D2.3 Data products validation report

Wp2: Earth Observation data products and services

Oscar Rosario Belfiore, Salvatore Falanga Bolognesi, Guido D'Urso, Carlo De Michele (ARIESPACE), Alfonso Calera Belmonte (AgriSat), Violeta Domnica Poenaru (ROSA), Stylianos Kotsopoulos (AgroApps)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 130709.

Disclaimer

Copyright message

Document Information

Grant Agreement Number	130709	Acronym		DIANA
Full Title	Detection and Integrated Assessment of Non-authorised water Abstractions using EO			
Horizon 2020 Call	EO-1-2016: Downstream applications			
Type of Action	Innovation Action			
Start Date	1st January 2017	**Duration**		36 Months
Project URL	http://diana-h2020.eu			
Document URL	-			
EU Project Officer	Iulia SIMON			
Project Coordinator	Polimachi Simeonidou, Anna Osann			
Deliverable	D2.3: Data products validation report			
Work Package	WP2 – Earth observation data products and services			
Date of Delivery	**Contractual**	M24	**Actual**	M24
Nature	R – Report	**Dissemination Level**		Public
Lead Beneficiary	ARIESPACE			
Lead Author	Carlo De Michele	**Email**	carlo.demichele@ariespace.com	
	ARIESPACE	**Phone**	0039 081 195 64 282	
Other authors	Oscar Rosario Belfiore, Salvatore Falanga Bolognesi, Guido D'Urso, (ARIESPACE), Alfonso Calera Belmonte (AgriSat), Violeta Domnica Poenaru (ROSA), Stylianos Kotsopoulos (AgroApps)			
Reviewer(s)	AgroApps			
Keywords	Data products, Data validation, Assessment, Accuracy, Crosscheck.			

Document History

Version	Issue Date	Stage	Changes	Contributor
0.0	19.11.18	DRAFT		ARIESPACE
0.1	12.12.18	DRAFT	Italian Pilot area	ARIESPACE
0.2	18.12.18	DRAFT	Spanish Pilot area	AGRISAT
0.3	19.12.18	DRAFT	Verification of the Meteorological Products	AGROAPPS
0.4	20.12.18	DRAFT	Romanian Pilot area	ROSA
0.5	21.12.18	DRAFT	Internal Review	ARIESPACE
0.6	26.12.18	DRAFT	Review	AGROAPPS
0.7	28.12.18	FINAL	Minor revision	ARIESPACE

Table of Contents

Table of Figures

Table of Tables

Executive summary

Earth Observation (EO) products provide key information about water abstraction through the identification of irrigation activities and the estimation of abstracted/consumed water volumes. However, like for in-situ non-EO-based approaches, this information needs to be compared to legal reference data to be able actually to detect possible cases of illegal irrigation.

The overall aim of WP2 is to implement an EO methodology and production line for detecting irrigated areas and estimating irrigation requirements and possibly actual water use. It includes data products and services based on a combination of EO data acquired by various platforms as well as meteorological and complementary data derived from different data sources, especially from pilot areas.

This document aims to describe the validation process of the data products provided by DIANA.

For each Diana data product –i) Crop classification, ii) Maps of irrigated area, iii) Crop evapotranspiration, iv) Net Irrigation requirements, v) Gross irrigation requirements, vi) Meteorological products - the validation and evaluation were performed.

From the methodological point of view, data products listed in this document were defined considering "*Users' and stakeholders' Requirements Analysis*" and *"Data requirements manual"* described respectively in the deliverable D1.1 and D2.2. This approach ensures to meet the realities in which to apply DIANA services addressing the users' requirements properly regarding spatial, temporal and spectral resolution and by extending the operational capabilities of the platform offered.

It is worth to note that considering a wide range of data – acquired by different sources - the validation process could be adapted and consequently changed during the running of services in each pilot area.

Bearing this in mind, the information contained in this document will represent an easy reference guide about the data product validation adapted to each different context.

1 Introduction

1.1 Purpose and scope of the document

How good is a dataset?

According to Loew et al. (2017)[1], validating the uncertainties in satellite data products is a very challenging task. In this context, can be found in literature a wide range of technical approaches - as well as methods and terminology- used by the Earth Observation (EO) communities.

The primary goal of DIANA is to provide users with tools that help them fulfil their mission and perform their daily routine operations more easily and better, and then quantifying the quality of the data products and services, is a fundamental objective to achieve.

This document details the results on the validation of the algorithms and methodologies considered to deliver the products, tools and services provided by the DIANA project.

These validation results are based on a combination of Earth Observation (EO), meteorological, modelled and *in-situ* data, collected for the different pilot areas considered in the project (Italy, Spain and Romania).

The algorithms and methodologies are described in detail in the "EO Methodology for DIANA services" (D2.1).

In the context of the data products validation, this report is the first version of an evolving document. This early version will be completed with the "*Data products validation report (2)*" (D2.4) scheduled for the M36.

In this document are also reported the preliminary results of the "*Diana Benchmark Exercise"*. A specific task delineated during technical meetings, which aims to improve the detection of irrigated areas process, taking into account the soil moisture data estimated from EO data. More details about the approaches and the algorithms tested are reported in *Section 5*.

[1] Loew, A., Bell, W., Brocca, L., Bulgin, C. E., Burdanowitz, J., Calbet, X. & Kinzel, J. (2017). Validation practices for satellite-based Earth observation data across communities. Reviews of Geophysics, 55(3), 779-817.

2 Italian Pilot area: Data products validation for the year 2018

2.1 Map of Irrigated Areas

For the irrigation season 2018, the detection of irrigated areas was performed using a supervised "multi-temporal classification" based on a time series of Vegetation Indices (Vis). The classification process based on temporal pattern recognition exploits the captured differences from the canopy on the VI to assign each pixel to a vegetation class. These classes need to be defined based on field inspections and knowledge about crop phenology and crop management. The proposed methodology is founded on the assumption that the hydrologic deficit typical of the semi-arid environments, as for the Mediterranean basin, the only detectable crops are those that are irrigated. In order to follow the phenological development of crops in the irrigation season, the considered approach is based on the use of a time series of the multispectral satellite images, opportunely processed in a semi-automatic workflow.

This application is based on the utilisation of data from the Multispectral Instrument (MSI) on board of Sentinel 2A & 2B platforms. To perform the irrigated areas detection a time series of S2A & B was selected. In detail, considering a cloud cover less than 20%, 44 images captured for the year 2018 were chosen (Table 1).

Tile	Granule	n° S2 images per month	Acquisition Time (yyyy-mm-dd)
T33	**TVF**	5	2018/01/**13-18-21-26-28**
		2	2018/02/**15-17**
		1	2018/03/**24**
		4	2018/04/**03-08-21-26**
		3	2018/05/**18-26-31**
		5	2018/06/**02-10-17-25-30**
		7	2018/07/**02-05-10-20-25-27-30**
		9	2018/08/**01-04-06-09-11-19-21-24-29**
		8	2018/09/**05-08-10-15-23-25-28-30**

Table 1 - S2 images used in the irrigated areas detection process

For EO applications, based on the multi-temporal approach (i.e. change detection, land surface phenology, land cover classification, etc.) an atmospheric correction is one of the most important steps, with the aim to convert the original digital data, generally in Digital Number (DN), into the

specific physical magnitudes (Caselles & Lopez Garcia, 1989)[2]. In other words, the surface reflectance for each considered input data is required.

Atmospheric and topographic corrections can be applied to satellite images before classification in order to normalise radiance and digital number (DN) values (Young et al., 2017)[3]. Atmospheric correction aims at determining the true surface reflectance values by removing the atmospheric effects resulting from the scattering and absorption of electromagnetic radiation by gases and aerosols when passing through the atmosphere to the satellite sensor (Hadjimitsis et al., 2010)[4]. Previous studies have reported that atmospheric correction is one of the most important corrections, especially when working with multiple scenes at different temporal scales (Song et al., 2001[5]; Vanonckelen et al., 2013)[6]. Topographic correction is the process of reducing the variation of image values resulting from differences in surface terrain illumination and shadows cast during image acquisition (Vanonckelen et al., 2013)[6]; these effects are especially common in rugged or mountainous areas. Studies have reported various effects of topographic correction. For example,

Vanonckelen et al., (2013)[6] reported that topographic correction improved classification accuracy from 78 to 89% in mountainous areas, while other studies showed that topographic correction might not significantly improve accuracy in land cover classification routines (Carpenter et al., 1999[7]; Goslee, 2012[8]; Mitri & Gitas, 2004[9]; Zhang et al., 2011[10]).

---

2 Caselles, V., Lopez Garcia, M. J. (1989). An alternative simple approach to estimate atmospheric correction in multitemporal studies. International Journal of Remote Sensing, 10(6), 1127–1134. http://doi.org/10.1080/01431168908903951.

3 Young, N.E., Anderson, R.S., Chignell, S.M., Vorster, A.G., Lawrence, R., Evangelista, P.H. (2017). A survival guide to Landsat preprocessing. Ecology 98 (4), 920–932. http://dx.doi.org/10.1002/ecy.1730.

4 Hadjimitsis, D. G., Papadavid, G., Agapiou, A., Themistocleous, K., Hadjimitsis, M., Retalis, A., Michaelides, S., Chrysoulakis, N., Toulios, L., Clayton, C. R. I. (2010). Atmospheric correction for satellite remotely sensed data intended for agricultural applications: impact on vegetation indices. Nat. Hazards Earth Syst. Sci. 10 (1), 89–95. http://dx.doi.org/10.5194/nhess-10-89-2010.

5 Song, C., Woodcock, C.E., Seto, K.C., Lenney, M.P., Macomber, S.A. (2001). Classification and change detection using Landsat TM data: when and how to correct atmospheric effects? Remote Sens. Environ. 75 (2), 230–244. http://dx.doi.org/10.1016/S0034-4257(00)00169-3.

6 Vanonckelen, S., Lhermitte, S., Van Rompaey, A. (2013). The effect of atmospheric and topographic correction methods on land cover classification accuracy. Int. J. Appl. Earth Obs. Geoinform. 24, 9–21. http://dx.doi.org/10.1016/j.jag.2013.02.003.

7 Carpenter, G.A., Gopal, S., Macomber, S., Martens, S., Woodcock, C.E. (1999). A neural network method for mixture estimation for vegetation mapping. Remote Sens. Environ. 70 (2), 138–152. http://dx.doi.org/10.1016/S0034-4257(99)00027-9.

8 Goslee, S.C. (2012). Topographic corrections of satellite data for regional monitoring. Photogramm. Eng. Remote Sens. 78 (9), 973–981. http://dx.doi.org/10.14358/PERS.78.9.973.

9 Mitri, G., Gitas, I. (2004). A performance evaluation of a burned area object-based classification model when applied to topographically and non-topographically corrected TM imagery. Int. J. Remote Sens. 25 (14), 2863–2870. http://dx.doi.org/10.1080/01431160410001688321.

10 Zhang, Z., De Wulf, R. R., Van Coillie, F. M., Verbeke, L. P., De Clercq, E. M., Ou, X. (2011). Influence of different topographic correction strategies on mountain vegetation classification accuracy in the Lancang Watershed. China. J. Appl. Remote Sens. 5 (1), 053512. http://dx.doi.org/10.1117/1.3569124. T.N. Carlson, R.R. Gillies, E.M. Perry.

In this application, the candidate images were selected considering a cloud coverage less than the 20 % of the full scene. Nevertheless, with the aim to achieve a correct interpolation of the NDVI time-series, the clouds removing and gap filling technique is executed, using the Whittaker smoother developed by Eilers (2003)[11] (a weighted spline with second-order finite difference penalty) and a smoothing parameter that preserves fidelity to data rather than data smoothness (Atzberger & Eilers, 2011)[12]. The smoother the result, the more it will deviate from the input data. A balanced combination of the two goals is the sum (Q):

$$Q = S + \lambda R \quad (1)$$

$$S = \sum_t (VI_t - VI_t^*)^2 \quad (2)$$

$$S = \sum_t (VI_t - 3VI_{t-1}^* + 3VI_{t-2}^* - 3VI_{t-3}^*)^2 \quad (3)$$

The lack of fit to the data S (Eq. 2) is measured as the usual sum of squares of differences. The roughness of the smoothed curve R (Eq. 3) is expressed here as third order differences. The smoothing parameter (λ) is chosen by the user. Penalised least squares aim to find the series VI(t) that minimises Q (Eq. 1). The larger the parameter λ, the greater is the influence of R on the goal Q and the smoother will be VI (t) (at the cost of the degradation of the fit).

The Whittaker smoother is applied independently for each raster and each pixel to produce smoothed and gap-filled VI. Diversely from other methods, the Whittaker adapt the filtering to each single pixel within the image, thus providing the maximum adaptability for the image itself. Whittaker is based on two assumptions (Chen et al.,2004)[13]:

i) That the time series of vegetation index follows an annual cycle of growth and decline as the index is primarily related to vegetation density and plant vigour.

ii) That clouds and poor atmospheric conditions produce a negative bias in the vegetation index values, requiring that sudden drops in vegetation index, which are not compatible with the gradual process of vegetation change, are regarded as noise and will be removed. Past the filtering, we used shape constraints, based on statistical analysis of raster data, for reduce edge effect.

11 Eilers, P. H. C. (2003). A perfect smoother. Analytical Chemistry. http://doi.org/10.1021/ac034173t

12 Atzberger, C., Eilers, P. H. C. (2011). Evaluating the effectiveness of smoothing algorithms in the absence of ground reference measurements. International Journal of Remote Sensing, 32(13), 3689 3709, 2011.

13 Chen, J.; Jönsson, P.; Tamura, M.; Gu, Z.H.; Matsushita, B.; Eklundh, L. A simple method for reconstructing a high quality NDVI time series data set based on the Savitzky–Golay filter. Remote Sens. Environ. 2004, 91, 332–344

According to (Garcia-Pedrero, Gonzalo-Martin, Fonseca-Luengo, & Lillo-Saavedra, 2015)[14], to provide agricultural services based on EO data, a correct delineation of agricultural parcels is a fundament requirement, and the high-resolution satellite images and machine-learning algorithms play a key role for these purposes. Hence, in this work to detect the irrigated areas, a supervised classification was applied using as input data the NDVI time series. In detail, the performance achievable from different Machine Learning Algorithms (MLA) were tested (Kuhn & Johnson, 2013)[15]:

- Artificial Neural Network (ANN),
- Single Decision Tree (DTs),
- *k*-nearest neighbour (k-NN),
- Random Forest (RF),
- Support Vector Machine (SVM).

To train these algorithms ground truth data were considered. During the irrigation season 2018, the Sannio Alifano Consorzio staff conducted field inspections, collecting 1200 points balanced for three classes (same size of training sample) distributed in the whole area: i) Bare soil or rainfed, ii) herbaceous and iii) tree crop. Subsequently, with the aim to obtain a robust validation of the irrigated areas map, the ground truth dataset (pixels) were separated into training and test samples using random sampling stratified by class, with 25% of pixels used to train the model, and 75% of the pixel used to validate the model. Most machine-learning algorithms have user-defined parameters that may affect classification accuracy. Although default values are often suggested for these parameters, empirical testing to determine their optimum values is needed to ensure confidence that the best possible classification has been produced. The relative difficulty of running parameter optimisation for different classifiers is often cited as a major consideration in selecting an algorithm. One commonly used method is k-fold cross-validation. In this method, the training data are randomly split into k disjunct subsets (e.g. 10). The model is then run k times, each time withholding one of the subsets, which is used for validation. The results of each run are assessed using the withheld data, and the results are averaged across all k replicates. In this way, is possible to test a range of values for all combinations of the parameters empirically, and the combination that yields the best performance, commonly defined based on overall classification

14 Garcia-Pedrero, A., Gonzalo-Martin, C., Fonseca-Luengo, D., & Lillo-Saavedra, M. (2015). A GEOBIA methodology for fragmented agricultural landscapes. Remote Sensing, 7(1), 767-787.
15 Kuhn, M., & Johnson, K. (2013). Applied predictive modeling. (Vol. 26). New York: Springer.

accuracy or the kappa statistic, is selected. Parameter tuning, in which an optimal value for the parameter is estimated for classifications, was performed using 10-fold cross validation for each model. For consistency, the tuneLength parameter was set to 10 so that 10 values of each parameter were assessed. All variables were also centred and rescaled for consistency, before classification. The best parameters were used to create a model that was then applied to the validation or test data.

Following this approach, to estimate the thematic accuracy of the classification process, an error matrix was computed for each considered algorithms, considering the classical accuracy measures reported in literature, as: Producer's Accuracy (PA), User's Accuracy (UA) and Overall Accuracy (OA) (Story & Congalton, 1986)[16], (Congalton & Green, 2009)[17].

The results of the thematic accuracy validation are plotted in the following graph (Figure 1).

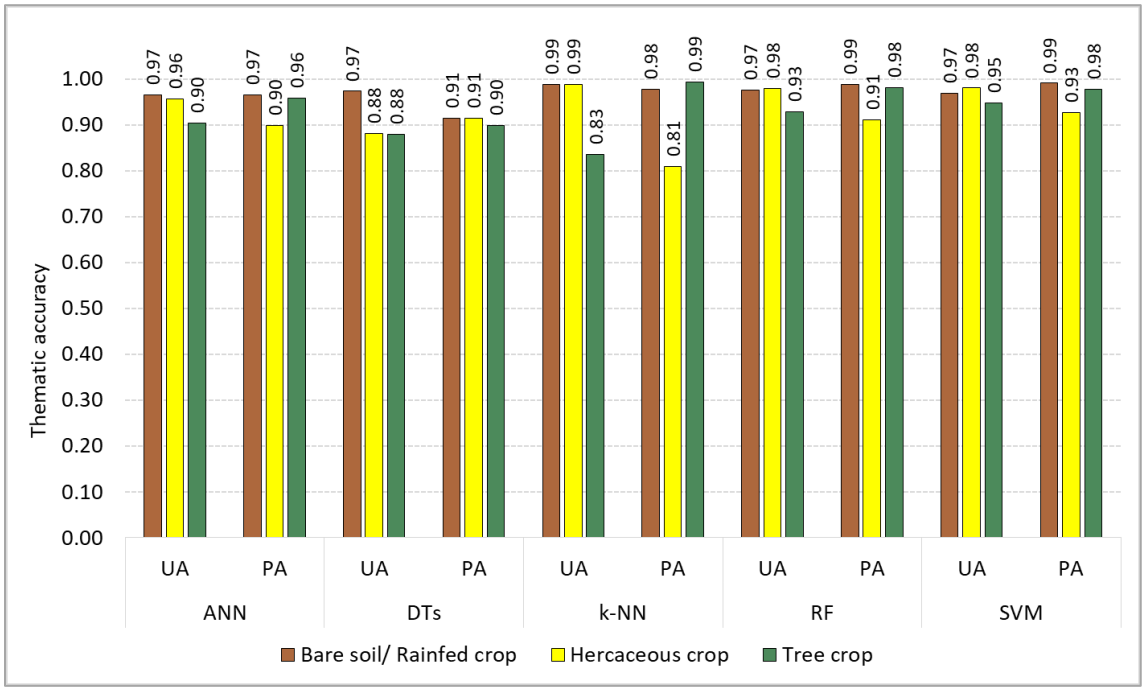


Figure 1 - User's Accuracy and Producer's Accuracy of the thematic class distinguished by the considered MLA. All numbers refer to the independent validation dataset not used during training.

The results of the accuracy assessment show high values of the accuracy measures for the SVM and RF, with a negligible difference between the two algorithms. In conclusion, considering the

16 Story, M., Congalton, R. G. (1986). Accuracy assessment: a user's perspective. Photogrammetric Engineering & Remote Sensing, 52(3), 397–399. http://doi.org/10.1111/j.1530-9290.2010.00257.x

17 Congalton, R. G., Green, K. (2009). Assessing the Accuracy of Remotely Sensed Data: Principles and Practices. The Photogrammetric Record (Vol. 2). http://doi.org/10.1111/j.1477-9730.2010.00574_2.x

results reported in Table 7, the SVM algorithm was selected to deliver the Map of irrigated areas (Figure 2) as data products on DIANA platform.

Machine Learning Algorithms	Overall Accuracy
SVM	0.97
RF	0.96
ANN	0.94
k-NN	0.93
DTs	0.91

Table 2 - Machine Learning Algorithms ranked by OA.

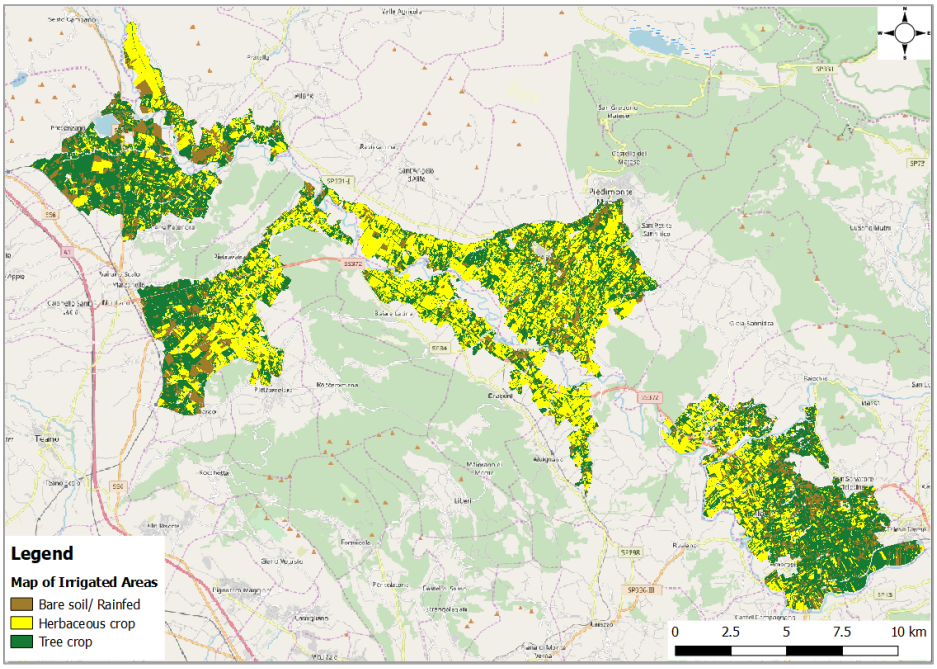


Figure 2 – Map of Irrigated areas obtained by the SVM algorithm.

2.2 Crop Evapotranspiration

The estimation of Crop Evapotranspiration (ET_p) was performed by the direct calculation based on the Penman-Monteith equation. More details are reported in the deliverable 2.1 (Paragraph 3.2).

The literature is abundant in EO - based ET models or model-variants and validations of these models in different environments, surfaces and management, confirming that EO is a mature technology ready to be transferred to operational applications in irrigation management. Several papers have demonstrated the accuracy of the methods mentioned above (Rubio et al., 2006)[18].

18 Rubio, E., Colin, J., D'Urso, G., Trezza, R., Allen, R., Calera, A., González, J., Jochum, A., Menenti, M., Tasumi, M., Kelly, C., Vuolo, F. (2006). Golden day comparison of methods to retrieve et (Kc-NDVI, Kc-analytical, MSSEBS, METRIC). AIP Conference Proceedings, 852, 193–200. http://doi.org/10.1063/1.2349344.

Comparison between different methods and micrometeorological measurements, i.e. Eddy Covariance fluxes data are shown, for example, in Rubio et al., 2006; D'Urso et al., 2010[19]. Unfortunately, in the Italian Pilot area, such measurements are not available. For this purpose, with the aim to validate the crop evapotranspiration, two approaches are followed:

- **Validation of LAI**: The Leaf Area Index represents the main input variable related to the crop development in the calculation of ET using the Penman-Monteith approach. The error on LAI propagates in the estimation of ET in a significant way. Field non-destructive measurements of LAI have been used to compare with the EO-product utilised in the calculations. In detail, during the irrigation season 2018, the field measurements of LAI have been executed using the portable canopy digital analyser (LAI-2000 Plant Canopy Analyzer, LI-COR), under conditions of diffuse illumination at sunset. The LAI value at each location is resulting as the average of 3 repetitions of 8 below canopy readings taken within a 5 m radius of the georeferenced location. An opaque cover (view cap) on the optical sensor, with an open wedge of 45°, was used to avoid the influence of neighbouring obstacles, such as the operator (Gower and Norman, 1991[20]; Li-Cor, 1992[21]). In conclusion, considering a sample of 33 values collected for herbaceous (Maize and Alfalfa) and tree crops (Apple, Peach and Hazelnut) the validation was executed (Figure 3).

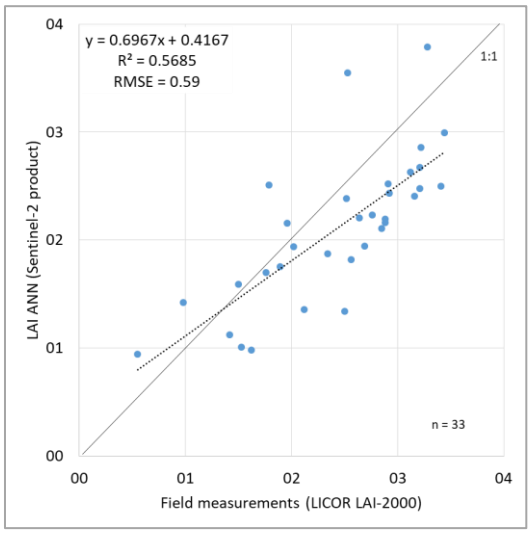


Figure 3 - Scatterplot of field LAI measured vs. Sentinel-2 LAI product.

19 D'Urso, G., Richter, K., Calera, A., Osann, M. A., Escadafal, R., Garatuza-Pajan, J., Hanich, L., Perdigão, A., Tapia, J. B., Vuolo, F. (2010). Earth Observation products for operational irrigation management in the context of the PLEIADeS project. Agricultural Water Management, 98(2), 271–282. http://doi.org/10.1016/j.agwat.2010.08.020.

20 Gower, S. T., Norman, J. M. (1991). Rapid estimation of leaf area index in conifer and broad-leaf plantations." Ecology 72.5 (1991): 1896-1900

21 Li-Cor LAI-2000 Plant Canopy Analyser: Instruction Manual. Nebraska Li-Cor, Inc., Lincoln (1992) (179 pp.)

- **ET_a vs. ET_p:** Crosscheck comparison between the actual Evapotranspiration (ET_a) derived from MODIS data (MOD16A2) and ET_p PM-FAO56 derived from Sentinel-2 data. Following the *Analytical Approach,* proposed by D'Urso & Menenti (1995)[22], the daily ET_p values were estimated using the standard FAO vegetation parameters (Allen et al.,1998)[23], while for the actual ET_a MODIS were used as input the same satellite data (LAI and albedo MODIS products) used by the MOD16 ET algorithm. More details are reported in the annexes section (8.1). Subsequently, the daily ET_p products were aggregate to match the temporal composite resolution of the MODIS ET_a (8-days). Moreover, the comparison was assessed by masking for each image the MODIS ET "fill values" pixels corresponding to non-vegetated areas (e.g. urban/build-up, barren, sparse vegetation) for which the MOD16 algorithms do not calculate the ET. The comparison was assessed by temporal and spatial pattern analysis between the ET_a derived from MOD16 ET and the Analytical ET_p, derived from Sentinel-2. Indeed, considering the difference of the spatial resolution of the two sensors (500 m MODIS, 10 m S2), the validation was performed by considering only the fully irrigated pixels – detected by the Map of Irrigated Areas – which had dimensions coherent with the MODIS resolution. For the irrigation season 2018, following this *criteria*, a sample of 12 pixels was selected from the MODIS Grid (Figure 4).

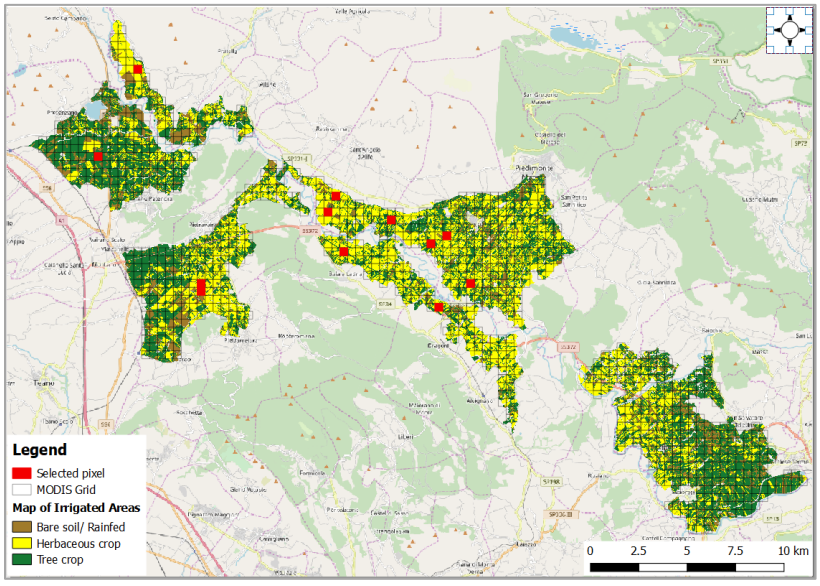


Figure 4 - MODIS pixel selected and their location in the Sannio Alifano pilot area.

[22] D'Urso, G., Menenti, M. (1995). Mapping crop coefficients in irrigated areas from Landsat TM images. In Remote Sensing for Agriculture, Forestry, and Natural Resources (Vol. 2585, pp. 41-48). International Society for Optics and Photonics.

[23] Allen, R., Pereira, L., Raes, D. Smith, M. (1998). Crop evapotranspiration - Guidelines for computing crop water requirements. FAO Irrigation and drainage paper 56. Fao, Rome, 1998, 300.9: D05109.

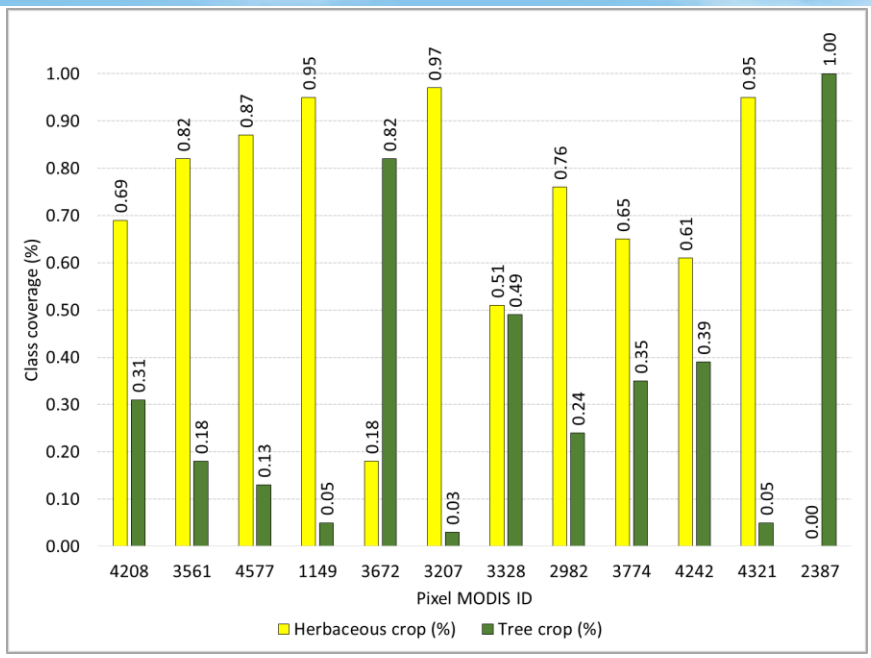


Figure 5 – The class percentage of irrigated surface included in each selected pixel MODIS.

The statistical spatial assessment was performed at the pixel level by the R-square, and the Root Mean Square Error (RMSE ET_a MOD16A2- ET_p PM-FAO56) calculated as:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(ETa_{MOD16A2} - ETp_{PM-FAO56})}{N}} \quad (4)$$

where, N is the number of valid ET MODIS value for each pixel during the considered period (irrigation season 2018).

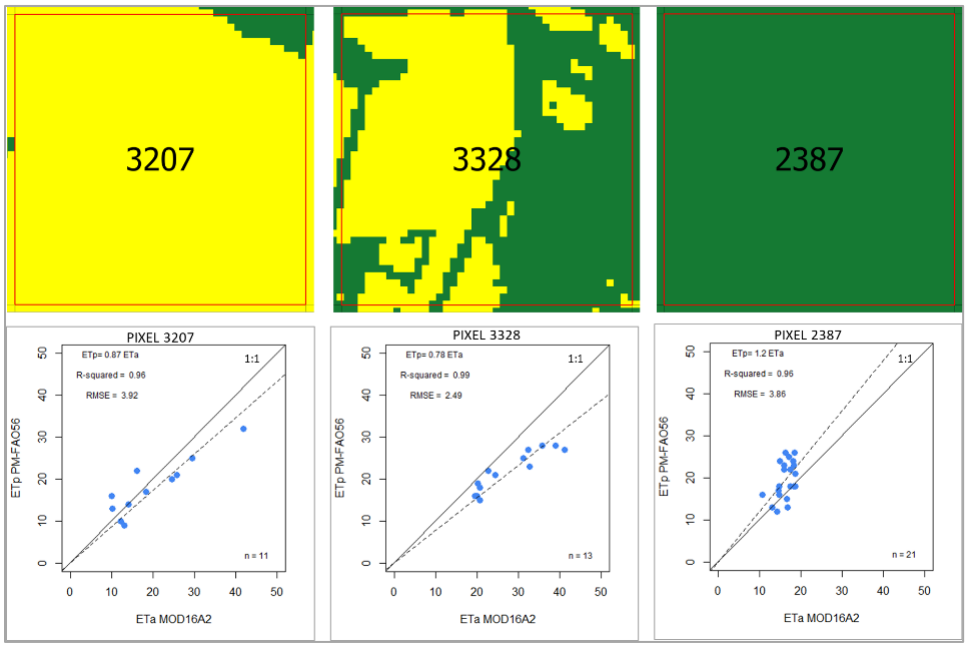


Figure 6 - An example of selected pixel details and the corresponding scatterplot of ET_a vs ET_p.

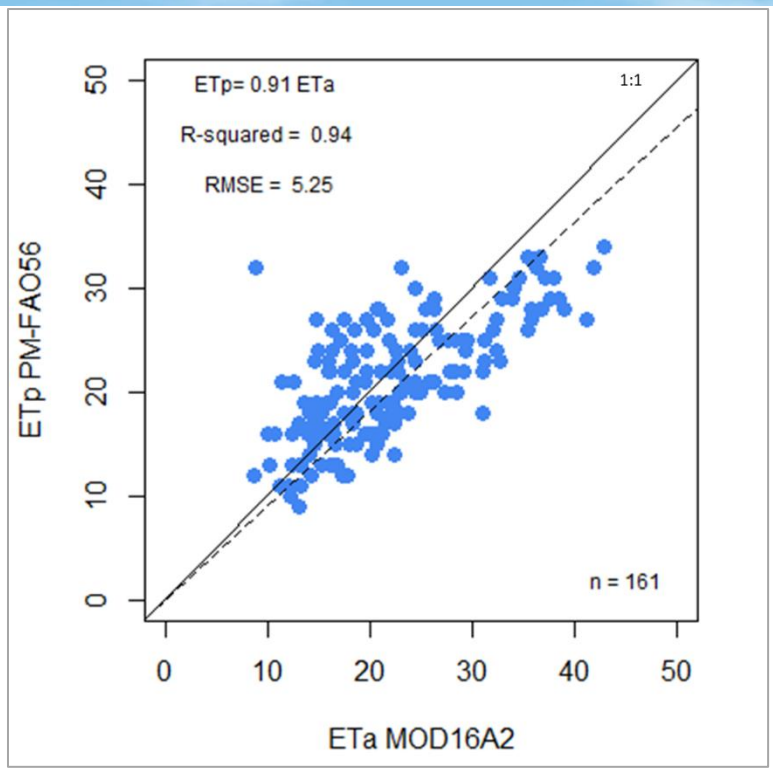


Figure 7 - Comparison between ET_a MOD16A2 (mm/8-days) and ET_p PM-FAO56 (mm/8-days) values for the irrigation season 2018 (from April to September).

In conclusion, in the results analysis context, it is necessary to take into account that footprints of the two sensor types might still vary considerably due to geolocation errors, view-angle effects and different point-spread functions (PSF). Different band locations and widths have also to be considered. Other differences can be explained from the different algorithms used and from the different source of the meteorological dataset. While the Analytical Approach uses daily meteorological data (ERA-Interim produced by ECMWF), the MOD16 algorithm use daily meteorological data produced by NASA's GMAO at the spatial resolution of 0.5° x 0.6°using a global circulation model and both ground and satellite-based observations.

2.3 Net Irrigation Water Requirements

To validate the Net Irrigation Water Requirements (NIWR) data a crosscheck comparison was performed between the Total Precipitation (T_p) estimated by the reanalysis model and the data recorded by the agrometeorological ground stations located in the Sannio Alifano area. In detail, for the irrigation season 2018, regarding the reanalysis model the ERA-Interim dataset was considered (ECMWF), while for the local rainfall data was considered the T_p recorded by the ground station of the CFM network (Centro Funzionale Multirischi) managed by the Civil Protection Department[24]. At this stage, six CFM ground station were selected, considering their intersection with the ERA-Interim Grid (cell size ≈ 12 Km) (Figure 8). Subsequently, the analysis has been conducted cumulating the daily T_p values recorded in the entire irrigation season 2018 (from April to September). An overview of this comparison is plotted in Figure 9.

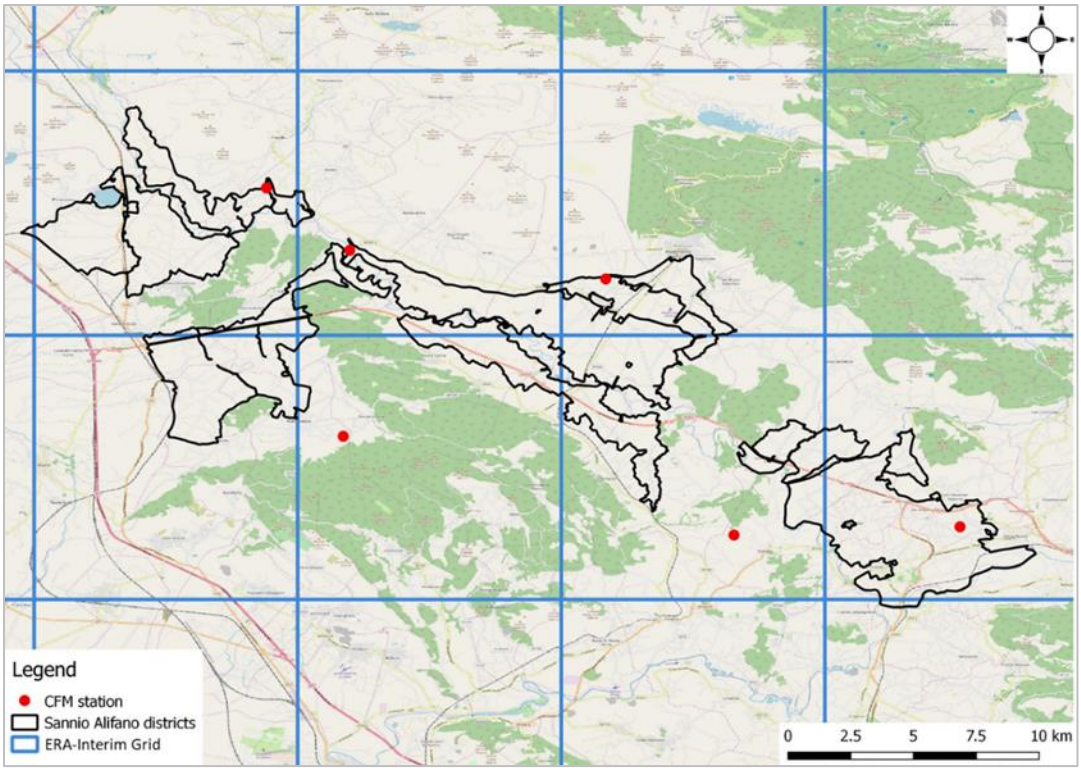


Figure 8 - CFM agrometeorological stations and their intersection with the ERA-Interim Grid.

[24] http://centrofunzionale.regione.campania.it/#/pages/dashboard

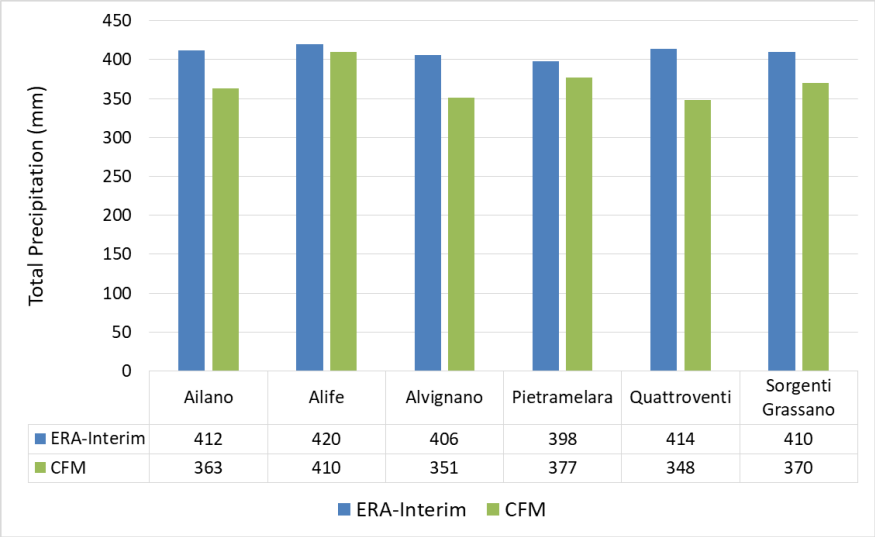


Figure 9 - Comparison of ERA-Interim and CFM total precipitation cumulated for the Irrigation season 2018 (mm/6-months).

Following this approach, it was possible to adjust – where needed – the ERA-Interim data product, then to consider a more accurate Net Precipitation (P_n) involved into NIWR computation process. In detail, with the aim to obtain the effective precipitation, the ERA-Interim T_p values have been reduced take into account the canopy development described using the LAI and fractional vegetation cover (fc) computed from the Sentinel-2 data (Braden, 1985)[25].

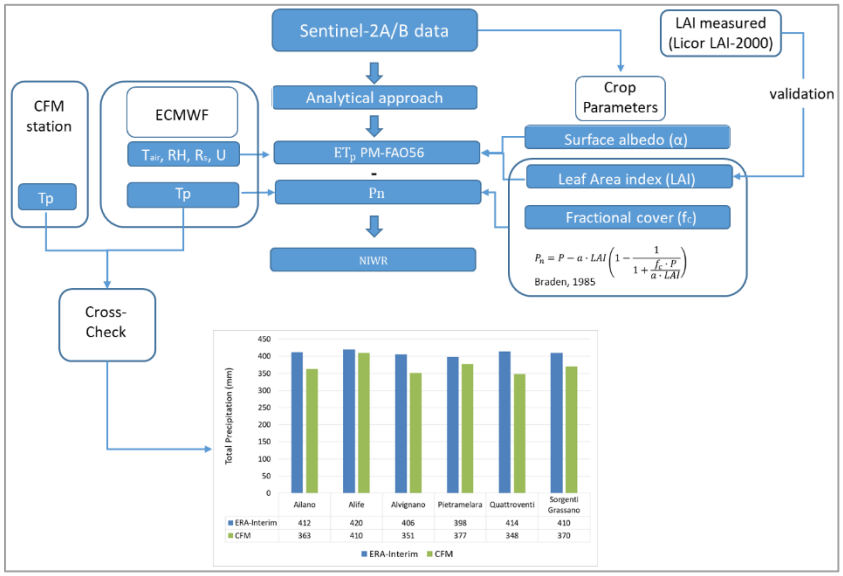


Figure 10 - The concept of the NIWR validation process.

[25] Braden, H. (1985). Ein energiehaushalts-und verdunstungsmodell for wasser und stoffhaushaltsuntersuchungen landwirtschaftlich genutzer einzugsgebiete. Mittelungen Deutsche Bodenkundliche Geselschaft. Retrieved from https://scholar.google.it/scholar?q=48.+Braden%2C+H.+Ein+Energiehaushalts-+und+Verdunstungsmodell+for+Wasser+und+Stoffhaushaltsuntersuchungen+landwirtschaftlich+genutzer+Einzugsgebiete.+Mittelungen+Dtsch.+Bodenkundliche+Geselschaft+1985%2C+42%2C+294–299.&btnG=&hl=it&as_sdt=0%2C5#0

2.4 Gross Irrigation requirements

To validate the Gross Irrigation Water Requirements (GIWR) data products, a comparison between the water volume applied and those deduced from EO data was conducted. In detail, the efficiency coefficients are applied to estimate the GIWR from NIWR, then with the aim to take into account in a lumped way of all water losses occurring in the distribution network of irrigation canals and pipelines.

To quantify the goodness of this approach, a comparison of the water volume applied and GIWR estimated from EO data was performed considering a sample of farms. The information was provided, by the Sannio Alifano Consorzio Staff, regarding cadastral coordinates - Municipality, Sheet and Parcel –irrigation scheduling and identity of landowners. The analysed sample is based on four farms with tree crop (Table 3), equipped by a drip irrigation system, then the irrigation efficiency considered is 0.9. The results are shown in the following graph (Figure 11)

Farm	Tree crop	Irrigation Period (2018)	Irrigation system	Efficiency
1	Peach	May to August	drip irrigation	0.9
2	Hazelnut	July-August		
3	Apple	June to September		
4	Peach	June to September		

Table 3 - Specifications of farms selected.

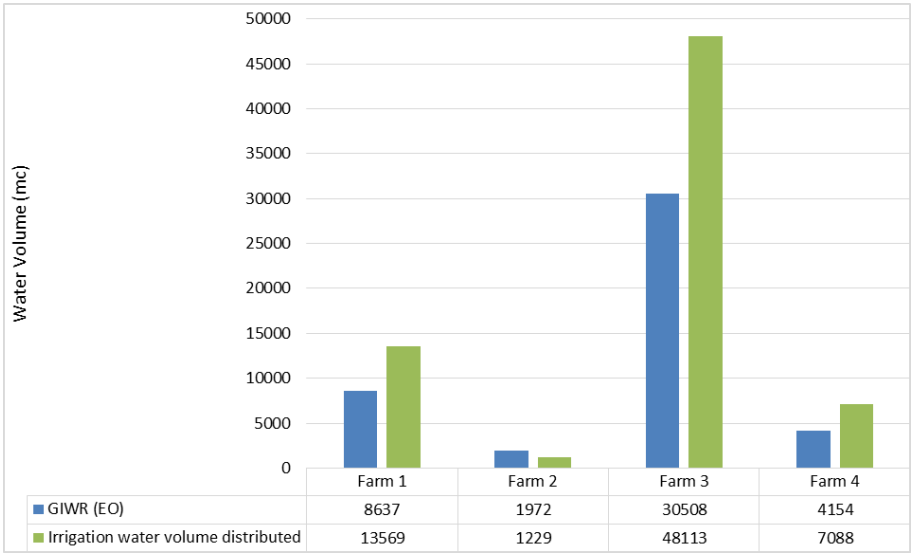


Figure 11 - Crosscheck comparison performed to validate the GIWR data product.

3 Spanish Pilot area: Data products validation for the year 2018

As mentioned, the validation approach of the data products elaborated in DIANA requires two types of processes. On the one hand, those concerning the technical quality of the products, and on the other hand, the usability of these products, what is closely related to the DIANA platform that is being built. Both validation aspects require users involvement.

Aware of the difficulty of the validation procedure, for the Spanish pilot areas, whose location is shown in Figure 12, the validation process began with the own selection of the pilot areas. Besides those already previously included, like “Mancha Oriental” and “Bembezar Margen Derecha”, other demonstration areas like “Bajo Jalón” and “Tierra del Vino” were selected by the Spanish Ministry, together with the Spanish DIANA team. By this way, is easier to involve these users to whom DIANA products are destined into the validation process from the earlier steps.

Figure 12 - Location of Spanish pilot areas, whose limits are superimposed on the Iberian Peninsula map.

3.1 Identified water managers' needs for the detection and monitoring of water abstractions

During the first year of the project, stakeholders expressed their requirements for a better knowledge of water abstractions and specifically on irrigation uses of irrigated areas and abstraction volumes. In the Spanish pilot area involved stakeholders are:

- Water users association:

- Junta Central Mancha Oriental, JCRMO (Central Board of Mancha Oriental) and
- Federación de Regantes de Andalucía, FERAGUA

-Water authorities:

Ministerio para la Transición Ecológica, Dirección General del Agua

River Basin Authority (Hydrological Planning Office): Júcar river, Ebro river, Duero river, Guadalquivir river

The operational requirements emerged included different aspects:

- continuous monitoring of irrigated areas and volumes abstracted from well-established areas;
- ensuring the reliability of self-declarations on water abstractions;
- optimising field inspections to ensure compliance with legal requirements on water abstraction;
- ensuring compliance in regards to seasonal water restrictions in the event of drought management.

Complementary requirements included:

- regularising historical water rights using better identification of water users;
- ex-post assessment on the implementation and efficiency of water management systems, providing different time scales (crisis or structural) as a base for to adopting future actions;
- adjustment of water prices and implementation of a volume-based fee system;
- support for irrigation scheduling to increase efficiency.

DIANA service "Non-authorized water abstraction detection and monitoring for control optimisation" designed and implemented to fulfil these requirements through three main products:

- Maps of irrigated areas

- Maps of Net Irrigation Water Requirements (Gross Irrigation Water Requirements)
- Time series of EO images,

Maps of Irrigated Areas

The detection of irrigated areas (providing a clear identification of the exact location and areal extent) requires land-use/land-cover maps that allow distinguishing irrigated from non-irrigated crops. The successful and widely used methodology implies the use of a supervised "multi-temporal classification" based on a time series of EO images. These images provide the temporal evolution of crops and vegetation during growing season through spectral reflectances, derived Vegetation Indices, (like NDVI, the Normalized Difference Vegetation Index, EVI, SAVI and others), and even, when available, surface temperatures, and radar-based soil moisture by using Sentinel1 imagery. The classification process based on temporal pattern recognition is based on the resulting differences of the canopy on the parameters mentioned above that allow relating each pixel or a group of pixels to a vegetation class. The classes need to be defined from fieldwork and previous knowledge of crop phenology in a given area. The crop classification is the prerequisite for the identification of irrigated areas and the exact period of irrigation.

The identification of plots that receive supplemental irrigation (i.e. less amount of water, but in a well-selected timeframe) usually implies more difficulties: under this practice crops show lower contrast compared to the same non-irrigated crop. Supplemental irrigation is usually used when water stress occurs, and it is employed both in extensive herbaceous annual crops and woody crops. In this case, precipitations data is needed to distinguish irrigation (the vegetation index, reflecting the plant water status, does not differentiate between water coming from rainfall or irrigation).

A benchmarking exercise is ongoing in DIANA to tackle those Mediterranean crops like vine, olive, and almond, among others, which exhibit sparse ground cover and where the deficit controlled irrigation is a usual practice. This exercise tries to exploit in an integral approach new reflectance bands of Sentinel-2, like those placed on the SWIR wavelength, radar-based soil moisture from Sentinel1 imagery, and temperature from Landsat 8, besides conventional Vegetation Index time series.

Maps of Net Irrigation Water Requirements, NIWR maps

A soil water balance following the FAO56 approach, integrating into it a reflectance-based crop coefficient, what we call EO-based soil water balance, has been applied to the pilot areas for the

2016 to 2018 years. The year 2018 products are currently ongoing. By this way, calculation of NIWR in a pixel by pixel way and the daily step is performed.

Either the Irrigated Area maps, either the NIWR maps need to be adapted to the users' requirements. So, the raw product, usually expressed into a pixel by pixel basis is aggregated into the spatial field scale, which is the object of the water administrative rights. For NIWR products, two-time scales are utilised for temporal aggregation: monthly and annual. By this way, it is feasible to compare with authorised irrigated areas and authorised amounts of water for irrigation.

Time series of EO images

In the framework of authorised amounts of water for irrigation, the "Annual Exploitation Plan, AEP" is a key tool. The AEP updates for each irrigation campaign and each field the authorised amount of water, by using parameters like piezometric evolution for groundwater abstractions, amount of water on the reservoirs, and so on. According to the stakeholder requirements, the AEP compliance enforcement requires continuous monitoring during the whole irrigation campaign.

Time series of images on real-time provides valuable help to monitor continuously the AEP. It requires some infrastructure for delivering the images in a proper way for helping the field inspection, what is accomplished by the DIANA platform. Table 4 shows the products delivered to the users for the Spanish pilot areas submitted for validation by the users, according to their requirements.

PILOT AREAS, SPAIN	PRODUCTS DELIVERED TO WATER MANAGERS and THE RIVER BASIN AUTHORITIES 2016-2018				
	EO BASIC PRODUCTS Time series			Maps of Irrigated areas 2016 -2018	EO-based Soil Water Balance
	RGB	NDVI	Kcb		Net (Gross) Irrigation Water Requirements Monthly/annual)
Tierra del Vino	√	√	√	√	√
Bajo Jalón	√	√	√	√	√
Bembezar MD	√	√	√	√	√
Mancha Oriental	√	√	√	√	√

Table 4 - Products delivered to stakeholder, water managers and river basin authorities, submitted for validation.

3.2 Validation of technical quality: accuracy of the maps of Irrigated Areas and NIWR map

Field inspections in the case of the area of the Mancha Oriental have allowed to evaluate the reliability of this methodology of crop classification, reaching 90-95% in the case of the determination of irrigated areas of arable crops, and 70% in the case of the determination of irrigated areas of woody crops.

Numerous works support the EO-based soil water balance for the calculation of the NIWR maps at field scale (for a review, see Calera et al., 2017)[26], and therefore, the calculated NIWR represents the optimal quantity of water to supply the crop. However, individual decisions made by farmers about the water to apply may not match this amount. But, the so calculated optimal amount of water to supply to the crop currently is being used for advising about water to apply in sustainable water use.

Beyond of technical aspects, the delivered NIWR maps are being checked directly by the users. The Water Directorate of Spanish Environmental Ministry has begun a program to evaluate the results obtained in DIANA, by comparing the NIWR maps against data directly obtained by the river basin authorities from water meter data, and other sources, in different areas and different environments. This program aims to evaluate the application of the same and homogeneous EO-based methodology and its potential use for the whole Spanish territory.

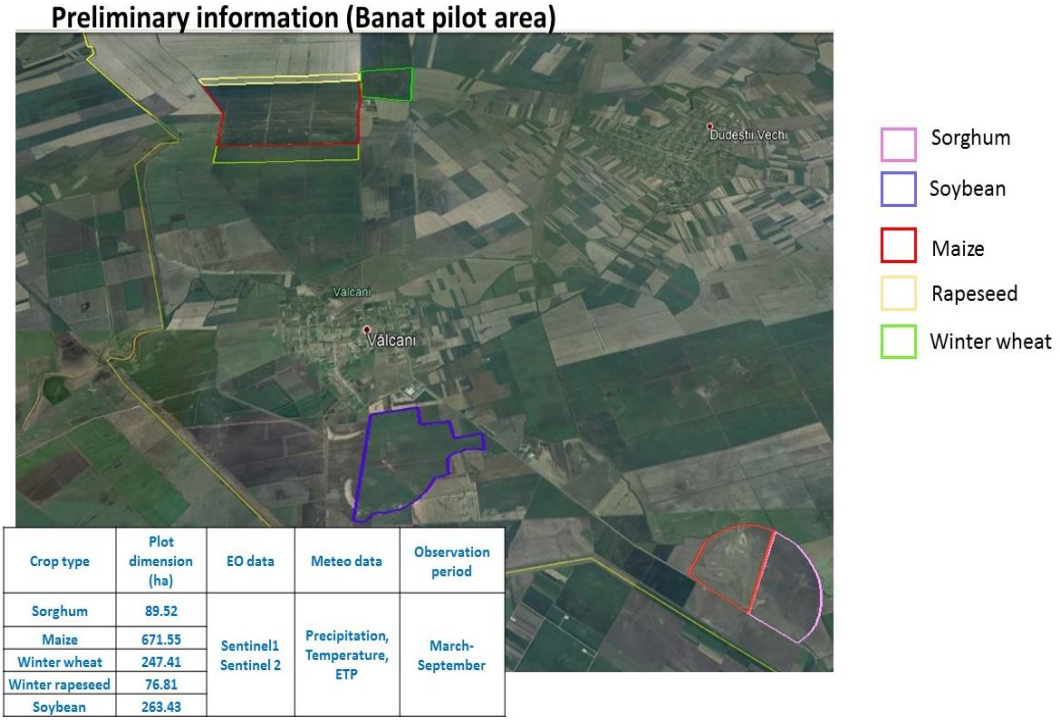
[26] Calera, A., Garrido-Rubio, J., Belmonte, M., Arellano, I., Fraile, L., Campos, I., Osann, A. (2017). Remote sensing-based water accounting to support governance for groundwater management for irrigation in La Mancha Oriental Acquifer, Spain. WIT Transactions on Ecology and the Environment, 220, 119-126.

4 Romanian Pilot area: Data products validation for year 2018

4.1 Map of irrigated areas

Banat region is located in the South-West of Romania bordered by the Mures river in the north, the confluence between Danube and Cerna rivers in the south, Romanian – Serbian border in the west and the Jiu river in the east. From the morphological point of view, this plain is a flat relief unit, with uniform appearance but heterogeneous in what concerns the lithology and soil, while flat surfaces are frequently separated by abandoned meanders.

Test area (Figure 13) covering SC Emiliana West farm has a special particularity consisting in heavy, dark, clayey soil type that becomes very solid when wetted. Also, depending on the rainfall regime during the season, the groundwater level could increase to near the land surface (0.4 – 4m). These aspects influence the NDVI classification results related to irrigated and non-irrigated parcels (Figure 14): NDVI has high values outside the irrigated plots while irrigated plots show medium values. Therefore, we focused on multi-temporal Sentinel-1 data analysis by applying supervised classification on irrigated and non-irrigated plots using as input sigma nought time series. Sentinel-1 data were downloaded from the COPERNICUS SCIHUB. Sensor configuration mode acquisition is shown in Table 5.



Crop type	Plot dimension (ha)	EO data	Meteo data	Observation period
Sorghum	89.52	Sentinel1 Sentinel 2	Precipitation, Temperature, ETP	March-September
Maize	671.55			
Winter wheat	247.41			
Winter rapeseed	76.81			
Soybean	263.43			

Figure 13 - Banat pilot area: main crops in 2018 and data needs.

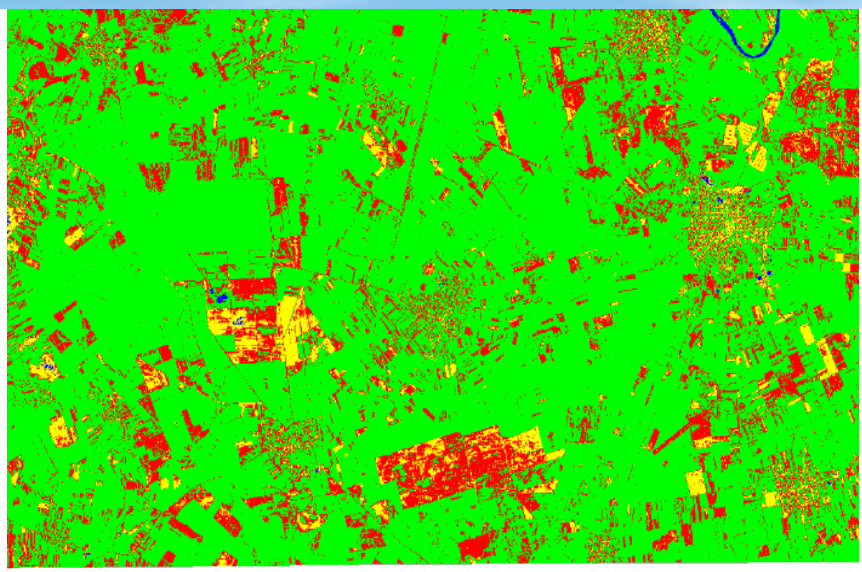


Figure 14 - NDVI classification: Yellow: bare soils, blue: water, green: non-irrigated areas, red: irrigated areas.

Sentinel-1 data	Characteristics
Orbit	Descending node
Polarisation	VV + VH
Mode	Interferometric Wide – GRD (ground range)
Incidence angle	Ranging from 30^0 - 40^0
Relativ orbit	153

Table 5 - Sentinel-1 data acquisition configuration.

A total of 107 Sentinel-1 data covering September 2016 – September 2018 were prepared for supervised classification by applying special processing steps like: orbit corrections, radiometric calibration, speckle filtering, terrain correction and normalisation. Machine learning algorithms results are shown in Table 6. The best results are obtained when VV + VH polarisation is used.

Machine Learning Algorithm	Crop type	Accuracy
Random forest	Sunflower, soybeans, maize, sorghum	0.95
K-Nearest Neighbor Classification[27] (KNN)	Sunflower, soybeans, maize, sorghum	0.84
Minimum Distance Classifier (MD)	Sunflower, soybeans, maize, sorghum	0.93

Table 6 - Machine learning classification results.

4.2 Crop evapotranspiration

EO-derived crop evapotranspiration (maximum) (ETC) is the value of evapotranspiration under standard conditions, as defined by FAO56: disease-free, well-fertilized crops, under optimum soil

[27] Campos, Guilherme O.; Zimek, Arthur; Sander, Jörg; Campello, Ricardo J. G. B.; Micenková, Barbora; Schubert, Erich; Assent, Ira; Houle, Michael E. (2016). On the evaluation of unsupervised outlier detection: measures, datasets, and an empirical study. Data Mining and Knowledge Discovery. doi:10.1007/s10618-015-0444-8. ISSN 1384-5810

water conditions. It can be derived by using a coefficient-based approach (KC* x ET0; KC*: reflectance-based crop coefficient; several approaches can be used for estimating KC* from Vegetation Indices, by using basal crop coefficient relationships).

To calculate ETC, the Crop coefficient-Reference evapotranspiration methodology described in the FAO56 manual (Allen et al.1998) was used. More concretely a "single" crop coefficient model where the estimation of ETC, in the absence of water stress, is calculated as:

$$ETC = KC \cdot ET0 \quad (5)$$

where, ET0 is the reference evapotranspiration and KC is the crop coefficient. ET0 is calculated from different climatic variables, while KC can be calculated through its linear relationship with the NDVI. In this case, the equation used has been the next:

$$KC = 1.25 \cdot NDVI + 0.1 \quad (6)$$

Thus, from NDVI data, the ETC can be calculated via KC relationship.

The NIWR and GIWR products are related as follows:

$$NIWR = ETC - PP \quad (7)$$

$$GIWR = NIRW/Ɛ \quad (8)$$

where, PP is the precipitation and Ɛ is the irrigation efficiency. So, from the ETC we can obtain the other products.

Accumulated ETC calculation in Banat area

The software used to calculate the accumulated ETC is Tonipbp, developed in UCLM (Albacete, Spain). The program allows to accumulate values of ETC pixel by pixel from temporal series of NDVI and reference ET (ET0) data during the growing crop as follow:

$$KC = 1.25 \cdot NDVI + 0.1 \qquad ETC = KC \cdot ET0 \quad (9)$$

It is necessary to describe the growing crop cycle to define the period to accumulate the ETC per pixel. For the Banat area, the growing cycle was defined for the irrigated crops in 2017: soybeans, sunflowers, maize and sorghum.

The inputs of TONIpbp are: selected time series NDVI (cloud-free), daily ET0, the range of NDVI values and the shape (layer) of the area of interest. The software creates daily NDVI values interpolating between the dates of NDVI images, applies the correlation NDVI- KC defined above and multiplied by the daily ET0 values. In order to capture the growing cycle, it is necessary to define the values of NDVI representing the beginning and the end of the crop cycle. In this case, it has been established 0.3 NDVI as the green up, the moment when the crop starts the vegetative development and 0.45 NDVI as the maturity physiological is reached. So, when one-pixel reaches

0.3 of NDVI, TONI will use this date to start to accumulate the ETC, and it will stop in the date when the value of NDVI decrease under 0.45. The daily ET0 and precipitation were obtained from the weather station of Sannicolau Mare.

The images used were NDVI Sentinel-2A, Sentinel-2B and Landsat 8 for 2017 cloud free. The dates used are summarized in the next list.

- ndvi3_20170102_S2A_32634_136000T34TDS_02.img
- ndvi3_20170303_S2A_32634_136000T34TDS_00.img
- ndvi3_20170402_S2A_32634_136000T34TDS_00.img
- ndvi3_20170516_landsat8_32634_186028_cc_26_rec.img
- ndvi3_20170704_S2A_32634_036000T34TDS_00.img
- ndvi3_20170719_S2B_32634_036000T34TDS_01.img
- ndvi3_20170808_S2B_32634_036000T34TDS_02.img
- ndvi3_20170815_S2B_32634_136000T34TDS_00.img
- ndvi3_20170914_S2B_32634_136000T34TDS_00.img
- ndvi3_20171002_S2A_32634_036000T34TDS_00.img

The Figure 15 is an example of the curve described by the evolution of NDVI images values for one pixel (red point in the map) in a soybean plot in an irrigated plot of Banat area.

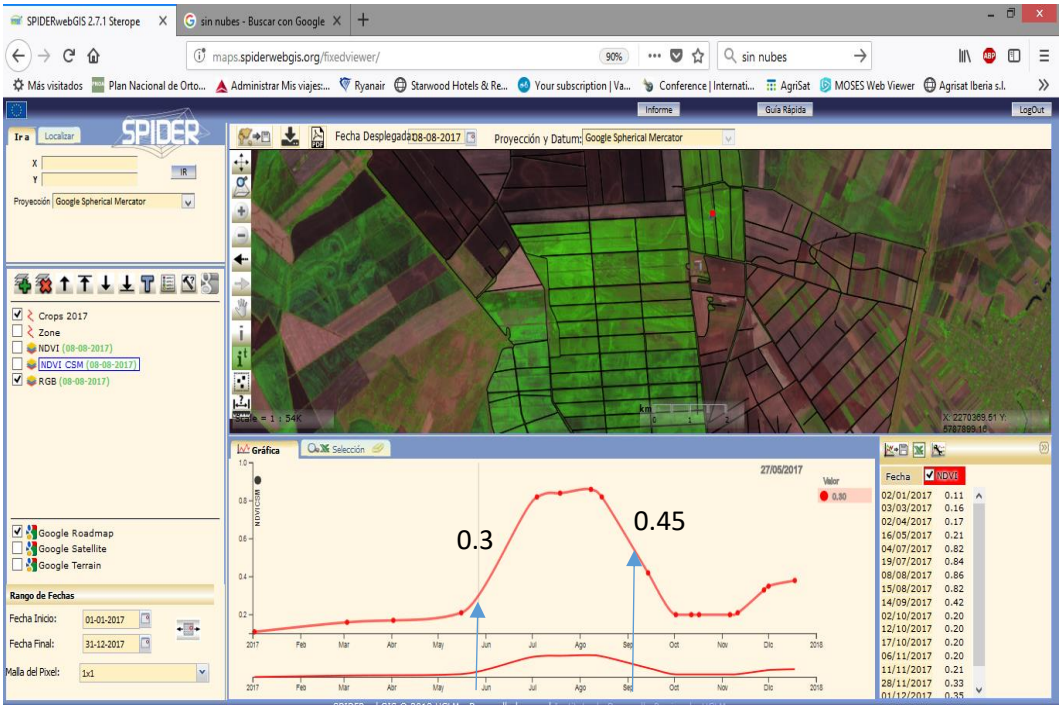


Figure 15 - NDVI evolution for a pixel in a Soybean plot in Banat area in 2017.

For this pixel, the software starts to accumulate in the 26-05-2017 (0.3 NDVI) and stops in 11-09-2017.

The final product is the accumulated ETC (mm) during the growing cycle pixel by pixel. This product is represented in the next figure (Figure 16, provided in .img format). Notice the south plots are incomplete due to the extension of the satellite Image does not cover the whole plot.

Figure 16 - Evapotranspiration accumulated at the end of the crop growing cycle.

Every pixel value of the ETC map shows the crop evapotranspiration in millimetres, mm (l/m2), usually understood as the crop water requirements (CWR). This ETC is the accumulated over the growing season Units are mm for the total CWR growing cycle (mm(year)).

ETc average (mm): Once we have the ETC value pixel by pixel for the whole growing cycle in a raster format, the zonal statistic has been run in QGIS program. Previously, it has been applied a buffer of minus 30 meters to the original vector layer where the boundaries of the plots are defined in order to avoid the boundary effects of the images. Finally, it has been obtained the ETC average per plot.

The value it has been provided for every polygon, in the attribute table of the original vector layer (column called ETc_lbl).

Figure 17 shows the map with assigned values of ETC.

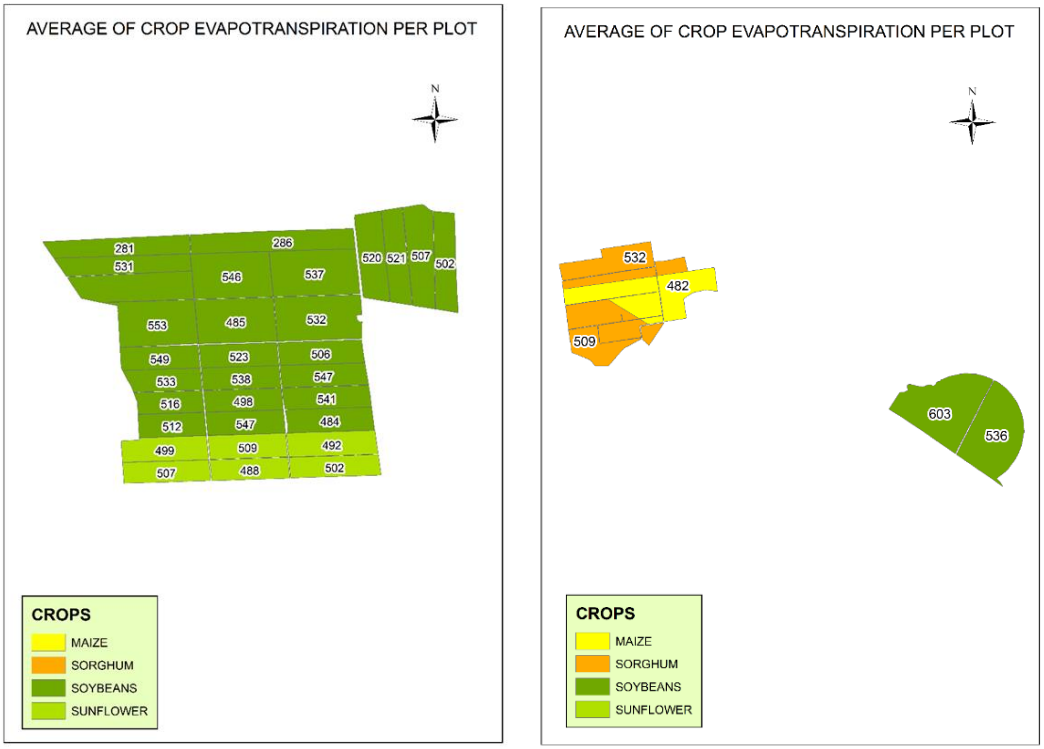


Figure 17 - Average of Accumulated Evapotranspiration per plot.

4.3 Net Irrigation Water Requirements (mm) NIWR = ETc-PP

PP is the precipitation during the growing cycle. In order to accumulate the precipitation during the growing cycle, the dates of the beginning of the cycle (0.3 NDVI) and the end (0.45 NDVI) have been established studying groups of plots and crops which have similar cycles. Finally, seven groups were defined (Table 7). Figure 18 shows the map with assigned values of precipitation.

Group	Precipitation (mm)	Start. Date	Finish. Date
Sorghum	84.4	21/05/2017	16/08/2017
Corn	89.6	16/05/2017	24/08/2017
Soybeans1	86.3	15/03/2017	03/09/2017
Soybeans2	125.3	04/05/2017	03/09/2017
Soybeans3	125.3	26/04/2017	03/09/2017
Soybean4	89.6	16/05/2017	26/08/2017
Sunflowers	82	23/05/2017	19/08/2017

Table 7 - Groups of crops and plots associated by their similarity of the length of their growing cycles.

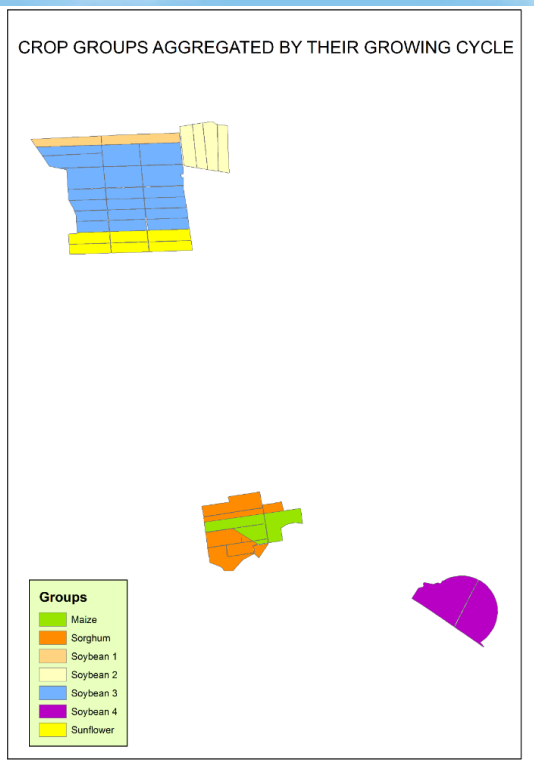


Figure 18 - Groups of crops and plots associated by their similarity of the length of their growing cycles.

The value it has been provided for every polygon, in the attribute table of the original vector layer (column called NIWR_lbl). Figure 19 shows the map with assigned values of NIWR.

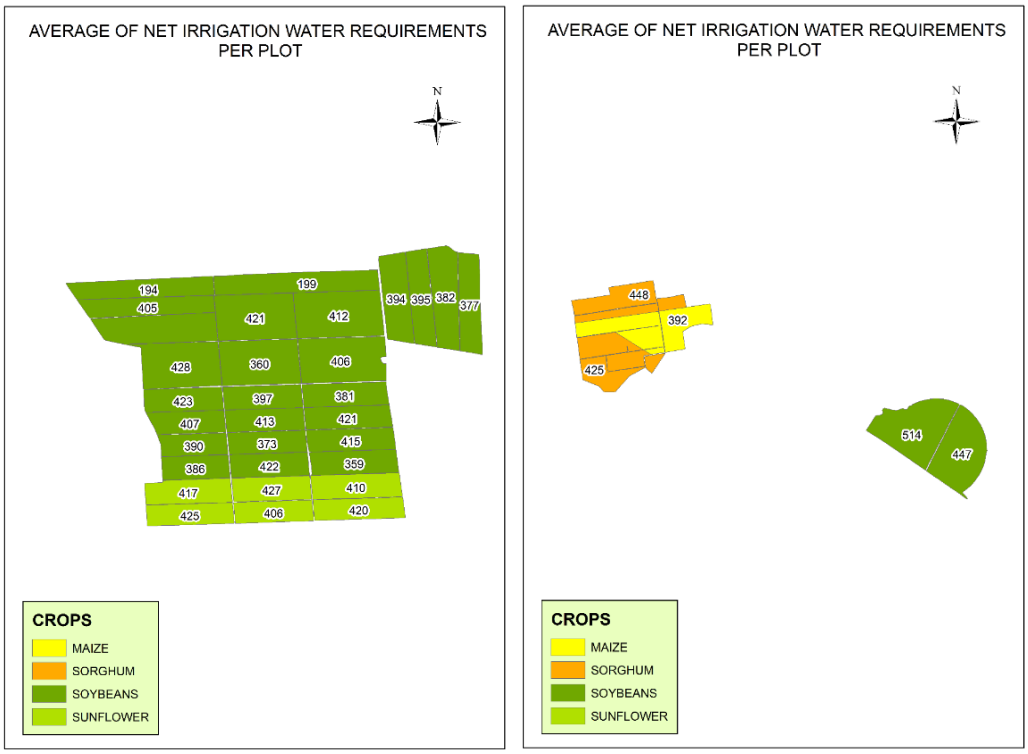


Figure 19 - Net Irrigation Water Requirement Average per plot.

4.4 Gross Irrigation Water Requirements (mm) (NIWR/Ɛ)

Derived from applying an efficiency irrigation coefficient to the net irrigation water requirements NIWR. In this case it has been used 0.85 as a regular value but we need to keep in mind that these coefficients are very specific parameters depending mainly on the type of irrigation system and on the effectiveness of the application of the water (open channels, pumping, sprinkling, dripping, irrigating climate conditions, etc.).

The value it has been provided for every polygon, in the attribute table of the original vector layer (column called GIWR_lbl). Figure 20 shows the map with assigned values of precipitation.

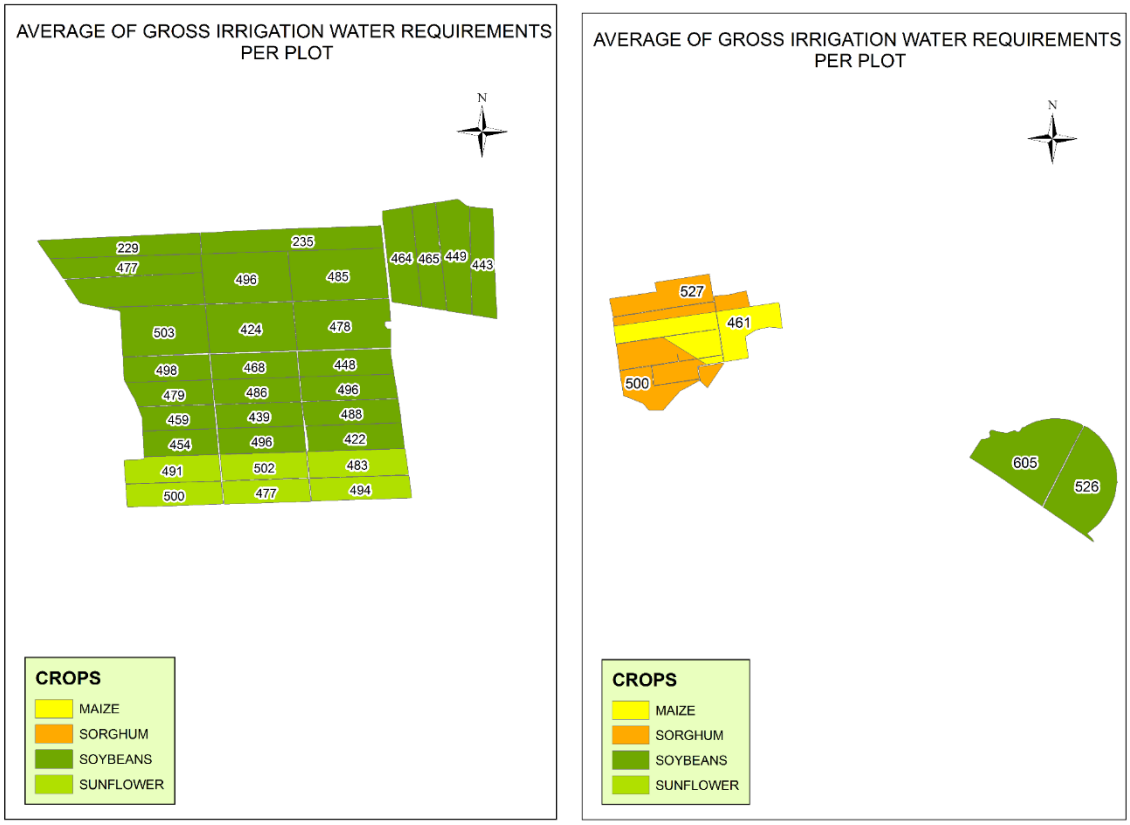


Figure 20 - Gross Irrigation Water Requirement Average per plot.

These results are validated by SC Emiliana West Rom, also providing the information required in the DIANA project like: parcel, crop planning and irrigation requirement vs irrigation estimation.

5 Diana Benchmark exercise

5.1 OPTRAM

Background

Sadeghi et al., (2017)[28] proposed a physical "OPtical TRApezoid Model" (OPTRAM) to remote sensing of soil moisture (SM) EO data.

OPTRAM is based on the pixel distribution within the shortwave infrared transformed reflectance (STR)-Normalized Difference Vegetation Index (NDVI) space (STR-NDVI space), defined as:

$$STR = \frac{(1 - R_{SWIR})^2}{2R_{SWIR}} \tag{10}$$

$$W = \frac{\theta}{\theta_s} = \frac{i_d + s_d \cdot NDVI - STR}{i_d - i_w + (s_d - s_w) \cdot NDVI} \tag{11}$$

$$STR_d = i_d + s_d \cdot NDVI \tag{12}$$

$$STR_w = i_w + s_w \cdot NDVI \tag{13}$$

$$W = \frac{STR - STR_d}{STR_w - STR_d} \tag{14}$$

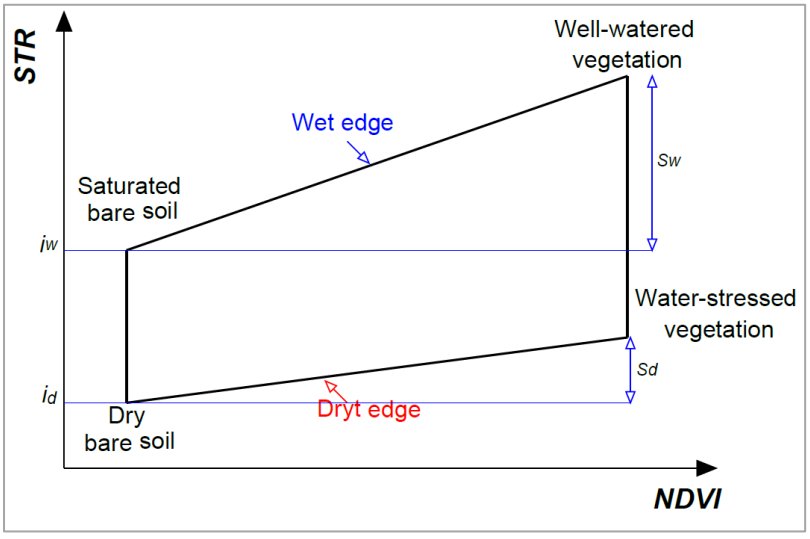


Figure 21 – Sketch illustrating parameters of the optical trapezoid model.

The advantage of the STR-θ space over the conventional triangle/trapezoid approach (LST-θ space) (Carlson et al., 1994)[29] is that it only requires optical data and it can be universally

[28] Sadeghi, M., Babaeian, E., Tuller, M., Jones, S. B. (2017). The optical trapezoid model: A novel approach to remote sensing of soil moisture applied to Sentinel-2 and Landsat-8 observations. Remote sensing of environment, 198, 52-68.

[29] Carlson, T. N., Gillies, R. R., Perry, E. M. (1994). A method to make use of thermal infrared temperature and NDVI measurements to infer surface soil water content and fractional vegetation cover. Remote sensing reviews, 9(1-2), 161-173.

parameterised for a given location because STR-θ is not affected by ambient atmospheric factors. However, the STR-θ space is sensitive to oversaturated pixels (due to ponding water), and then, occasionally, the index W assumes values higher than 1 (Sadeghi et al., 2017)[28].

This section aims to show the testing of the optical trapezoidal model (OPTRAM) for the remote sensing monitoring of soil moisture, at high resolution, for detection of irrigated crop.

Satellite imagery (EO data)

To obtain the Sentinel-2 data for the study area, the Copernicus Open Access Hub (previously known as Sentinels Scientific Data Hub) was considered. A selection of 44 Sentinel-2 was chosen, according to the following *criteria*: cloud-free images for the area of interest; and 2A processing level, corrected for Bottom Of Atmosphere (BOA) reflectance values. Images were selected into the time-period of 3 January – 30 September 2018. To compute the NDVI the band 4 and 8 were selected, while the STR index (eq. 10) computation was performed using the band 12, opportunely resampled at 10 m.

Test sites

The OPTRAM model was evaluated for some parcels located in the Presenzano district (Figure 22). For each considered parcel, with tree crops, the Sannio Alifano Consorzio staff provided the volume and the irrigation time, in addition to other information reported in Table 8.

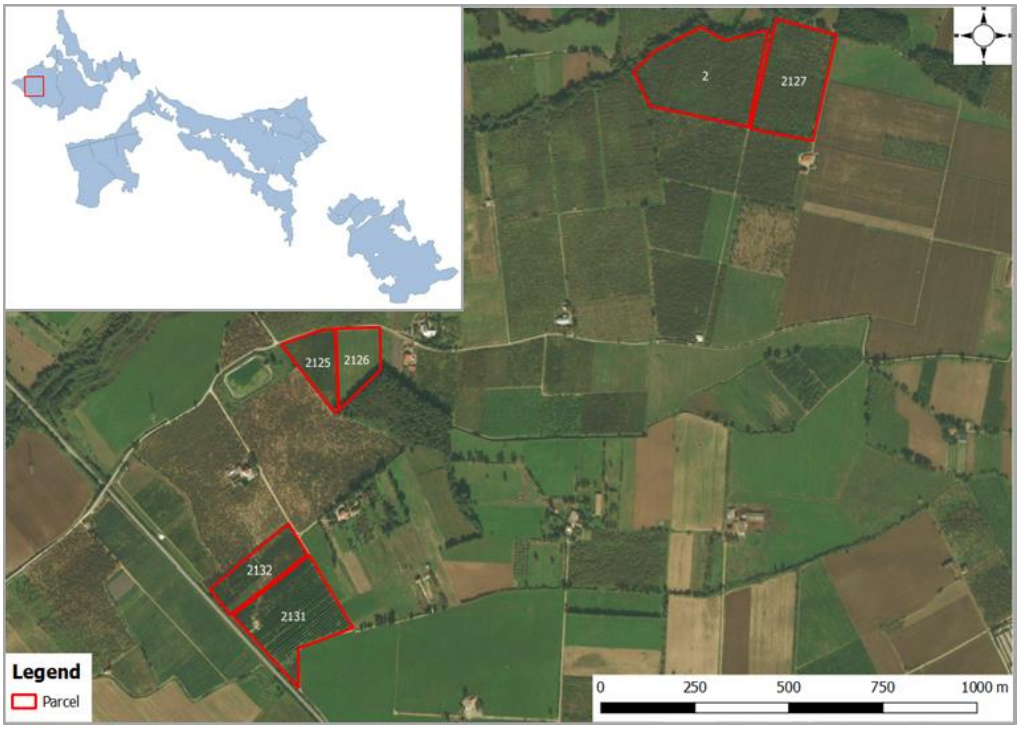


Figure 22 – Selected parcels and their location in the Presenzano district.

ID Plot	Crop type	Plot dimension (ha)	Hydrants ID	Irrigation scheduling	EO Data
2125	Peach	1.8	1082	Volume & Irrigation Time	Sentinel-2
2126		1.9			
2127	Hazelnut	4.9	1145		
2		6.4			
2131	Apple	4.9	552, 548, 547		
2132	Peach	2.4	1219		

Table 8 - Data related to of parcel holding irrigation water rights (Figure 24).

Following are reported more details, about a first test executed on the parcels with hazelnut trees (Figure 23), irrigated from 26 July to 2 August.

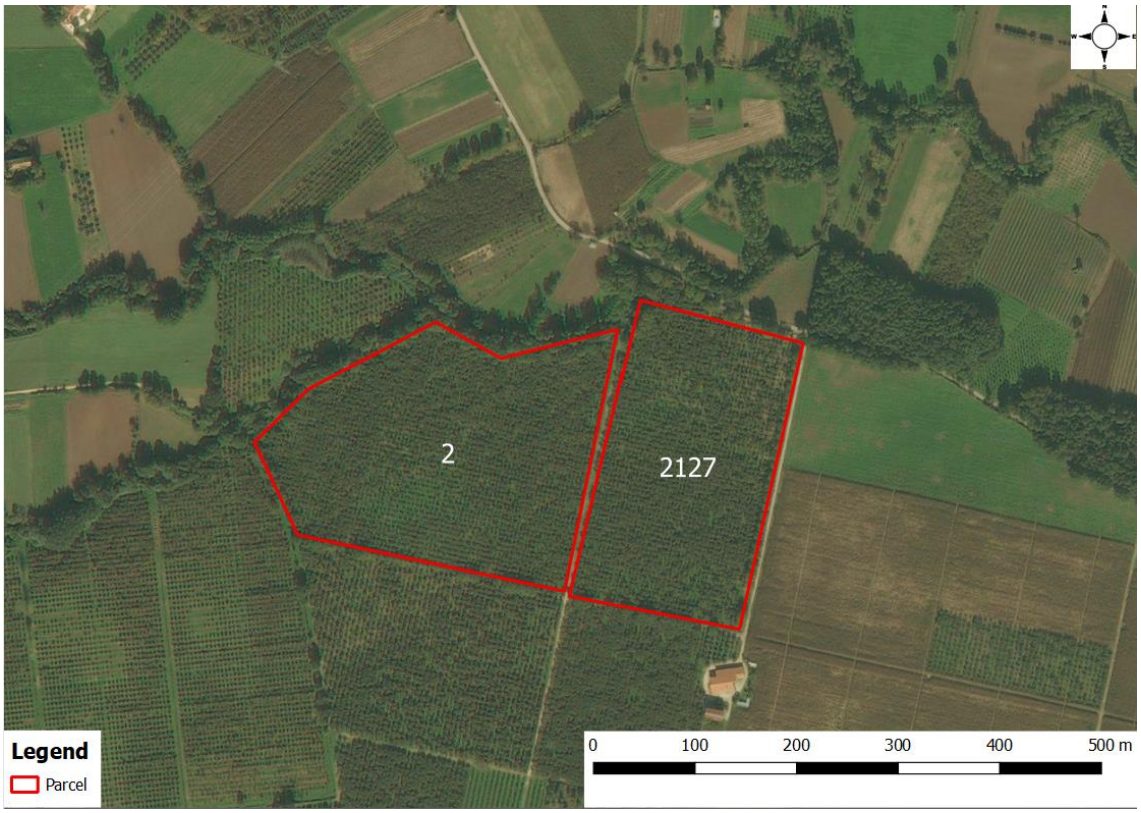


Figure 23 - Map of parcel holding irrigation water rights (halzenut).

Results

In this exercise, for a first analysis, the pixel distributions of the NDVI-STR were plotted, considering the S2 images collecting during the irrigation period (Figure 24.), but to parameterise the OPTRAM model, all the 44 S2 imagery were considered. In detail, the wet edge lines (continuous line) and dry edge (dash-dot line) were identified by a visual inspection and provided

the values of slope and intercepts for the dry edge (s_d and i_d) and for wet edge (s_w and i_w) respectively equal to s_d=2; id=0; i_w=0 and s_w=9.5 (Figure 25).

Following the eq. 14, for each acquisition the soil moisture index (W) was computed. Subsequently, the time series of the W index was compared with rainfall and irrigation scheduling for the considered parcels (Figure 26 and 27).

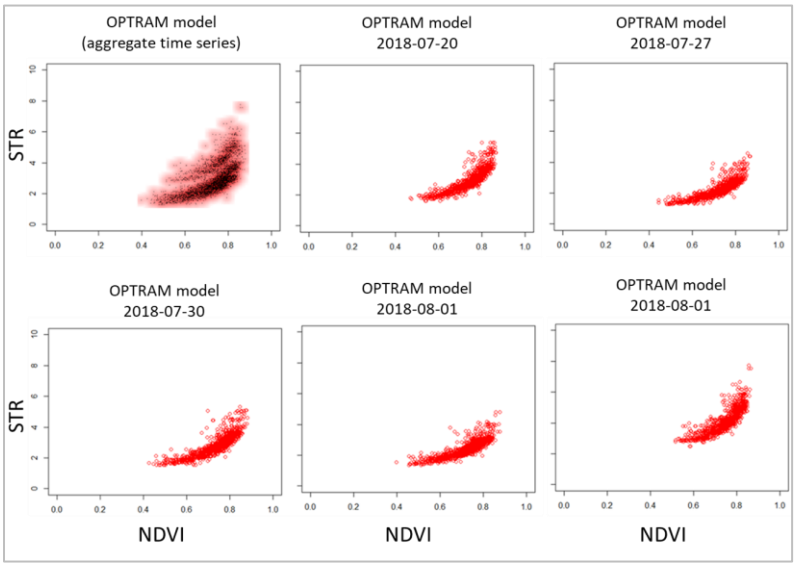


Figure 24 – Scatter plot of STR-NDVI space related to plots 2 and 2127, for aggregate time series (from 20 July to 2 August) and for single Sentinel-2 acquisitions.

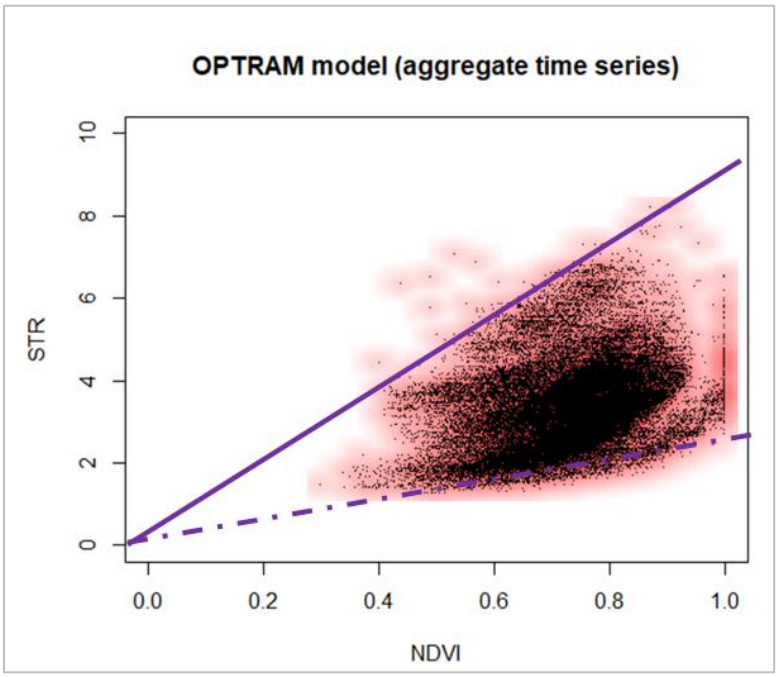


Figure 25 - Scatter plot of STR-NDVI space to plots 2 and 2127, for aggregate time series (from 3 January to 30 September).

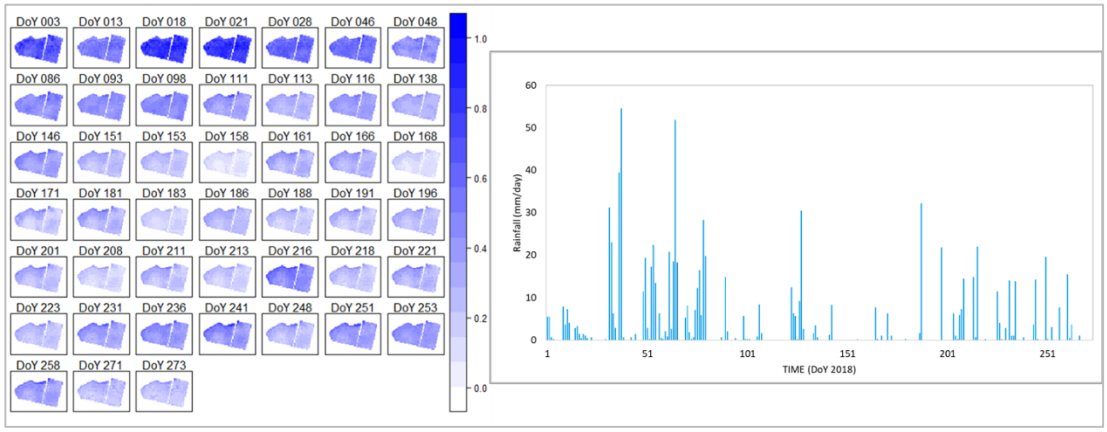


Figure 26 - Time series of soil moisture index (W) (left), and rainfall (CFM station) (right), related to plots 2 and 2127 for a number of 44 acquisitions of Sentinel-2.

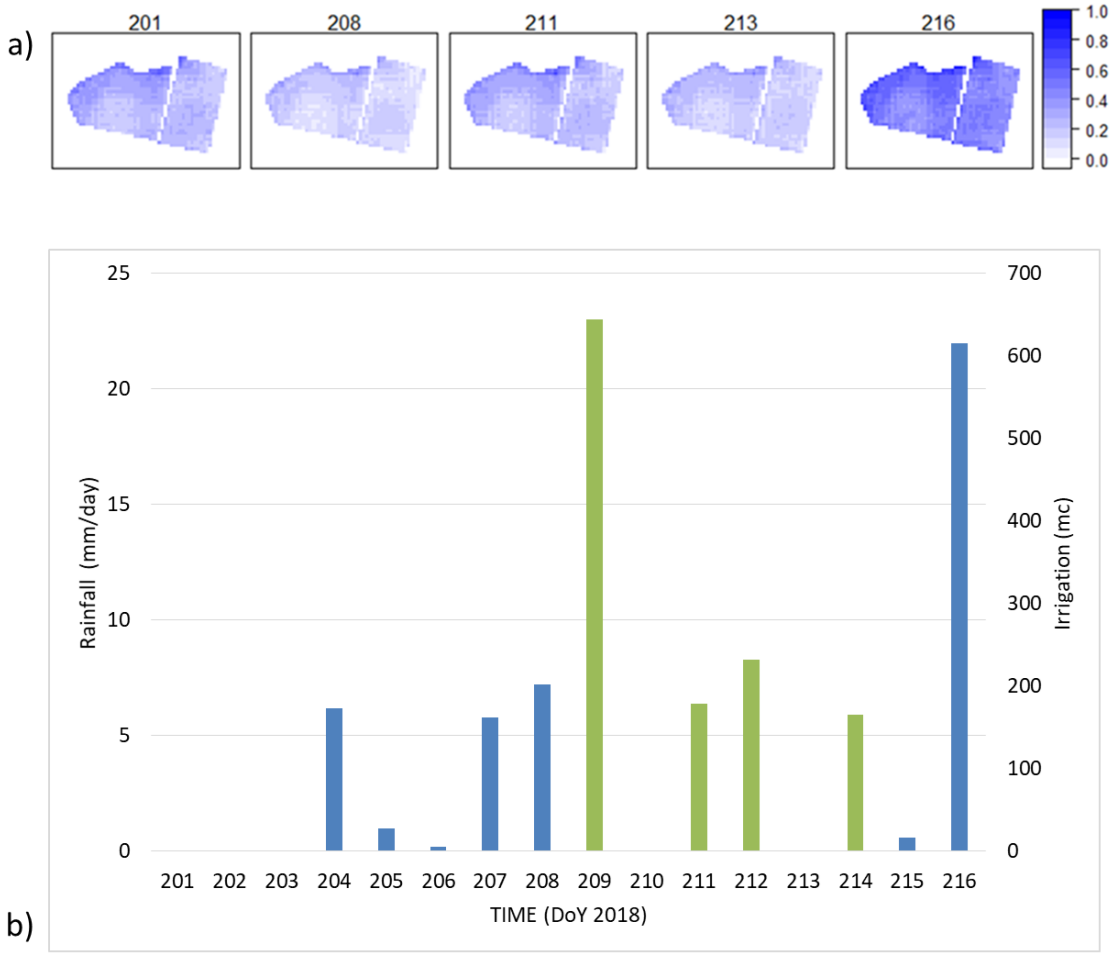


Figure 27 – a) Time series of soil moisture index (W), b) CFM Rainfall (in blue) and irrigation water volume (in green).

5.2 TOTRAM

The Thermal-Optical TRApezioid Model (TOTRAM) is one of the most popular approach used to estimated soil moisture from EO data. Also called "*trapezoid"* or *"triangle"* model is based on the pixel distribution of the Land Surface Temperature (LST) and NDVI space.

The TOTRAM model was applied in the area of Tarazona de la Mancha (Mancha Oriental, Spain) (Figure 12). Particularly, the goodness of this approach was tested taking into account typical Mediterranean permanent crops, usually under drip irrigation and characterised by a large fraction of bare soil. In detail, the test was executed on tree crops like vine, almond and olive.

In this exercise, the pixel distribution of the thermal and optical data was obtained from the LANDSAT 8 OLI-TIRS images acquired from April to August 2018.

The dates used are summarised in the next list:

- April, 19th
- July, 17th
- July, 24th,
- August, 2nd,
- August, 9th,
- August, 24th

5.3 TU WIEN algorithm

Radar signal modelling

Change detection method developed by Wagner for soil moisture estimation at the global scale (1 Km resolution) using SAR scatterometer sensor was used for modelling backscattered radar signal over wheat and rapeseed[30,31]. This method considers temporal stability of the spatial soil moisture pattern supposing soil roughness and vegetation conditions are not changed when a dense time series of SAR data is available. In a particular day, soil moisture is computing by comparing wetting and drying trend for radar backscattered signal using historically lowest and highest backscattered values as references (eq.15 – eq.17):

$$\sigma^{o}(\theta,t) = \sigma^{0}_{dry}(30) + \beta(\theta - 30) + Sm_{s}(t) \quad (15)$$

$$S = \sigma_{wet} - \sigma_{dry} \quad (16)$$

$$SM_t = \frac{\sigma^0(t) - \sigma^0_{dry}(30)}{\sigma^0_{wet}(30) - \sigma^0_{dry}(30)} \quad (17)$$

where,

- S is radar sensitivity,
- β is slope (local incidence angle),
- $m_s(t)$ is the surface soil moisture normalised at an incidence angle of 30°.

As a result, surface soil moisture is expressed in volumetric units (%) and scaled between zero and one. This method is applied only for positive temperature values when the soil is not covered with snow or is frozen.

Soil moisture retrieval

Soil moisture estimation implies calibration of the change detection method whose main steps are described as follows.

Step 1. Normalising the backscattering coefficients. To adjust local incidence variation during acquisition, VV and VH-polarized backscattered coefficients are normalised by using Lambert low from optics[32]:

$$\sigma^0_{\theta_{ref}} = \sigma^0_{\theta} \frac{cos^2\theta_{ref}}{cos^2\theta} \quad (18)$$

where,

$\sigma^0_{\theta_{ref}}$ and σ^0_{θ} are backscattering coefficients observed at incidences angles θ and θ_{ref}. As is

[30] Wagner, W. L. (1999). A study of vegetation cover effects on ERS scatterometer data. IEEE Transactions on Geoscience and Remote Sensing, 37 (2): 938-948.

[31] Wagner, W. P. F. (2008). Temporal stability of soil moisture and radar backscatter observed by the Advanced Synthetic Aperture Radar (ASAR). Sensors, (8): 1174-1197.

[32] Topouzelis, K. S. (2016). Incidence angle normalization of Wide Swath SAR data for oceanographic applications. Open Geosciences, 8(1), 450-464.

mentioned earlier, the incidence angle of Sentinel-1 data is set to 39°.

Step 2. Determine dry and wet references. After normalization and meteorological data analysis, the dry and wet references are determined by statistical methods of noise analysis.

Step 3. Calculate surface soil moisture. Soil moisture is estimated by applying Eq. 17, comparing reference backscattered signal with dry and wet references values.

Step 4. Analyse the vegetation contribution to the soil moisture. The estimated soil moisture content values are used to model vegetation contribution and to validate the estimated results since no measurements have been done in the test area.

Surface soil moisture estimation

Investigations were focused on establishing the dry and wet references based on seasonal variation which reflect vegetation phenology. This variation can induce backscatter changes that required vegetation correction. Sentinel-1 images acquired between September 2016 and February 2017 were analysed separately to see the radar temporal response during the seeding and dormancy crop stages in correlation with precipitation and air temperature values (Figure 28).

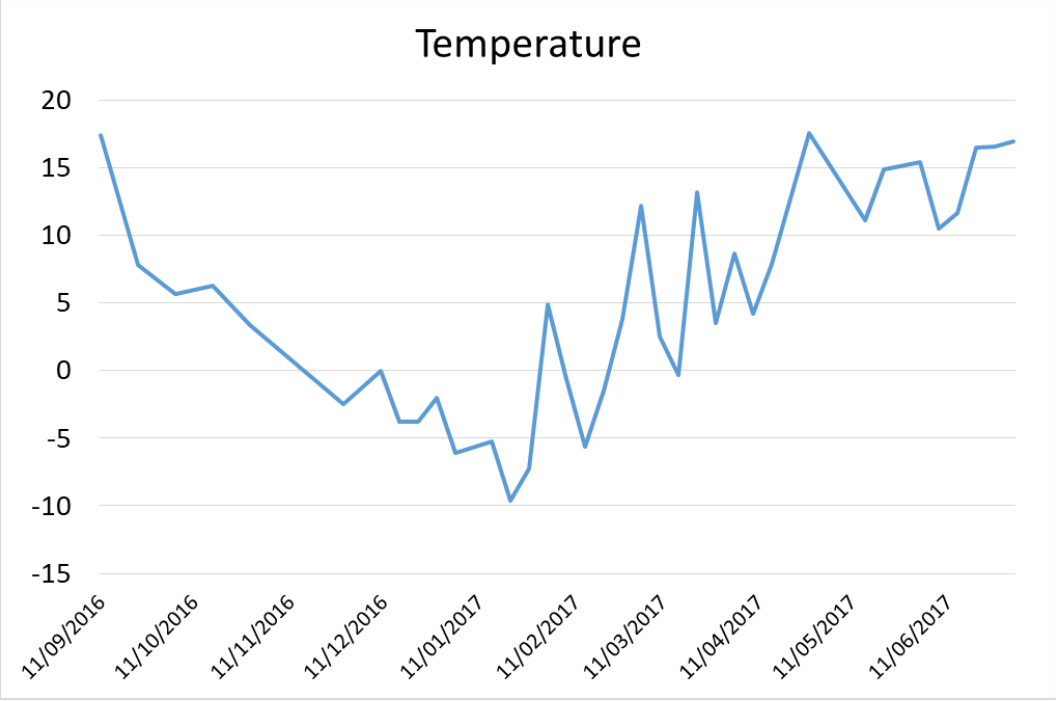


Figure 28 - Thermic regime in Banat pilot area. Negative values have not been considered in modelling (it means snow cover and frozen land surface that has negative influences on soil moisture estimation).

Investigations were focused on establishing the dry and wet references based on seasonal variation, which reflect vegetation phenology. This variation can induce backscatter changes that required vegetation correction. Thus, the change detection algorithm was used to scale dry and wet references of the backscattering coefficients over the whole SAR datasets in order to estimate surface soil moisture (Figure 29).

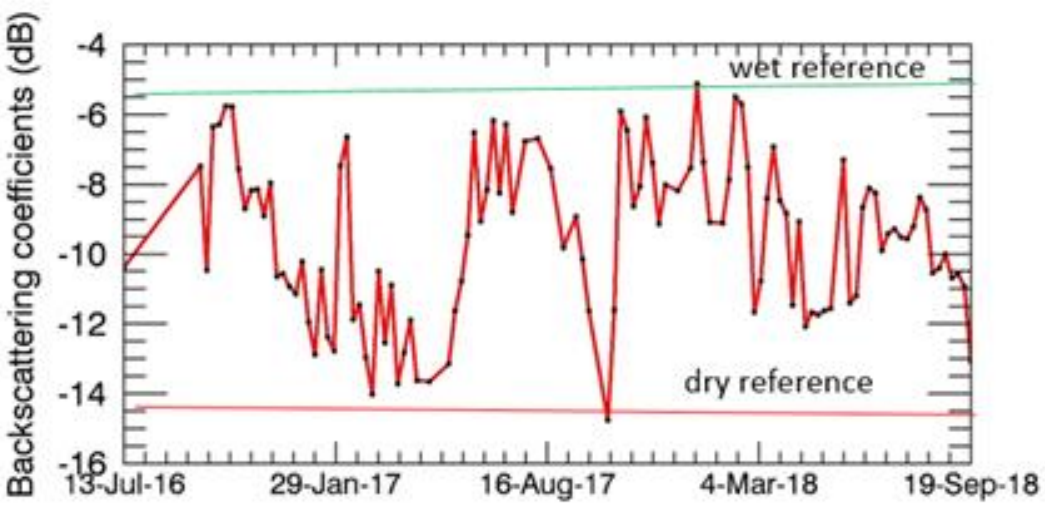


Figure 29 - Dry and wet reference characteristics to Banat pilot area.

Temporal evolution of the vegetation indices like NDVI and LAI was investigated in order to determine the contribution from soil and vegetation in the radar backscatter response and to validate the soil moisture results. Since no a priory information is considered, the performance of the surface soil moisture estimation is assessed as a function of LAI, NDVI and SAR acquisition far from a rainy episode or in the context of high air temperature recorded. The result is depicted in Figure 30.

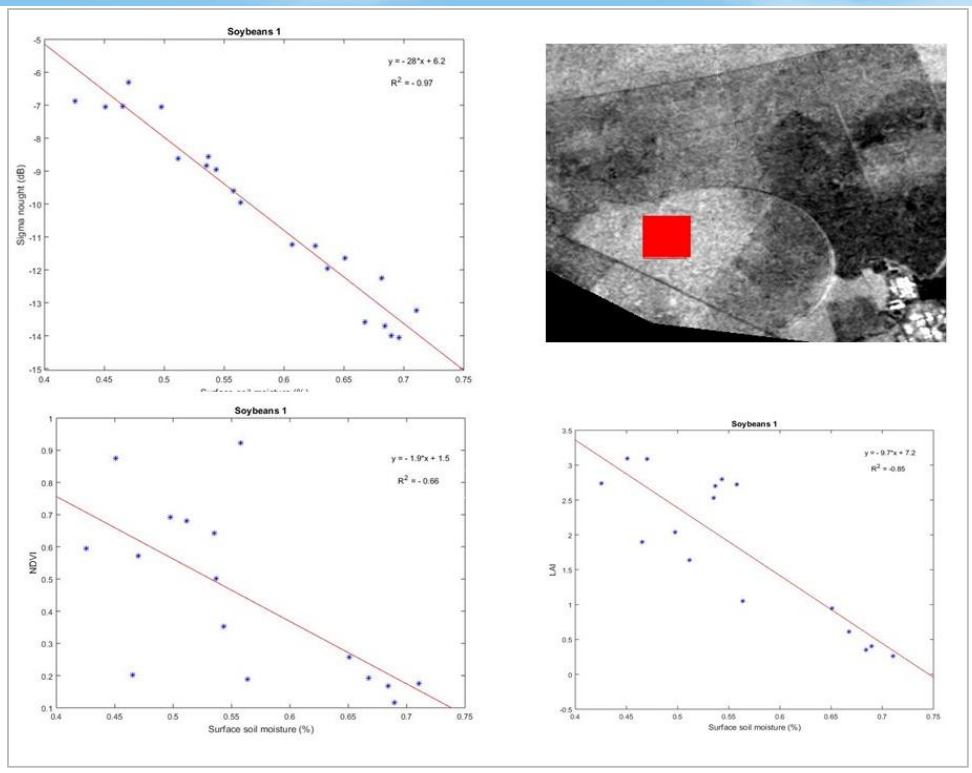


Figure 30 - Surface soil moisture for Soybeans area 1. Good correlation is observed between vegetation indices and estimated surface soil moisture (0.85 for LAI and 0.66 for NDVI).

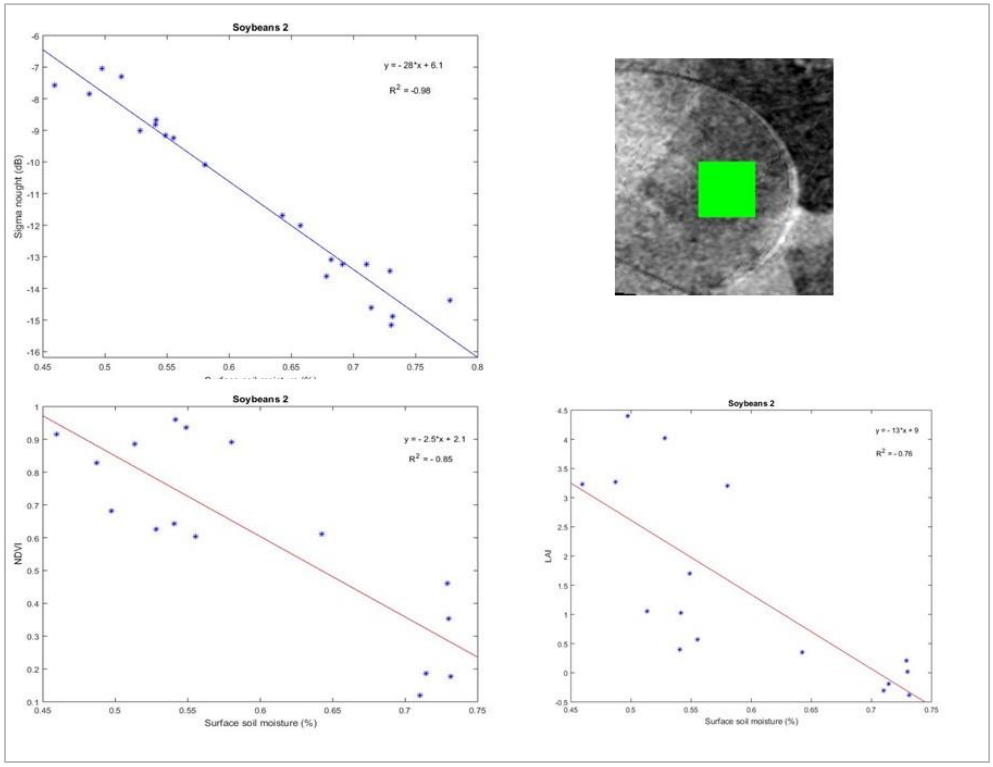


Figure 31 - Surface soil moisture for Soybeans area 2. Good correlation is observed between vegetation indices and estimated surface soil moisture (0.85 for NDVI and 0.76 for LAI).

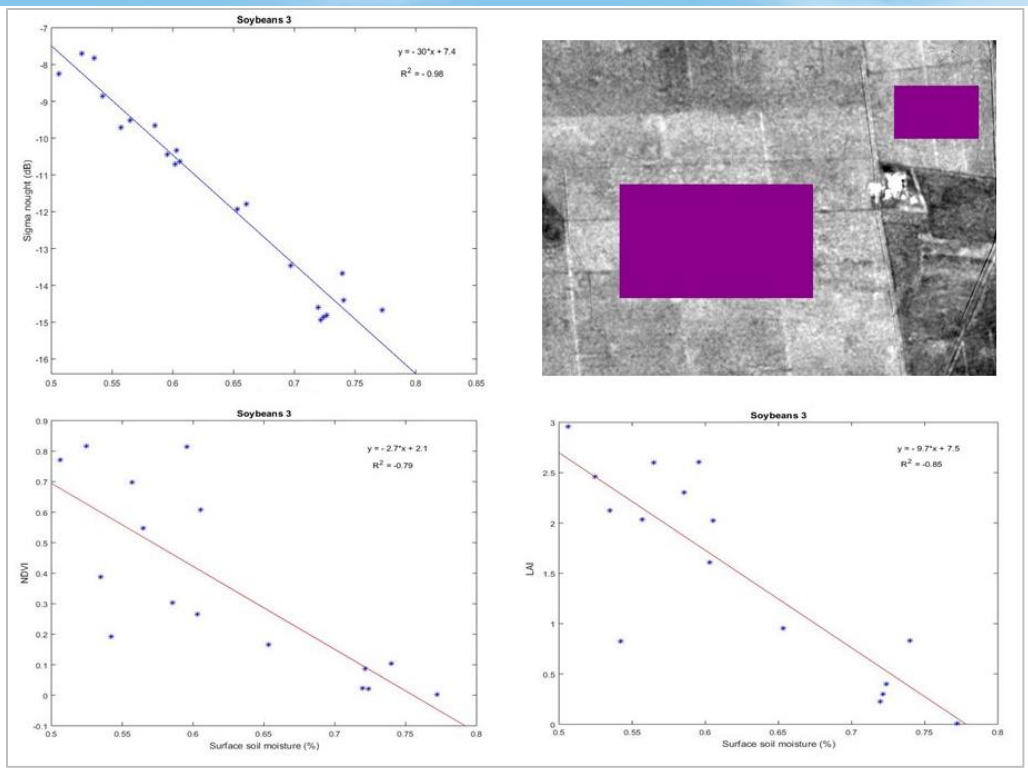


Figure 32 - Surface soil moisture for Soybeans area 3. Good correlation is observed between vegetation indices and estimated surface soil moisture (0.85 for LAI and 0.79 for NDVI).

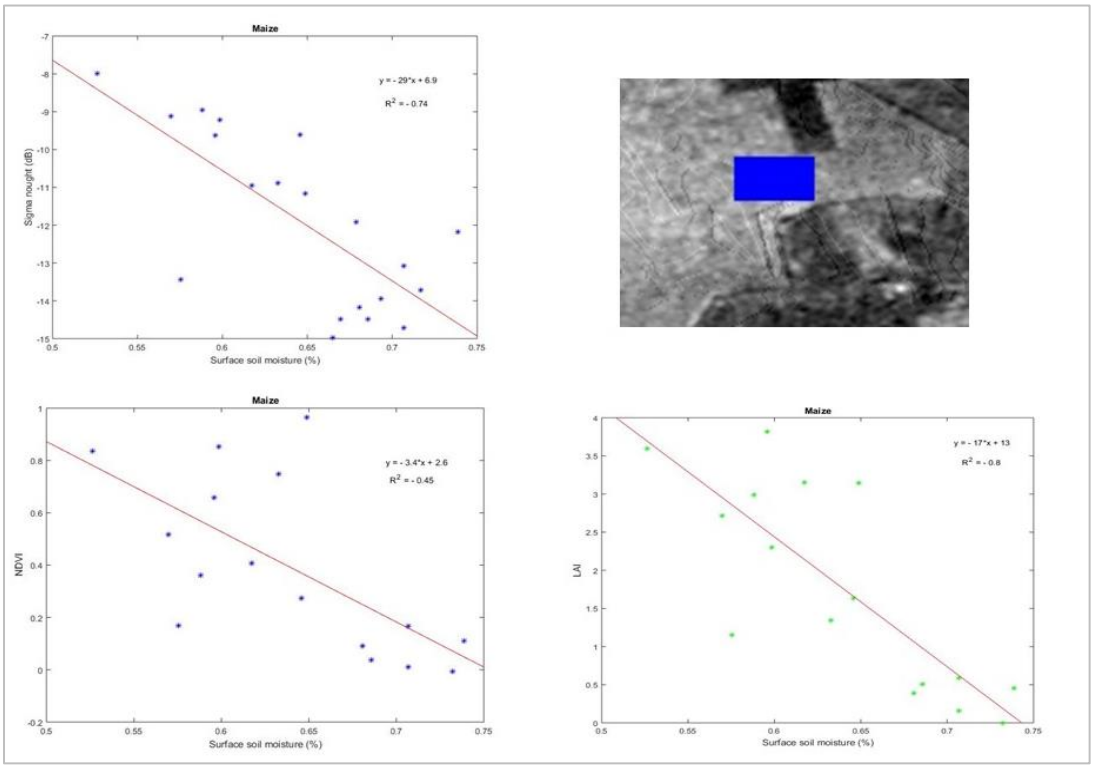


Figure 33 - Surface soil moisture for maize. Good correlation is observed between vegetation indices and estimated surface soil moisture (0.8 for LAI and 0.45 for NDVI).

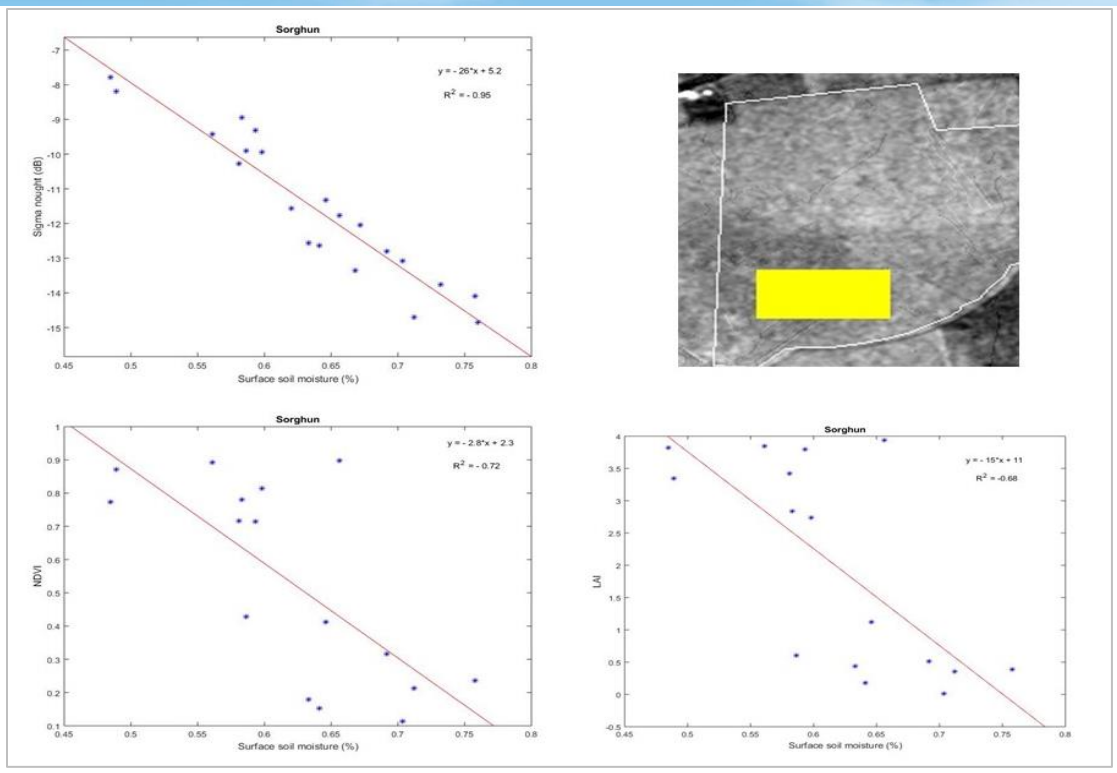


Figure 34 - Surface soil moisture for sorghum. Good correlation is observed between vegetation indices and estimated surface soil moisture (0.68 for LAI and 0.72 for NDVI).

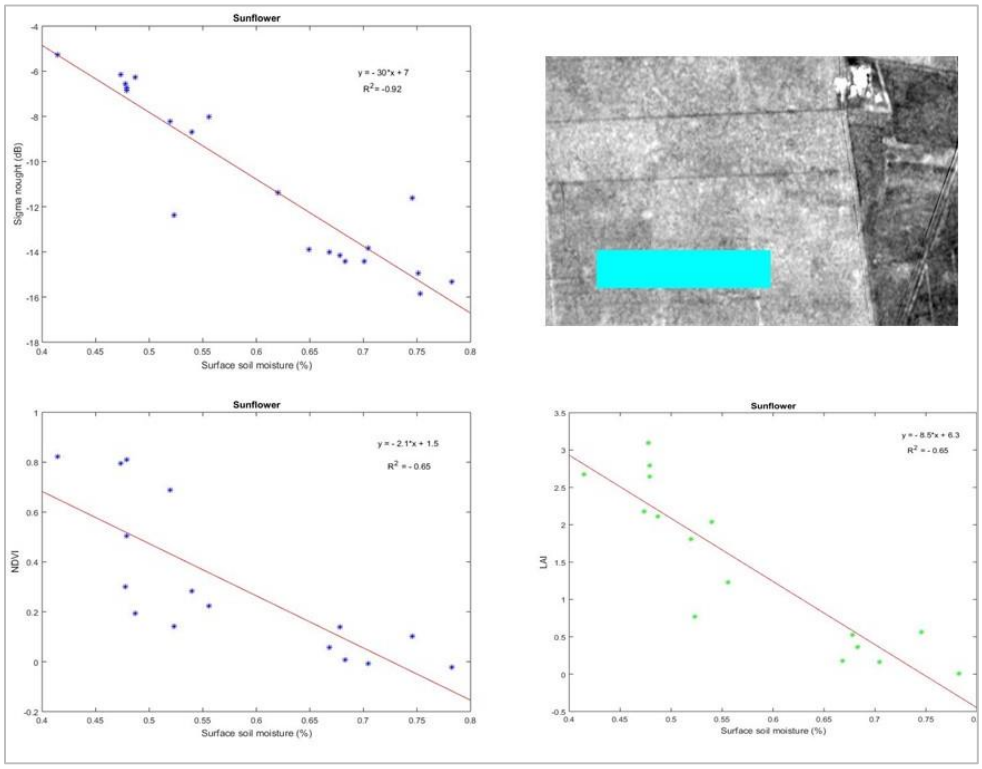


Figure 35 - Surface soil moisture for sunflower. Good correlation is observed between vegetation indices and estimated surface soil moisture (0.89 for LAI and 0.78 for NDVI).

These estimates could be validated using crossing methods like OPTRAM and TOTRAM since in-situ measurements require data to be collected at the same time with satellite passage.

Sentinel-1 data has been processing in ENVI SarScape 5.4 in order to reduce bias noise induced by the Sentinel-1 sensor. Backscattering coefficients calculated with ESA SNAP open source software have a bias of - 3dB comparing with ENVI SarScape.

6 Verification of the Meteorological Products

6.1 Introduction

Forecast verification is the process and practice of determining the quality of forecasts in the spatial and temporal domain[33]. This section describes the methodology that followed to verify weather forecast fields and the gridded observational data produced by the statistical and dynamical downscaling processes during the project's 1st pilot phase.

6.2 Data and Methods

For the verification of the meteorological parameters, the MET[34] code (Model Evaluation Tools) version 5.2 was used. MET is a collection of functions written in C that was developed by the Developmental Testbed Center to provide to the atmospheric community highly configurable and state-of-the-art verification tools. These tools are based on a variety of verification techniques, including standard verification scores for gridded and point statistics as well as more advanced statistical measures, such as neighbourhood, object-based and intensity-scale decomposition approaches for spatial statistics. The observational data that used against the meteorological products were coming from the MADIS[35] system.

The meteorological products of the temperature and humidity at 2 m height, the wind speed at 10m high, the daily accumulated precipitation and the occurrence of the precipitation event were compared against observational data from surface weather stations. The occurrence of a precipitation event was treated as a categorical variable, while all the other atmospheric fields were treated as continuous variables. To match the meteorological products and observations on the horizontal plane inverse distance weighted interpolation of 9 points was used for temperature, humidity and wind speed, while nearest neighbour interpolation was applied for precipitation.

[33] Murphy, Allan H., and Robert L. Winkler. (1987). A general framework for forecast verification. *Monthly Weather Review.* 115.7 (1987): 1330-1338.

[34] http://www.dtcenter.org/met/users/index.php

[35] https://madis.noaa.gov/

6.3 Statistical Measures of Verification of Continuous Variables

The statistical measures that used to verify the continuous atmospheric fields of temperature and humidity at 2m high, wind speed and direction at 10m high, and the amount of daily precipitation are the following:

- **Mean Absolute Error (MAE):** Mean absolute error is a measure of the absolute error between forecast and observation. MAE is less influenced by large errors and does not depend on the mean error. A perfect forecast would have MAE=0.
- **Mean Error (ME):** Mean error is a measure of overall bias of the forecast. An identical forecast would have ME=0.

Statistical Measures of Verification of Categorical Variables

The statistical measures that were used to verify discontinuous atmospheric fields, like precipitation are describing below:

- **Probability of Detection (POD)**: Probability of detection is the fraction of the events where correctly forecasted to occur. POD ranges between 0 and 1, and the best forecast has a POD value equal to 1.
- **False Alarm Rate (FAR):** False alarm rate is the proportion of forecasts of the event occurring for which the event did not occur. FAR ranges between 0 and 1 and the best forecast have a FAR value equal to 0.
- **Equitable Threat Score (ETS):** Equitable threat score is based on CSI, corrected for the numbers of hits that would be expected by chance. ETS ranges from -1/3 to 1. The perfect forecast would have an ETS equal to 1, and a no skill forecast would have an ETS equal to 0.

Results for the Continuous Variables

Lead Time (hours)	ME	ME_BCL	ME_BCU	MAE	MAE_BCL	MAE_BCU
12	1.558	-0.684	3.693	8.989	7.787	10.226
24	2.792	1.350	4.409	8.867	7.998	9.839
36	9.666	7.784	11.354	12.356	10.872	13.872
48	11.470	9.656	13.355	13.861	12.489	15.274
60	1.639	-0.547	3.734	8.134	6.924	9.395
72	2.937	1.450	4.553	9.047	8.181	10.002
84	9.040	6.938	10.881	12.052	10.684	13.423
96	10.513	8.725	12.432	13.059	11.636	14.492
108	0.026	-2.478	2.299	9.352	7.854	10.789
120	2.108	0.220	3.914	9.596	8.565	10.655
132	7.886	6.024	9.823	11.673	10.205	13.199
144	10.824	8.997	12.676	13.014	11.615	14.393
156	-1.764	-4.580	0.896	10.349	8.627	12.127
168	4.430	-4.218	11.591	12.841	7.604	18.736

Table 9 - Relative Humidity at 2m Height (All Pilots).

Lead Time (hours)	ME	ME_BCL	ME_BCU	MAE	MAE_BCL	MAE_BCU
12	0.086	-0.243	0.394	1.374	1.206	1.551
24	0.310	0.118	0.516	1.057	0.937	1.178
36	-1.074	-1.366	-0.779	1.670	1.486	1.885
48	-1.051	-1.294	-0.800	1.500	1.321	1.693
60	0.369	0.007	0.703	1.474	1.276	1.678
72	0.333	0.122	0.525	1.129	1.005	1.266
84	-0.812	-1.120	-0.520	1.611	1.414	1.812
96	-0.803	-1.056	-0.527	1.485	1.311	1.656
108	0.653	0.318	1.029	1.528	1.311	1.760
120	0.657	0.448	0.865	1.259	1.116	1.404
132	-0.408	-0.724	-0.078	1.617	1.413	1.845
144	-0.559	-0.840	-0.297	1.431	1.235	1.616
156	0.927	0.561	1.340	1.690	1.434	1.962
168	0.893	0.650	1.151	1.450	1.274	1.619

Table 10 - Temperature at 2m Height (All pilots).

Lead Time (hours)	ME	ME_BCL	ME_BCU	MAE	MAE_BCL	MAE_BCU
12	1.255	0.941	1.604	1.681	1.453	1.925
24	1.123	0.834	1.422	1.800	1.596	1.996
36	0.923	0.595	1.266	1.855	1.625	2.091
48	1.173	0.816	1.524	1.957	1.684	2.233
60	1.455	1.097	1.801	1.872	1.615	2.149
72	1.429	1.092	1.773	2.118	1.894	2.365
84	1.346	0.999	1.707	2.172	1.941	2.411
96	1.244	0.822	1.648	2.088	1.805	2.398
108	1.576	1.186	1.940	1.812	1.504	2.148
120	1.527	1.180	1.870	2.198	1.958	2.442
132	1.300	0.926	1.717	2.163	1.900	2.446
144	1.030	0.620	1.389	1.962	1.719	2.225
156	1.527	1.156	1.931	1.908	1.624	2.230
168	1.706	1.379	2.053	2.340	2.097	2.620

Table 11 - Wind Speed at 10m Height (All pilots)

Lead Time (hours)	ME	ME_BCL	ME_BCU	MAE	MAE_BCL	MAE_BCU
12	-13.871	-33.155	5.725	78.054	66.103	88.798
24	-4.608	-18.475	9.837	73.998	65.894	82.311
36	-24.428	-44.040	-4.903	97.538	85.264	110.088
48	-17.070	-33.671	1.233	89.558	78.570	100.931
60	-47.755	-68.621	-27.251	91.992	75.243	107.672
72	-27.691	-43.531	-12.404	88.220	78.250	97.968
84	-59.218	-82.263	-37.047	128.246	112.502	143.962
96	-15.734	-34.868	1.942	94.524	82.359	106.333
108	-67.598	-91.806	-42.770	106.212	89.615	124.668
120	-47.160	-62.666	-32.135	89.631	80.008	99.363
132	-80.150	-101.550	-59.200	130.566	116.889	145.161
144	-23.094	-42.211	-5.299	92.394	81.067	104.963
156	-90.272	-114.244	-68.040	113.794	96.046	132.418
168	-51.924	-71.125	-32.222	109.461	96.343	123.235

Table 12 - Mean Sea Level Pressure (Pa) (All pilots).

Results of the Categorical Variables

Lead Time (hours)	PODY	FAR	ETS
24	0.737	0.650	0.234
48	0.909	0.608	0.273
72	0.889	0.686	0.222
96	0.818	0.633	0.234
120	0.889	0.600	0.311
144	0.810	0.685	0.185
168	0.722	0.667	0.222

Table 13 - Precipitation (All pilots).

7 Annexes

7.1 Background of the ET comparison performed in the Italian Pilot area.

Comparison at pixel scale between ET estimated by two different algorithms based on the logic of P-M equation: the global MODerate Resolution Imaging Spectroradiometer (MODIS) ET product (MOD16 ET) and the "Analytical Approach" (D'Urso et al., 2006)[36]. Both of them use as input a combination of daily meteorological data and remotely sensed vegetation properties and not require thermal infrared remote sensing of Land Surface Temperature (LST) as input data. However, they have been developed for different purpose and applications and usually use different sources of data input. MOD16 ET is a global product and provides key information useful to water resource management and to calculate regional water and energy balance and soil water status from the regional to the global scale (Mu et al., 2007)[37]. Due to its long-term data, it allows to quantify the effects of changes in climate, land use and ecosystems disturbances on regional water resources and land surface energy change. MODIS ET product was validated worldwide with observed latent heat flux for 46 field-based eddy covariance towers, from 232 global watersheds for each of the seven biome types (Mu et al., 2011)[38]. MOD16 ET algorithm uses as input vegetation properties (LAI and albedo) derived from the same MODIS sensors. The Analytical approach is an operative method to estimate the ET at moderate-high spatial resolution It exploits agrometeorological data measured in situ and the crop parameters (crop height - hc, albedo – α - and Leaf Area Index - LAI) estimated from remotely sensed data. It is a consolidated method applied in numerous studies (Akdim et al., 2014)[39] and in numerous satellite-based irrigation advisory services (Vuolo et al., 2015)[40] thanks to its applicability in areas with limited field measurements and without knowledge of crop type. However, it has only site-specific validation. The main goal of this task is to compare the ET retrieved by these two different methods by using as input the same vegetation parameters retrieved by different satellites (LAI and albedo MODIS for MOD16 ET products and LAI and albedo Sentinel-2 for ET_p Sentinel-2) during the irrigation

[36] D'Urso, G., Calera Belmonte, C. (2006). Operative approaches to determine Crop Water Requirements from earth observation data: methodologies and applications, AIP Conference Proceedings (2006).

[37] Mu, Q., Heinsch, F., Zhao, M. end Running, S. (2007). Development of a global evapotranspiration algorithm based on MODIS and global meteorology data. Remote sensing of Environment, 519-536 (2007).

[38] Mu, Q., Zhao, M., Running, S. (2011). Improvements to a MODIS global terrestrial evapotranspiration algorithm. Remote Sensing of Environment, 1781-1800, (2011).

[39] Akdim, N., Alfieri, S., Habib, A., Choukri, A. (2014). Monitoring of irrigation schemes by remote sensing: phenology versus retrieval of biophysical variables. Remote Sensing, 5815-5851, (2014).

[40] Vuolo, F., D'Urso, G., De Michele, C., Bianchi, B., Cutting, M. (2015). Satellite-based irrigation advisory services: A common tool for different experiences from Europe to Australia. Agricultural water management, 82-95 (2015).

season (April to September) of the year 2018. The evaluation is carried out on the Italian Pilot area "Sannio Alifano" irrigation district, characterised by an extremely heterogeneous and fragmented landscape and a typically Mediterranean climate.

MOD16 algorithm

The MODIS sensor resides aboard the Terra and Aqua platforms, offering a view of the earth's surface every 1-2 day. Scientists from a variety of disciplines, including oceanography, biology, use standard MODIS products and atmospheric science and are useful to both global change research and resource management. The MOD16 algorithm estimates the total daily ET as the sum of evaporation from the wet canopy surface, the transpiration from the dry canopy surface and the evaporation from the soil surface. The total daily latent heat flux (λE, [Wm^{-2}]) and the potential ET (λE_{pot}) are calculated as the sum of the evaporation from the wet canopy surface ($\lambda E_{wet,c}$), the transpiration from the dry canopy (λE_{trans}) and the evaporation from the soil (λEsoil):

$$\lambda E = \lambda E_{wet,c} + \lambda E_{trans} + \lambda E_{soil} \quad (19)$$

$$\lambda E_{pot} = \lambda E_{wet,c} + \lambda E_{pot,trans} + \lambda E_{wet,soil} + \lambda E_{pot,soil} \quad (20)$$

Actual plant transpiration estimation is based on the logic of the Penman-Monteith (Monteith, 1965)[41] equation, while the potential is calculated following the Priestley-Taylor method (Priestley et al., 1972)[42]:

$$\lambda E_{trans} = \frac{sA + \rho C_p (e_{sat} - e)/r_a}{s + \gamma(1 + \frac{r_s}{r_a})} \quad (21)$$

$$\lambda E_{pot,trans} = \frac{1.26\, sA_c}{s + \gamma} \quad (22)$$

where, s [Pa K^{-1}] is the slope of the curve relating saturated vapour pressure (e_{sat}, [Pa]) to temperature [K]; A [W m^{-1}] is the available radiative energy portioned between sensible and latent heat fluxes on land surface and A_c is the part of A allocated to the canopy; ρ [Kg m^{-3}] is the air density; Cp [J Kg^{-1} K^{-1}] is the specific heat of air; γ [Pa K^{-1}] is the psychrometric constant and r_a and r_s [s m^{-1}] are respectively the aerodynamic and the surface resistance. The aerodynamic resistance

[41] Monteith, J. Evaporation and environment. Symposia of the Society for Experimental Biology (1965).

[42] Priestley, C., Taylor, R. (1972). On the assessment of surface heat flux and evaporation using large-scale parameters. Monthly weather review (1972).

is calculated following the Biome-BGC model (Thornton, 1998)[43] as parallel resistance to conductive (biome dependent) and radiative (function of the air temperature) resistance and therefore not includes the effects of the wind speed. The surface resistance is the effective resistance to evaporation from land surface or transpiration from vegetation canopy. In the case of bare soil, it is not calculated, while the canopy resistance is estimated as:

$$r_{s,canopy} = \frac{1}{[C_L m(T_{min}) m(VPD) LAI]} \quad (23)$$

where, C_L [m s^{-1}] is the potential stomatal conductance per unit leaf area when all environmental parameters are at the optimum levels (full stomatal opening condition), and $m(T_{min})$ and m(VDP) are two multipliers that limit the stomatal conductance depending on minimum air temperature and vapour pressure deficit (VPD, [Pa]).

The "Analytical approach"

The "Analytical Approach" (D'Urso and Menenti, 1995)[22] is also referred in the Food and Agriculture Organization (FAO) in the Irrigation and Drainage Paper No. 56 (Allen et al., 1998)[23] as the "one-step" or "direct" approach. The FAO-56 did not provide specific means for estimating crop ET (ET_c) from satellite imagery. However, since it was published, substantial progress and applications have been obtained in ET_c estimation by remote sensing (RS) (Pereira et al., 2015)[44]. The Analytical Approach is an operative approach for mapping the ET_c by using ground-based standard meteorological data (solar radiation R_s, wind speed U, air temperature T and humidity RH) and crop-specific vegetation parameters obtained from RS: the surface albedo (α), the leaf area index (LAI), the crop height (h_c) and the stomatal resistance (r). The calculation is performed at the pixel level, as shown in the following equation:

$$ET = f(\alpha, LAI, h_c, r, R_s, U, T, RH) = \frac{s(R_n - G) + \rho C_p (e_{sat} - e)/r_a}{s + \gamma(1 + \frac{r_s}{r_a})} \quad (24)$$

where, the terms R_n-G represents the available radiative energy. For the estimation of the potential ET ($ETp_{analytical}$) were used the LAI and Albedo retrieved from the Sentinel-2. For the remains, two crop parameters (r and h_c) a simplified estimation were introduced. The stomatal

[43] Thornton, P. (1998). Regional ecosystem simulation: combining surface-and satellite-based observations to study linkages between terrestrial energy and mass budgets. Graduate Student Theses, Dissertations, Professional Papers. 10519. https://scholarworks.umt.edu./etd/10519.

[44] Pereira, L., Allen, R., Smith, M., Raesd, D. (2015). Crop evapotranspiration estimation with FAO56: Past and future. Agricultural Water Management, 147, 4-20.

resistance describes the average resistance of vapour flow through an individual well-illuminated leaf. This resistance is crop specific and differs among crop varieties and crop management. Moreover, it is also influenced by climate and water availability. Thus, estimating the stomatal resistance is one of the major complicated issues in the ET estimation, and the information available in the literature is often oriented toward physiological or eco-physiological studies. In the procedure proposed by D'Urso et al. (2006)[36] the ET estimation was performed assuming that the canopy resistance is at the minimum value (r=70 ms^{-1}). Therefore, the estimated ET represents the maximum value at "potential condition" from a well-watered and diseases free crop during the active growing stage.

The crop height, together with the LAI, influences the aerodynamic proprieties of vegetation canopy. It is strictly related to the crop type and the estimation of the crop height, D'Urso et al. (2006)[36] suggests using a mean crop height (constant) value. Later studies conducted in semi-arid areas confirm this assumption: Aghdasi (2010)[45] demonstrates that a percentage change of 50% in h_c corresponds to a variation of the order of 5% of ET. Therefore, in this contest was used a constant value of 0.5 m, valid for the Mediterranean climatic conditions [Petropoulos et al., 2018)[46]. This last assumption makes it possible to assess the ET also without a previous land cover classification in term of crop height.

Satellite data

In the present task, MODIS ET product and Sentinel-2 data input (LAI and albedo) were used. For MODIS ET product were analysed 24 images. For Sentinel-2 were analysed 47 images, used to derive LAI and albedo product for ET_p calculation. ET_p was calculated at the pixel level (10m) and daily time step, subsequently aggregate at MODIS time step (8-days) and spatial resolution (500 m) during the irrigation season 2018. Observations from the MODIS sensors are used to derive operationally global land surface evapotranspiration product (MOD 16) on 8-day (MYD16A2). The 8-day composite product (Collection 6) is the improved level 4 MODIS land data product computed globally every day at 500 meters spatial resolution. The MODIS ET algorithm, following a technique proposed by Mu et al. (2007)[37], runs at daily basis and uses as input a combination of daily global meteorological data and remotely sensed vegetation property dynamics (Cleugh et

45 Aghdasi, F., Sharifi, M., Van der Tol, C. (2010). Assessment of crop water requirement methods for annual agricultural water allocation planning. EGU General Assembly Conference Abstracts. Vol. (12), 2010.
46 Petropoulos, G., Srivastava, P., Piles, M. Pearson, S. (2018)**.** Earth Observation-Based Operational Estimation of Soil Moisture and Evapotranspiration for Agricultural Crops in Support of Sustainable Water Management. Sustainability, 10.1, 2018, 181.

al., 2007)[47]: the albedo (MCD43A3) and the MODIS Leaf Area Index/FPAR (MOD15A2H). While the algorithm runs on a daily basis, it assumes that LAI and albedo do not vary during the compositing period. The daily meteorological data are provided by NASA's Global Modeling and Assimilation Office (GMAO). These data are produced every six hours using a global circulation model and both ground and satellite-based observations. They are distributed at the spatial resolution of 0.5° x 0.6° and hence are spatially smoothed to 0.5 km MODIS pixel level using a cosine function for the spatial data interpolation. Provided in the MOD16A2 product, there are layers for composited Evapotranspiration (ET), Latent Heat Flux (LE), Potential ET (PET) and Potential LE (PLE) along with a quality control layer (QC). The QC data layer directly inherits the QC data field from the corresponding LAI QC of the same 8-day and denotes if filled data input were used (Mu et al., 2007)[37]. The pixel values for the two Evapotranspiration layers (ET & PET) are the sum of all eight days within the composite period, and the pixel values for the two Latent Heat layers (LE & PLE) are the average of all eight days within the composite period. In this contest were used only the 8-day ET and PET layers.

Meteorological data

For this task was used ERA-Interim dataset. ERA-Interim is a dataset, showing the results of a global climate reanalysis from 1979 to date. ERA stands for 'ECMWF Re-Analysis' and refers to a series of research projects at ECMWF which produced various datasets (ERA-Interim, ERA-40, etc). ERA-Interim "runs" in near real time, data is published with a few months delay. The NWP system blends, or assimilates observations with a previous forecast to obtain the best fit for both. The result of this blending is called an analysis and is the starting point for the next forecast. In this manner, data is produced at increasingly later times. Analysed data is described as instantaneous, though it does represent an average over the model time step (30 minutes for ERA-Interim). Depending on the parameter, forecast data in ERA-Interim is either instantaneous or accumulated from the beginning of the forecast (twice daily forecasts starting at 00 and 12 UTC). Parameters such as precipitation and radiation are accumulated. The ERA-Interim dataset contains atmospheric and surface parameters: 6-hourly atmospheric fields on model levels, pressure levels, potential temperature and potential vorticity; 3-hourly surface fields and daily vertical integrals; monthly averages of daily means; monthly synoptic averages at 0 UTC, 6 UTC, 12 UTC, 18 UTC. The temporal coverage is from 1 January 1979 to present; the spatial coverage is global;

47 Cleugh, H., Leuning, R., Mu, Q., Running, S. (2007). Regional evaporation estimates from flux tower and MODIS satellite data. Remote Sensing of Environment, 2007, 106.3: 285-304.

the native horizontal resolution is ~80 km. The ERA-Interim dataset contains analyses (four times per day, at 00:00, 06:00, 12:00 and 18:00), as well as forecasts (from 00:00 and 12:00, with 3, 6, 9, and 12-hour steps, and more, into the future). Variables of the ERA-Interim data for the calculation of the ET_p is reported in the Table below.

Short name	Variables	units
SP	Surface pressure	(Pa)
Tp	Total Precipitation	(m)
SSRD	Surface solar radiation downwards	(W m^{-2} s)
STR	Surface thermal radiation	(W m^{-2} s)
T2	2-meter temperature	(K)
D2	2 meter dewpoint temperature	(K)
U10	10 meter U wind component	($m.s^{-2}$)
V10	10 meter V wind component	($m.s^{-2}$)

Table 14- Variables of the ERA-Interim data for the calculation of the ETp.

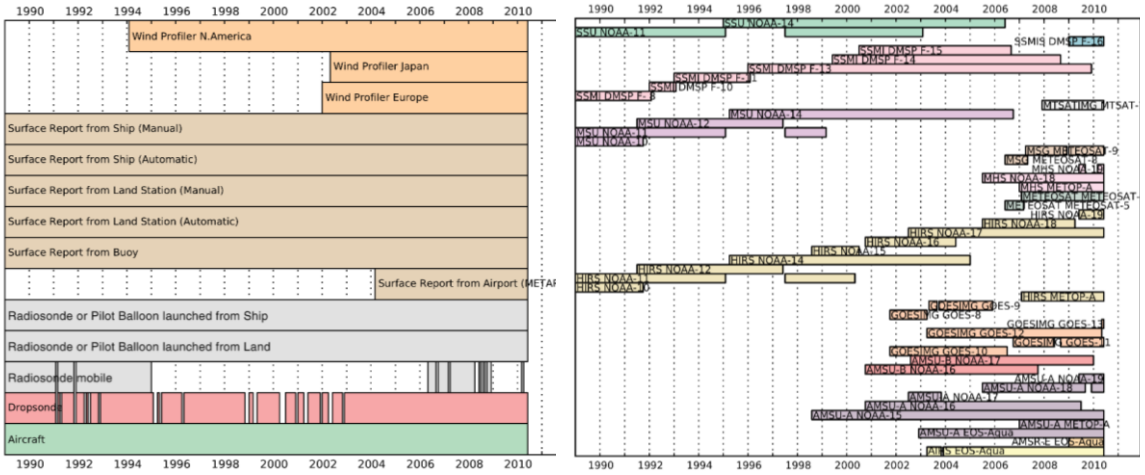


Figure 36 - Timeline of conventional observations assimilated (DEE, Dick P., et al. The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. Quarterly Journal of the royal meteorological society, 2011, 137.656: 553-597.).