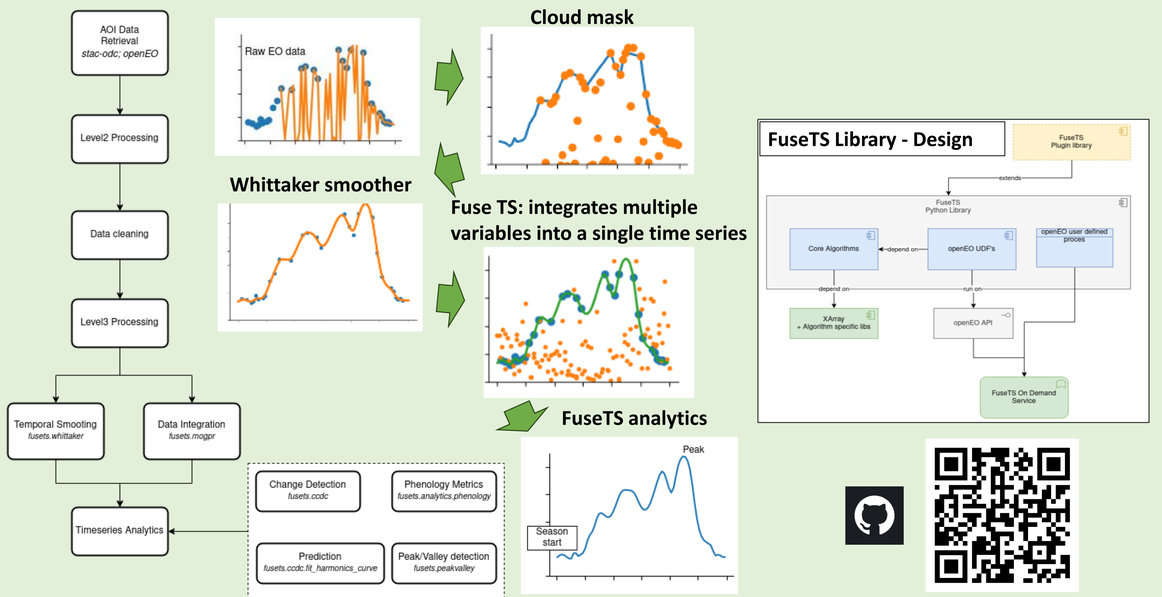


TS reconstruction for 2019:



- MOGPR, CropSAR and Whittaker were compared on their capacity of temporal reconstruction given croplands Castilla y León, Spain
- Methods provide very similar performances, **with Whittaker the fastest and slightly most accurate reconstruction against a reference NDVI image**
- Crop phenology indicators identified (SOS, POS, EOS): similar performances
- Altogether, Whittaker and CropSAR preferred given small temporal gaps….**MOGPR is only of interest in case of large time gaps (Whittaker would fail)… and having patience ☺**

FuseTS



https://github.com/Open-EO/FuseTS

FuseTS Library:

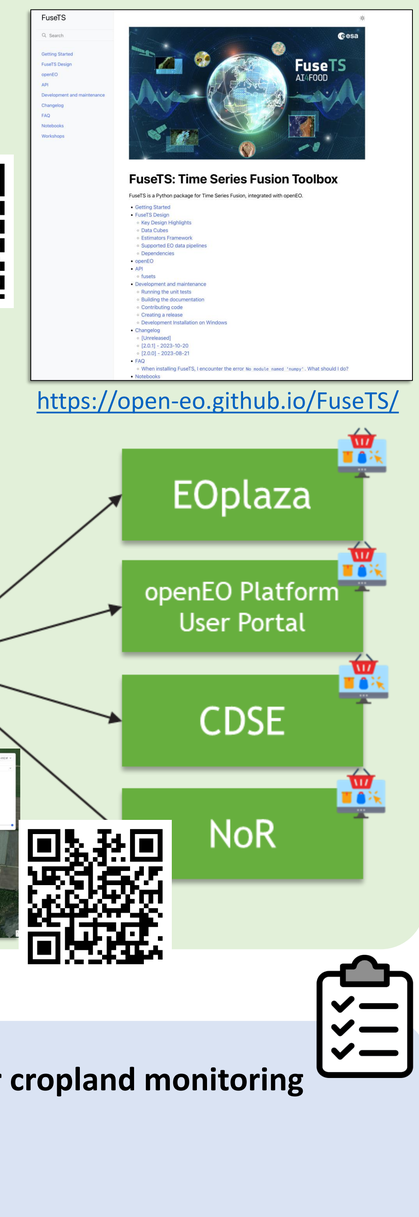
open-source GitHub repository

Data fusion algorithms:

- Whittaker (S2 TS)
- MOGPR (fusion S1 & S2 TS)
- CropSAR (fusion S1 & S2 TS)

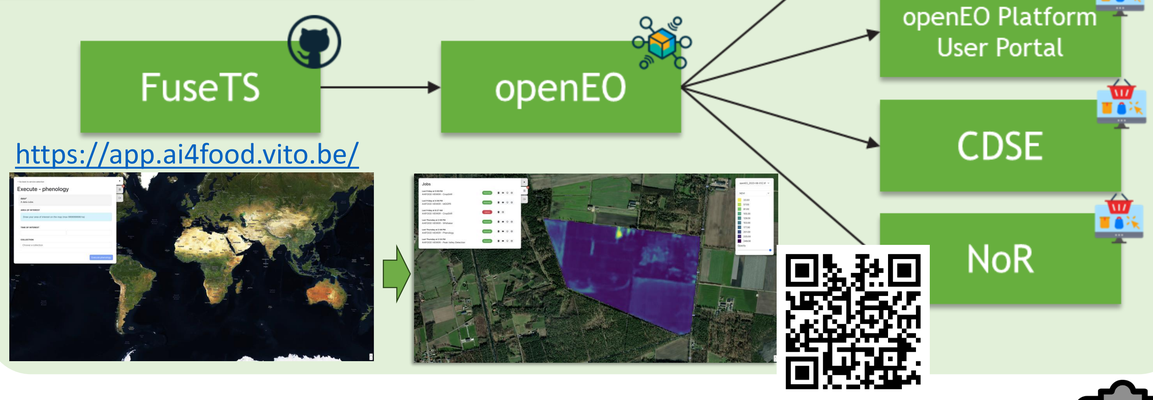
Time series analytics:

- Peak valley detection
- Phenology



https://open-eo.github.io/FuseTS/

FuseTS Service Offering:



Conclusions

- **AI4FOOD demonstrated a user-centric, end-to-end approach for advancing the use of Sentinel-1 and Sentinel-2 time series (TS) data for cropland monitoring**
- Effective TS methods (Whittaker, MOGPR, CropSAR) for filling gaps in single/multi-source data.
- Successful design and implementation of **FuseTS library on GitHub for cloud processing**
- Seamless access through **openEO** improves utilization as on-demand service.