Remote sensing for Agricultural Statistics: Perspectives with new sensors.

Raúl López-Lozano
Joint Research Centre of the European Commission, Ispra, Italy – Raul.lopez@jrc.ec.europa.eu

Javier Gallego
Joint Research Centre of the European Commission, Ispra, Italy – Javier.Gallego@jrc.ec.europa.eu

Abstract

This paper discusses the suitability of existing and forthcoming Earth Observation (EO) platforms for agricultural statistics, describing their main strengths and limitations for operational activities of crop area estimation and crop yield forecasting at national/sub-national scale. Although very high resolution platforms developed in the last ten years offer new capabilities, compared to medium-resolution sensors, there are still cost-inefficient (satellites) and present operational limitations (drones) for their use in crop area estimation, where platforms like Landsat or the future Sentinel-2 are preferred. In crop yield forecasting, where frequent observations and long-term historic archives are needed, low-resolution sensors are still the best choice, even if in in some cases the spatial resolution is not sufficient to achieve crop-specific information.

Keywords: Earth observation, Drones, Crop area estimation, Yield forecasting.

1. Introduction

Satellite imagery has been presented since the early times of remote sensing as the element that would bring the big revolution in agricultural statistics, substituting field observations (Mac Donald and Hall, 1980) both for crop area and yield estimation and forecasting. Practical experience has shown more limitations than initially expected. A very large number of pilot projects have been carried out. Some of them remain unpublished partly because of a low scientific consistency. In many cases they limit themselves to image classification and area estimation by pixel counting, an approach that has been shown to have a high risk of bias (GEOSS, 2009, Carfagna and Gallego 2005, Stehman 2013).

In the last ten years more advanced platforms open new possibilities for applications. In parallel, new EO programs are consolidated, as the European Copernicus Programme (http://www.copernicus.eu/), to provide solutions in the application of remote sensing to agriculture, environment and other fields.

This paper illustrates the suitability of current and forthcoming platforms in the domain of agricultural statistics for both crop area estimation and yield forecasting. Technical requirements are identified for both activities, based on available studies. Strengths and limitations in terms of spatial resolution, cost, and revisit frequency of the different EO platforms are discussed.

2. Existing methodologies for crop area estimation and yield forecasting using remote sensing

2.1 Area estimation

Satellite images can be used for crop area estimation in several ways:

- Building the sampling frame, in particular for stratification.
- As support documents for field surveys
- As covariate to apply regression, calibration or small area estimators.
- Only exceptionally photo-interpretation or classified images can be directly used for crop area estimation. This happens when the targeted crop can be extremely well identified (paddy rice) or when field survey is very problematic, as it happens in North Korea or Somalia, Colombia (coca) or Afghanistan (poppy).

In virtually all cases a nearly-full coverage is required. This is the major requirement. Examples in which sampling satellite images for crop area estimation can be cost-efficient are very rare, because a sampling plan with units adapted to the size of satellite images is very inefficient, unless the spatial correlation is close to zero (Gallego and Stibig, 2013)

2.2 Yield forecasting

Satellite EO systems can support operational activities in crop monitoring and yield forecasting providing spectral vegetation indices related to plant status (e.g., the Normalized Difference Vegetation Index, NDVI) or vegetation biophysical parameters such as the leaf area index (LAI) or the fraction of absorbed photosynthetically active radiation (fAPAR) through the application of radiative transfer modelling to the top-of-canopy reflectance provided by EO.

The existing link between vegetation indices/biophysical products and actual crop yield is often exploited to construct (semi-)empirical regressions between aggregated EO indicators and yields from official statistics (Balaghi et al., 2008; Lopez-Lozano et al. 2015). An example is given in Figure 1.

Alternative methods are based on data assimilation of EO indicators into crop models at regional scale. This implies updating dynamically crop model parameters to make leaf area dynamics simulated by the crop model match the observed EO indicators. Although successfully applied to into studies at large scale studies (at field, or experimental site level) few studies have attempted to implement assimilation methodologies at regional scale.

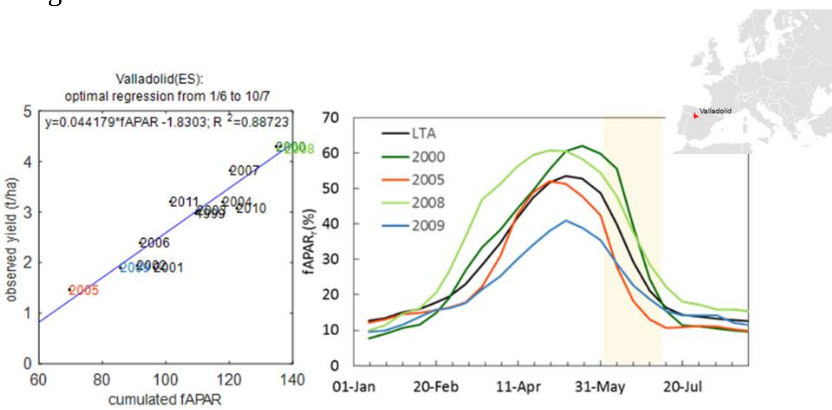

Figure 1. Empirical regression between cumulative fAPAR at regional level during crop grain-filling and official barley yield statistics in Valladolid (Spain). Extracted from Lopez-Lozano et al., 2015.

3. Technical requirements for remote sensing platforms

3.1 Revisit period

The revisit period is the time between two successive acquisitions of the same area. It indicates how frequently usable (cloud-free, homogeneous geometric conditions) images of a given remote sensing platform can be available. The revisiting time is mainly determined by the sensor field of view (FOV) that is translated into a footprint width. The overlapping area between contiguous N-S paths and, therefore, how frequently the same site can be observed is a consequence of the satellite swath and the latitude.

Digital image classification of satellite images in agriculture mostly identifies crops –or crop groups– by their different phenology. Several images along the growing season improve the ability to differentiate winter crops –with maximum activity in early spring and senescence in summer– from summer crops –sown in spring and reaching the peak of green biomass formation in summer.

For crop status monitoring and yield forecasting EO methods require continuous monitoring of vegetation canopy to construct time-series (as in Figure 1). Most operating platforms produce 10-days synthesis products: single values averaged over 10-days windows after eliminating non valid observations (cloud coverage, non-optimal acquisition geometry…). At least one valid observation every 10 days is thus a minimum threshold for crop yield forecasting purposes.

3.2 Spatial resolution

The spatial resolution of the images determines the size of objects that can be detected. It mainly depends on the optics of the sensor and can be very heterogeneous: few centimetres in unmanned air vehicles (UAV) up to several kilometres of some satellite platforms.

In crop area applications the use of satellite imagery for digital classification generally requires a spatial resolution of the images below the average plot size.

In crop monitoring and yield forecasting spatial resolution directly affects how crop-specific is the information on crop vegetative status provided by the EO indicators. This is crucial when the purpose is monitoring non-dominant crops for which specific EO time-series have to be identified. The capability of a platform to provide crop-specific information also depends on the difference between the average field size and the spatial resolution. Nevertheless, the requirements are slightly less constraining, since the purpose is only selecting a set of pixels containing the same crop, and aggregating the information at administrative unit level (national, sub-national...). The minimum spatial resolution depends also on the fragmentation of the agricultural landscape: Duveiller et al., (2010) considered 250 meters adequate for European and African croplands, whereas in China these requirements would be higher (around 100m).

3.3 Spectral resolution

Spectral resolution refers to the number and width of the spectral bands observed by an instrument. Most applications on crop area estimation are relying on VNIR (visible and near infrared) channels, which permit to identify healthy vegetation through the different spectral response of plant tissues in the near infrared (scattering) and visible (absorption by photosynthetic pigments). This identifies the periods of high and low photosynthetic activity to distinguish between crops with different phenological cycles. Some crops have a more specific response; for example paddy rice thanks to the low reflectance of water, or rapeseed if caught in the flowering period.

For crop monitoring and yield forecasting VNIR is the minimum requirement, as vegetation indices and many EO-derived products are based on near infrared and red wavelengths. Additional regions of the electromagnetic spectrum as the red-edge (transition between visible and near infrared) or the short-wave infrared provide access to additional biophysical parameters as the crop chlorophyll content, or leaf moisture content (Zarco-Tejada et al., 2005) but are seldom operationally used.

3.4 Image size

Traditionally, image size has been determined by the footprint width (eg. Landsat or SPOT 4). Other platforms (in particular VHR, see Section 4 for further details) can orientate the sensor towards a specific ground target (off-nadir). The FOV and footprint width are not, in these cases, an indicator of the achievable image size, as the sensor cannot observe simultaneously all the footprint width. It is better to consider the swath, defined as the strip width that the sensor is capable to observe in a single scan. For programmable VHR sensors the image size can be also larger than the nominal swath as the instrument permits to mosaic images acquired over multiple strips in a single pass orienting adequately the instrument optics. This is useful for small pilot areas, but makes little difference for operational use in large regions when complete image coverage is required.

The requirements in terms of image size depend on the extension of the region of interest. If the nominal swath permits it, the best possible option is acquiring the region into a single scene. That avoids the reprocessing of image mosaics to compensate for brightness differences associated with a variation in illumination/observation geometries among individual scenes.

4. Discussion: suitability of available and forthcoming platforms for agricultural statistics.

Table 1 presents a selection of existing and forthcoming remote sensing platforms of interest for vegetation monitoring. This selection includes platforms part of national and international EO programs like the Copernicus EO programme supported by the EC (Sentinel 2 and 3, PROBA-V) or NASA EO programme (MODIS, Landsat 8), and different public and private initiatives for EO. In this section the suitability of the different platforms for area estimation and crop monitoring are discussed.

4.1 Crop area estimation

For crop area estimation the key requirement is the ability to (nearly-) completely cover the target region at a low cost, possibly several times in the season. This excludes VHR images, such as SPOT 6/7, PLEIADES, WorldView and GeoEye. RapidEye might reach cost-efficiency in specific cases.

Another key parameter is the spatial resolution. A useful rule of thumb is that at least half of the pixels should be pure (not shared by different fields or landscape patches). In some landscapes with very large fields this can be achieved with moderate resolution (250-300 m), but in general medium resolution (20-30 m) is required. When the average plot size is between 0.1 and 0.5 ha, sensors providing 10 m resolution can make the difference (Sentinel-2). The revisit period is important, but not critical. Image classification improves if based on 3-4 images in the growing period, but it is not clear in which cases the improvement is enough to make up for cost increase.

Table 1. Selection of current and forthcoming remote sensing platforms for vegetation monitoring. * Sentinel 2 and 3 will be launched in 2015 and 2016.

PLATFORM-SENSOR	nr. of satellites	Spatial resolution (m)	Footprint width (km)	Nominal swath (km)	Revisit period (days)	Spectral bands	Price (€/km2)
PROBA-V HTTP://PROBA-V.VGT.VITO.BE/	1	300 (VNIR) 1,000 (SWIR)	2,285	2285	1	3 VNIR+ 1 SWIR	0
SENTINEL 3-OLCI* https://earth.esa.int/web/guest/missions/esa-future-missions/sentinel-3	2	300	1,270	1270	1	21 VNIR	0
TERRA/ACQUA-MODIS HTTP://MODIS.GSFC.NASA.GOV/	2	250, 500, 1,000	2,330	2330	1	36 (VNIR, SWIR,TIR)	0
METOP-AVHRR HTTP://WWW.EUMETSAT.INT/WEBSITE/HOME/SATELLITES/CURRENTSATELLITES/METOP/INDEX.HTML	2	1,000	2,900	2900	1	6 (VNIR, SWIR, TIR)	0
LANDSAT 8-OLI http://landsat.usgs.gov/landsat8.php	1	30	185	185	16	11 (VNIR, SWIR, TIR)	0
SENTINEL 2-MSI* https://earth.esa.int/web/guest/missions/esa-future-missions/sentinel-2	2	10 (VNIR) 20 (NIR-SWIR) 60 (Atm. corr)	290	290	5	13 (VNIR, SWIR)	0
DEIMOS/UK –DMC HTTP://WWW.DEIMOS-IMAGING.COM/	2	22	620	620	1-2	3-VNIR	0.13
SPOT 6/7-NAOMI HTTP://WWW.GEO-AIRBUSDS.COM/EN/147-SPOT-6-7-SATELLITE-IMAGERY	2	6	800	60	1-2	4VNIR	5.7
PLEIADES HTTP://WWW.GEO-AIRBUSDS.COM/EN/52-PLEIADES-VERY-HIGH-RESOLUTION-SATELLITE-IMAGERY	2	2	800	20	1-2	4VNIR	11.5
GEO-EYE HTTP://WWW.SATIMAGINGCORP.COM/SATELLITE-SENSORS/GEOEYE-1/	2	2	780	15	1-2	4VNIR	15.5
WORLDVIEW HTTP://WWW.SATIMAGINGCORP.COM/SATELLITE-SENSORS/WORLDVIEW-1/	3	2	780	16	1	8VNIR	15.5
RAPIDEYE HTTP://BLACKBRIDGE.COM/RAPIDEYE/	5	5	77	77	1	5 VNIR	1.13
UAVs	-	<1	3	0.2	-	Config.	-

Medium-resolution sensors are the reasonable solution in most cases: platforms like Landsat, Deimos, or the forthcoming Sentinel 2. The spatial resolution is sufficient, and they may become operational in combination with an intensive field survey. In cloudy areas a traditional drawback of medium-resolution platforms has been the revisiting period (16 days for the Landsat family). This is improved with large swath platforms as DEIMOS, or Sentinel 2 that will decrease the revisit period to 5 days when the 2 satellites of the constellation will be fully operational. Moreover, NASA and ESA have free distribution data policy for Landsat and the Sentinels and computing cost has been significantly reduced. This is probably the major revolution for the usability of EO in agricultural statistics.
Aerial photographs, both from small traditional aircrafts or from UAVs, better known as drones, provide an interesting alternative, A limitation for drones is their on-flight stability, although techniques producing image mosaics from a zig-zag flight seem to have a reasonably good geometry and a spatial

resolution around 5 cm or even finer. Such resolution may significantly change the way images are used to identify crops. While classical EO uses reflectance to identify crops and discriminating crops with similar phenological calendar is very difficult, 5 cm images can give additional information on the distance between rows and in some cases the shape of the leaves. New UAV sensors provide images with 2 cm resolution. In the example of Figure 2 we can see two maize fields easy to identify; the third field is more difficult, but a local expert from Malawi clearly identifies it as ground nuts. The ability of such images to identify specific crops still needs to be assessed more in depth. Unfortunately in many countries regulatory restrictions limit UAVs flights to stay within the direct view field of the operator, usually not more than 1 km. Small aircrafts with a human pilot may be more adapted for such sampling units. We have made a few simulations using as pseudo-truth the farmers' declarations in the Netherlands in 2012. Long and thin sampling units (eg. 100 km x 100 m) are around three times as efficient as square sampling units of the same area, because the intra-cluster correlation is lower.

Figure 2. Example of drone image with 2 cm resolution in Malawi in a test run by the ITA consortium.

4.2Crop monitoring and yield forecasting

As mentioned in Section 2.1, monitoring vegetation activity and forecasting crop yield generally requires frequent acquisitions to have, at least, 1 valid observation every 10 days. That makes low-resolution platforms like PROBA-V, METOP/NOAA-AVHRR, MODIS or Sentinel 3, with an almost daily global coverage, more suitable. In medium-resolution platforms like Landsat, the revisit period of 16 days is not enough to produce reliable time-series, especially in cloudy areas, whereas VHR platforms are not capable of covering large areas frequently (narrow swath) and their cost is elevated.

The spatial resolution requirements are not particularly high, but a resolution of 250-300 is needed to access crop-specific information, which places MODIS, PROBA-V and Sentinel 3 in full resolution mode as better options than 1-km platforms as METOP/NOAA-AVHRR. Different works have been conducted in crop monitoring and yield forecasting with SPOT-VEGETATION data at 1-km, but it's only possible to monitor dominant crops (Lopez-Lozano, 2015). Sentinel 2 would become a major step ahead, as it increases significantly the spatial resolution of low-resolution platforms, with an acceptable revisit frequency of 5 days, and free access policy. This would offer, for the first time, the possibility to incorporate continuous, high-resolution data on crop status into operational crop yield forecasting systems.

Another important requirement to be considered is having a long historical archive of images. Crop yield forecasting with remote sensing methods requires understanding how the inter-annual variability of EO indicators and observed yields are actually correlated. To do so, a minimum number of years in the archive must be available, as indicators form different platforms are not directly comparable. This is a point in favour of MODIS and other platforms as NOAA-AVHHR, with more than 15 years of continuous observations. The Copernicus programme for EO of the European Commission emphasizes the production of a long-term data archive of consistent EO products joining the archives of SPOT-

VEGETATION (1998-2014), PROBA-V (from 2014) and Sentinel 3-OLCI (from 2016) through dedicated inter-operable algorithms to produce biophysical indicators (Baret et al., 2013).

5. Conclusions

The present paper summarized how suitable the different satellite Earth Observation (EO) sensors are as a tool for agricultural statistics in two main domains: crop area estimation and crop yield forecasting. The main requirements in terms of both spatial and spectral resolution, revisit frequency and scene size are presented, and confronted with the capabilities of a selection for the current and forthcoming platforms. Medium resolution platforms –like Landsat 8 or the future Sentinel 2– combined with field visits are the currently the most suitable choice for area estimation, compared to very high resolution platforms or unmanned air vehicles, due to their good-balance between spatial resolution, revisit time and free access policy. For crop yield forecasting, requiring frequent observations, low resolution platforms/sensors like MODIS or PROBA-V are still the best options to monitor continuously large areas. Nevertheless, the forthcoming constellation Sentinel 2 will constitute a major step ahead combining high spatial resolution with important capabilities to acquire frequent images over large areas.

References

Balaghi,R. Tychon,B., Eerens, H., Jlibene, M. (2008), Empirical regression models using NDVI, rainfall and temperature data for the early prediction of wheat grain yields in Morocco. *Int. J. Appl. Earth Obs. Geoinf.*,10: 438–452

Baret, F. ; Weiss, M. ; Lacaze, R. ; Camacho, F. ; Makhmara, H. ; Pacholczyk, P. ; Smets, B. (2013): GEOV1: LAI, FAPAR Essential Climate Variables and FCover global times series capitalizing over existing products. Part1: Principles of development and production. *Remote Sensing of Environment*, 137: 299–309.

Carfagna E., Gallego F.J., (2005), The use of remote sensing in agricultural statistics, *International Statistical Review* 73 (3), 389-404.

Duveiller, G. and Defourny, P. (2010): A conceptual framework to define the spatial resolution requirements for agricultural monitoring using remote sensing. *Remote Sensing of Environment*, 114: 2637–2650.

Gallego F.J., Stibig, H.J., 2013, Area estimation from a sample of satellite images: the impact of stratification on the clustering efficiency. *Journal of Applied Earth Observation and Geoinformation* 22 pp. 139-146

GEOSS (2009). *Best practices for crop area estimation with Remote Sensing*. http://www.earthobservations.org/documents/cop/ag_gams/GEOSS%20best%20practices%20area%20estimation%20final.pdf

López-Lozano, R, Duveiller, G., Seguini, L., Meroni, M., García-Condado, S., Hooker, J., Leo, O. and Baruth, B. (2015): Towards regional grain yield forecasting with 1 km-resolution EO biophysical products: Strengths and limitations at pan-European level. *Agricultural and Forest Meteorology*, 206:12-32

Mac Donald R.B., Hall F.G., 1980, Global crop forecasting. *Science*, 208, 670-679.

Stehman, S. V. (2013). Estimating area from an accuracy assessment error matrix. *Remote Sensing of Environment*, *132*, 202-211.

Zarco-Tejada, P., Berjon, A. Lopez-Lozano, R., Miller, J. R., Martin, P., Cachorro, V., Gonzalez, M. R. and de Frutos, A. (2005): Assessing vineyard condition with hyperspectral indices: Leaf and canopy reflectance simulation in a row-structured discontinuous canopy. *Remote Sensing of Environment*, 99: 271–287.