Agricultural Monitoring, an Automatic Procedure for Crop Mapping and Yield Estimation: The Great Rift Valley of Kenya Case

Roberto Luciani, Giovanni Laneve, *Member, IEEE*, and Munzer JahJah

***Abstract*—Agricultural activities conducted in the Great Rift Valley of Kenya show a significant decline of productivity levels. This phenomenon is mainly related to limited availability of water resources, lack of supporting irrigation, and harvesting techniques ineffectiveness. Production risks reduction is closely related with a better use of water resources and a better understanding of the effects resulting from the multiple interactions between climate, agricultural vegetation, soil type, and crops management techniques. In this paper, a remote and automatic agricultural monitoring system is presented as an effective alternative to the most traditional *in situ* measurements and observations. We investigated the use of phenological information extracted from satellite imagery combined with crop calendar and supported by agro-ecological zoning (AEZ) in accurate crop classification and monitoring. Vegetation indices extracted from Landsat 8 imagery are capable to track the vegetation development through the year, then phenological profiles can be extracted and implemented into a multitemporal automatic classification process to detect agricultural areas and to discriminate among different crop species. The phenological profiles extracted by satellite imagery are compared with crop calendar data compiled by FAO for the area of interest. The classification procedure is supported by AEZs based on crop modeling and environmental matching procedures in order to identify crop-specific environmental limitations under assumed levels of inputs and management conditions. The FAO crop water productivity model AquaCrop is calibrated for wheat and maize yield mapping in the central highland of Kenya, handling both environmental and phenological data. The combined use of phenological data and AEZs results in a robust methodology with a classification overall accuracy of 91.35%. A good model performance is obtained relative to yield predictions, with *R* of 0.69 and 0.72.**

***Index Terms*—Agricultural monitoring, Aquacrop, crop classification, Kenya, Landsat 8, phenology, remote sensing, yield mapping.**

I. Introduction

AGRICULTURE is the most important economic activity in Kenya, although less than 8% of its territory is used for crops and sheep farming. Less than 20% of the land is suitable for agriculture, of which only 12% is classified as high agricultural potential (adequate precipitation) and about 8% of the soil as medium potential. The rest of the territory is arid or semi-arid. About the 80% of the workforce is engaged in agriculture or food production [1], [2]. A detailed land use-land cover (LULC) map of Kenya was released in 2000 by the FAO in the framework of the Africover project. This map was mainly based on Landsat imagery acquired since 1995 [3] and was specifically dedicated to agricultural areas identification and description. Since then, socioeconomic and political changes have brought about major changes that dealt with the intensification and the systematic expansion of agro-pastoral areas. The periodic production of LULC maps is necessary to understand the changes that have occurred in the last two decades and for the long-term prediction of the agricultural potential. Agricultural monitoring systems must be able to recognize agricultural areas, discriminate among different crop types, and finally evaluate crop health status. The identification of crop stress factors stands as a critical point in order to correct crop growth simulation models and properly estimate the expected crop yield.

Site-specific crop identification based on satellite imagery is important for agro-ecological analysis, climate modeling at regional scale, and agricultural policy development [4]. Timely and accurate data on crop yields both at regional and field scales are important for many applications: At a regional scale, it is an input into the decision-making process on food security issues, and at a field scale, it represent a site-specific input to improve field-management practices. The critical factor is the time needed to collect all the necessary data to properly carry out the monitoring and forecasting activities. Currently, data-collection systems are based on ground controls and air surveillance. These systems provide timely and accurate data, but show several disadvantages: Data are collected by various entities over extended territories, which results in the application of different methodologies and tools; furthermore, these strategies are time consuming and they require a significant amount of economic investments. A remote crop monitoring system is a significant step toward the introduction of standard measurements and uniform validation checks of cultivation supporting techniques effectiveness.

Remote sensing technology has been proved as a useful application for natural resources evaluation and management. Automated and low-cost remote sensing tools are suited for

Manuscript received October 9, 2018; revised April 1, 2019 and May 16, 2019; accepted May 20, 2019. Date of publication June 24, 2019; date of current version July 30, 2019. This work was supported by the Italian Space Agency in the framework of the ASI-Sapienza agreement, Satellite Based Agricultural Monitoring (SBAM) project. *(Corresponding author: Roberto Luciani.)*

R. Luciani and G. Laneve are with the Earth Observation and Satellite Image Applications Laboratory (EOSIAL), School of Aerospace Engineering (SIA), Sapienza University of Rome, Via Salaria 851-881, 00138 Italy (e-mail: roberto.luciani@uniroma1.it; giovanni.laneve@uniroma1.it).

M. JahJah is with the Italian Space Agency, Via del Politecnico, Rome 00133 Italy (e-mail: munzer.jahjah@asi.it).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSTARS.2019.2921437

monitoring crop growth and health status and providing farmers with timely information concerning the necessary actions to take to assure food security. Vegetation indices (VIs) derived from satellite images are well correlated with the parameters that define crops status and for nearly four decades the normalized difference VI (NDVI) has remained one of the most consistently and widely measured VIs across a wide variety of space-borne sensors, providing us with plurality of information and historical archives. NDVI data at high temporal frequency have been widely used to track seasonal phenology of green-up and senescence over a large variety of ecosystems from space using NOAA's Advanced Very-High Resolution Radiometer (AVHRR) ([5]–[8]) and NASA's Moderate Resolution Imaging Spectrometer (MODIS) [9]–[11]. Earth Observation data are recognized for strategic relevance in monitoring crops health status and productivity, providing the necessary information to food security and early warning systems. However, as researchers discovered, VIs and/or spectral data by themselves are often not up to the challenge of crop species discrimination. Further, both AVHRR and MODIS provide high-frequency observations, but their spatial resolution, in several cases, could be quite coarse. MODIS data are available at 250-, 500-, and 1000-m spatial resolution depending on the specific product [12]. AVHRR data, on the other hand, are available at spatial resolutions of 1.1 km for local area coverage and 4 km globally. Higher spatial resolution data from sensors such as the Landsat Thematic Mapper (30 m) have also been used in agricultural applications [13], [14]. The repeating period for Landsat is relatively infrequent (16 days), which presents challenges for agricultural applications that rely on high-frequency sampling during critical phases of the crop growth cycle. Recent studies have fused higher spatial resolution map products derived from Landsat with higher frequency MODIS or AVHRR data (e.g., [15], [16]). However, high-resolution map products derived from Landsat data such as the cropland data layer (CDL) [17], which is produced by the United States Department of Agriculture (USDA) National Agricultural Statistics Service (NASS), are generally only produced at regional scales and are therefore not available for much of the world. Further, medium-/low-resolution maps are not worldwide applicable due to the different extent of the cropland characterizing each country and finally the accuracy of such maps is not the same across the world [18], [19].

Methods that rely on widely available datasets such as MODIS could be more or less useful depending on the characteristics of the area of interest (AOI). Decades of scientific research have shown considerable progress toward assessing land use cover change (LUCC) [20]. Using air- or space-borne remote sensing data is a fast-advancing approach in this field [21]–[24], particularly due to its ability to provide regular spatially and temporally explicit data across large areas when compared to field-based assessments [25]. Considerable challenges in mapping LUCC using remote sensing data persist; the data are not always uniquely linked to land cover and are ambiguously related to land use, hence commonly requiring the use of heuristic, empirical (e.g., [26], [27]), or physically based models [28] to infer land properties. Further, land-use information must often be inferred based on integration with ground knowledge or user interpretation [26], [29]. Reliable, regular, and extensive ground assessments are expensive and challenging, often constraining remote sensing for mapping unambiguous land-cover properties only. Studies often extracted phenological indices such as leaf area index (LAI), fraction of vegetation cover, enhanced VI, NDVI, and land surface water index (e.g., [30]–[33]), and various band ratios and differences [34], from optical data. Mapping vegetation at a species level is quite challenging using Landsat products, especially in heterogeneous environment characterized by high variability of soil type and crop quality.

Simulation models properly calibrated can play a key role in the analysis of crop adaptability to climatic and soil conditions, as well as the effectiveness of cultivation-supporting technique in managing critical stress factors that may occur at different stages of vegetation development [35]. Simulations analysis introduces the possibility to choose through different possible and adaptable planning strategies for the management of the agricultural potential of a specific area [36]. However, it is necessary to remark that a universal calibration for such simulation models does not exist [37]. The first objective to achieve is to calibrate the simulation model according to the climatic local variability, using the least possible number of parameters to this purpose and taking into account the most relevant vegetation stress factors within the study area [38]. The current crop development through the growing season is described and introduced into the simulation model through phenological metrics and variables retrieved during the classification task.

In Section II, a brief description of the study area, the reference data, and the location of the sample sites on the ground are provided, as well as a brief description of the satellite data that have been used in this paper. In Section III, our methodologies related with crop mapping and with crop yield forecast are presented. Section IV is dedicated to discuss the results and validation of the previously described methodologies. Section V of this paper discusses the final conclusions.

II. Study Area and Materials

A. Study Area and Reference Data

The study was performed for the Nakuru County in 2016. Nakuru County (Fig. 1) covers the largest part of the Rift Valley and includes the bordering escarpments and plateaus: The western plateau rises to nearly 3000 m within the district territory. Nakuru County is particularly suited for the development of agro-pastoral activities. The climate is cool and wet with a mean temperature of 10–15 °C and annual average rainfall of about 1200–1400 mm [39]. Nakuru County, according to FAO agro-ecological zoning (AEZ), is divided into tropical alpine (TA), upper highland (UH, from 0 to 3), lower highland (LH, from 2 to 5), upper midland (UM, from 3 to 6), and lower midland (LM 5) zones.

ERA-Interim is a global analysis system of climatic data since 1979. The system is characterized by a four-dimensional variational analysis in a 12-h time window, with a spatial resolution of about 80 km distributed over 60 altitude levels (from the surface up to 0.1 hPa). Climatic data are available from the ECMWF Public Datasets web interface [40]. The archives made available

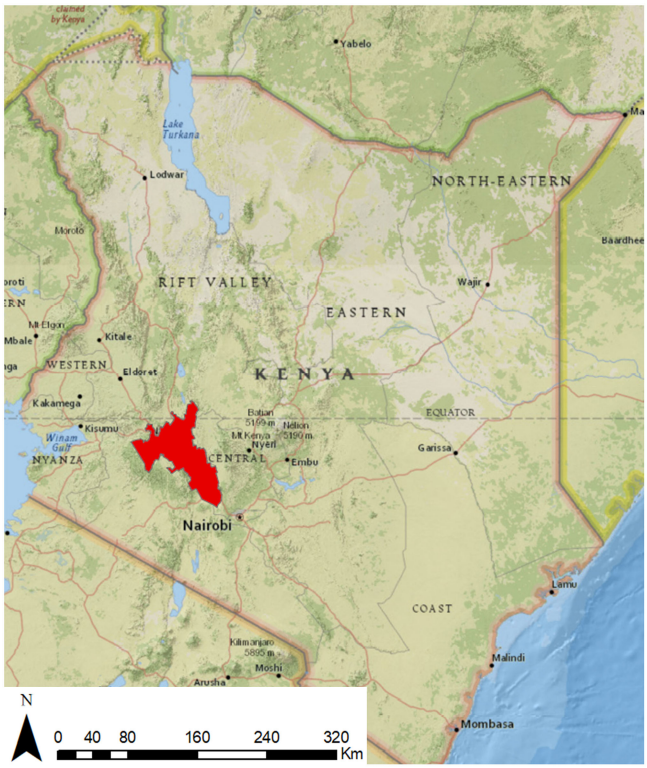

Fig. 1. Nakuru County (in red), in the Great Rift Valley of Kenya, has been selected as study area for this investigation.

TABLE I
SAMPLE SITES

Sample	Type	Lon	Lat
Sample field n. 1	Maize	35.92187 E	0.30612 N
Sample field n. 2	Maize	35.94662 E	0.34212 N
Sample field n. 3	Maize	35.92412 E	0.35112 N
Sample field n. 4	Maize	35.92412 E	0.35337 N
Sample field n. 5	Maize	35.90123 E	0.31421 N
Sample field n. 6	Maize	35.93482 E	0.31230 N
Sample field n. 7	Maize	35.93482 E	0.35304 N
Sample field n. 8	Maize	35.92540 E	0.30187 N
Sample field n. 9	Maize	35.94563 E	0.30187 N
Sample field n. 10	Maize	35.94563 E	0.34316 N
Sample field n. 11	Maize	35.92120 E	0.34316 N
Sample field n. 12	Maize	35.92120 E	0.34187 N
Sample field n. 13	Wheat	35.92187 E	0.30612 N
Sample field n. 14	Wheat	35.94212 E	0.34662 N
Sample field n. 14	Wheat	35.94662 E	0.34662 N
Sample field n. 16	Wheat	35.92171 E	0.30501 N
Sample field n. 17	Wheat	35.93141 E	0.30429 N
Sample field n. 18	Wheat	35.93141 E	0.34771 N
Sample field n. 19	Wheat	35.92514 E	0.34820 N
Sample field n. 20	Wheat	35.92514 E	0.30254 N
Sample field n. 21	Wheat	35.92401 E	0.31684 N
Sample field n. 22	Wheat	35.93611 E	0.31855 N
Sample field n. 23	Wheat	35.93611 E	0.30895 N

for the years 2015–16 were used for modeling the atmosphere over the study area. The following archives were implemented:

1) geopotential (from 1000 to 100 mbar);
2) relative humidity (from 1000 to 100 mbar);
3) temperature (from 1000 to 100 mbar);
4) horizontal wind speed (from 1000 to 100 mbar).

Top-soil (soil composition up to a depth of 20 cm) data were imported from the KENSOTER and SOTWIS databases [41]. Data on AEZ were provided by the Farm Management Handbook of Kenya (FMHK), Vol. II, Part A-B-C [39].

The zoning used reflects the methodology and subdivision criteria already adopted by FAO [42]. Phenological calendars were extracted from the FMHK and integrated with crop calendars already in use and distributed by FAO. The digital elevation model at 90-m spatial resolution from the Shuttle Radar Topography Mission (SRTM), distributed by United States Geological Survey (USGS), was implemented [43]. Testing sites' exact location within the study area and the associated crop species are reported in Table I. Crop yield observation provided by Food and Agriculture Organization Statistical service (FAOSTAT) [44] and USDA Foreign Agricultural Service (FAS) [45] was used to validate the simulation process results.

B. Landsat 8/OLI Dataset

A total of 23 geometric and terrain-corrected (L1T) scenes of Landsat 8 Operational Land Imager (OLI; Path 169/Row 60) between January 2016 (day of the year – DOY 28) and December 2016 (DOY 364) were obtained from the USGS archive. These scenes were then converted into top of atmosphere (TOA) reflectance values. Due to the relatively low-resolution image (low compared with crop field sizes of the AOI), a pan-sharpening technique was implemented to increase Landsat 8 OLI image resolution from 30 to 15 m. Pan-sharpened images have proved their effectiveness in classification or land-cover change detection problems [46], [47]. The fast intensity hue saturation (FIHS) method was applied on Landsat 8 OLI bands 2–5 as proposed by Johnson in 2014 [48], [49]. It has been demonstrated that is preferable to apply the image pan-sharpening before the VI calculation [47].

III. METHODOLOGY

This paper investigates a new phenology-based, site-specific classification approach able to discriminate different crop species using Landsat 8 OLI imagery. In particular, the phenology of the characteristic crops of the AOI and AEZ were taken from FAO sources (FAO crop calendar and FAO Farm Management Handbook). These datasets were used for implementing a multivariate decision tree (MDT). Decision tree (DT) classifiers have not been widely used by remote sensing community for the LULC classification. DT classifiers are of nonparametric nature and characterized by simplicity, flexibility, and computational efficiency [50]. An important advantage of classification trees is that they are structurally explicit, allowing for clear interpretation of the links between the dependent variable of class membership and the independent variables of remote sensing and/or ancillary data [51]. To construct a classification tree by investigative approach, it is assumed that a dataset of feature vectors and corresponding classes are available. The features

are identified and selected based on *a priori* problem-specific knowledge [52].

The DT is designed by separating the original dataset into homogenous subset on the basis of a test series applied to preliminary selected attribute values [52]. An operational example of the application of DT algorithm is the annual national-level coverage product of the United States CDL [17]. As stated by Brown de Colstoun *et al*. [53], DT classifiers can also accept a wide variety of input data, including nonremotely sensed ancillary data, in the form of both continuous and/or categorical variables. Further advantages of DTs include an ability to handle data measured on different scales, lack of any assumptions concerning the data-frequency distributions in each class, flexibility, and ability to handle nonlinear relationships between features and classes [54]. Peña-Barragán *et al*. [55] successfully combined object-based image analysis (OBIA) and DT methodology for the identification of 13 major crops cultivated in the agricultural area of Yolo County (California, USA). Their multiseasonal assessment of a large number of crop types and field status reported an overall accuracy of 79%. Brown de Colstoun *et al*. [53] used multitemporal ETM+/ Landsat-7 data and a DT classifier to map 11 types of land-cover classes, acquiring a final land-cover map with an overall accuracy of 82%. Grouping the 11 land-cover classes in forest versus nonforest classes, this same accuracy was 99.5%. Vieira *et al*. [14] combined OBIA and DT to map harvest-ready sugarcane in Brazil. The limitation of testing a single feature at a single node potentially leads to a much larger DT and could significantly reduce the ability to express and describe concepts resulting in less-representative classes. To avoid testing features more than once or to avoid replicating tests among subtrees, an MDT has been implemented. In this paper, we attempted for LULC and crop mapping of six crop species in Nakuru County, Kenya.

The AquaCrop model was then used to estimate the levels of foliar coverage, biomass, and harvest of two agricultural species (wheat and maize) in response to irrigation deficit in the subhumid and semi-arid climatic conditions of the Kenyan highlands. The FAO crop water productivity AquaCrop model simulates the obtainable biomass and the associated productivity in response to water resources availability with the specific purpose of limiting the number of required explicit parameters [56]. It is particularly suitable for application in dry, semi-arid, or drought-prone areas [57]. A crop forecasting model takes into account the plant interactions with soil (through which water and nutrients are extracted), and atmosphere (which determines the evapotranspiration, the demand for carbon dioxide, and the required energy for plant growth). Many models require a large number of input parameters and variables, not easily available for all types of crops and for all possible areas of application, especially in developing countries. Furthermore, these parameters are usually more familiar to scientists than to end users. Compared to other crop forecasting models, AquaCrop is parameterized through a reduced number of parameters, trying to balance simplicity, accuracy, and robustness.

The AquaCrop simulation model is based on the relationship between reference evapotranspiration (Et0) and harvest [58]: Evapotranspiration is divided into leaf transpiration and soil cover transpiration. The water content at roots level is modeled taking into account the water flow from and to the ground. The ground canopy cover replaces the LAI. Leaf cover expansion and senescence, stomata conductance, and production index (harvest index, HI) are all key parameters describing the response to water stress, which is finally summarized through the water productivity index that is susceptible of remote estimation through the equivalent water thickness (EWT) evaluation. It is also considered as a thermal stress factor for both high and low temperatures (pollination- and biomass-level stress) as well as the water resources management. Crop yield mapping and assessment at a local scale consists of two major steps: 1) crop mapping and 2) yield assessment at 15-m spatial resolution. Fig. 2 illustrates the main processing steps along with required input dataset. These steps are described in detail in the following subsections.

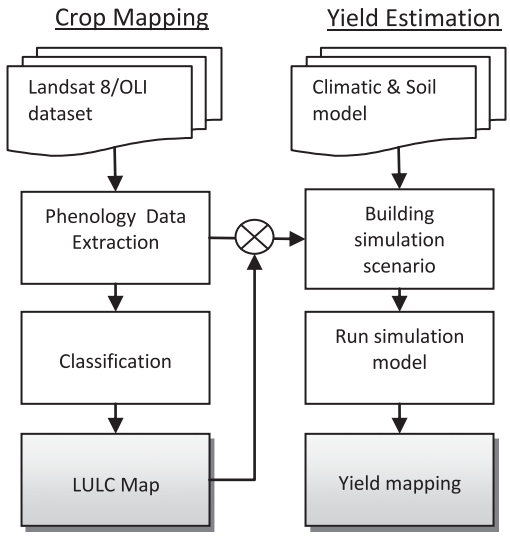

Fig. 2. Flowchart of the overall monitoring systems along with the required input dataset.

A. Crop Mapping

The developed approach for crop classification allows automatic mapping of crops at species level using a priori knowledge on crop calendar and without using ground truth data as training sites. The method is based on per-pixel estimation of phenological metrics from NDVI time series and per-pixel comparison with reference data. Tracking crops biomass development results in phenological signature identification; we applied rules on phenological variables, such as the dates of certain growing stages, to discriminate and classify crops at species level. Phenological profiles provide historical NDVI projection over a specific site and can be used to discriminate for different crop species and to identify crop rotation cycles and biomass seasonal trends [59]. Phenological stages were retrieved from crop growth time series as the main input to the classification algorithm. A pixel-based algorithm was designed to automatically retrieve vegetation phenological signature starting from NDVI time series. The NDVI values from cloud-contaminated pixels

were recovered by interpolating the available NDVI values over the closest and shortest period.

The NDVI time series were determined by interpolating the raw NDVI datasets. Spline smoothing has been achieved based on the description in [60]. This method was employed to find a spline function $S_h(t)$ that minimizes a criterion function $C_h(t)$ for a specific smoothing coefficient h ($\geq$0)

$$C_h = \sum_{i=1}^{m} \left\{ w_i [NDVI_i - S_h(t_i)]^2 + v_i p \int |S_h''(t)|^2 dt \right\} \quad (1)$$

where t_i $(i = 1, \ldots, m)$ is the regularly spaced time grid corresponding to the NDVI time series. Each point of the time series is associated with a weight $w(t_i)$, while v is the adaptative stiffness weight vector at i. The parameter h controls the spline function shape from an exact interpolation ($h = 0$) to straight line ($h \to \infty$). In order to switch from a manually selecting approach of h toward an automatic approach, generalized cross-validation was used by [60]. This procedure was improved and implemented in *GCVSPL*, a *Fortran* software package for spline smoothing and differentiation, designed by [61]; in this paper we used the MATLAB *MEX* interface for the *GCVSPL* package developed by [62].

The phenological function was then analyzed to retrieve the phenological metrics. Vegetation development is well described by the following phenological stages [63]:

1) start of season (SOS), or the onset of photosynthetic activity;
2) end of season (EOS), or the very end of the senescence period;
3) the maturity peak located at the NDVI maximum value (Max);
4) the dormancy period characterized by no photosynthetic activity and related to soil preparation or soil restoration practices.

The distance between EOS and SOS defines the length of growing period (LGP). From the NDVI-fitted function we retrieved five transition dates corresponding to the aforementioned phenological stages as illustrated in Fig. 3. Three dates ($\mathrm{DOY_{SOS}}$, $\mathrm{DOY_{EOS}}$, and $\mathrm{DOY_{Max}}$), corresponding to local minimum or maximum of the phenological function (first derivative equals to zero), describe SOS, EOS, and Max stages; two dates ($\mathrm{DOY_G}$, and $\mathrm{DOY_S}$) are determined when the increasing or decreasing rates (first derivative) are maximum and represents, respectively, the middle of the growing season (G) and of the senescence period (S). The difference between $\mathrm{DOY_{EOS}}$ and $\mathrm{DOY_{SOS}}$ stands for the LGP. The minimum NDVI value within the LGP represents the NDVI background value ($\mathrm{NDVI_b}$): The difference between $\mathrm{NDVI_{Max}}$ and $\mathrm{NDVI_b}$ represents the amplitude of the NDVI variation within the LGP.

The automated mapping task was conducted based on hierarchical rules within an MDT algorithm. Phenological variables, that represent the seasonal dynamic of NDVI and stages of phenological transition during the growing period, are useful in crop classification and discrimination tasks at species level based on crop calendar [64]. Due to the interannual and interregional variability of crop calendar, phenological variables alone may not be sufficient to fully separate crop cycles. Additional data were introduced into the classification problem based on AEZ to improve crop separability. The purpose of zoning for rural land-use planning is to separate areas with similar sets of potentials and constraints for development. AEZ, as applied in FAO studies, defines zones on the basis of combinations of soil, landform, and climatic characteristics. The Kenya AEZ study [65] distinguishes agro-ecological cells, which are the basic units for land evaluation and data processing, from agro-ecological zones, which are spatial units related to a soil map. The concept of the LGP is essential to AEZ, and provides a way of including seasonality in land resource appraisal. The determination of the beginning of the growing period is based on the start of the rainy season. By compiling an inventory of LGPs over a historical sequence of years, the frequency distribution of different annual numbers of LGP can be assessed. Kenya AEZ study identifies 22 occurring LGP patterns. In this paper, crop calendars have been imported for the AEZs of the study area, and decisional rules were designed according to local crop development characteristics.

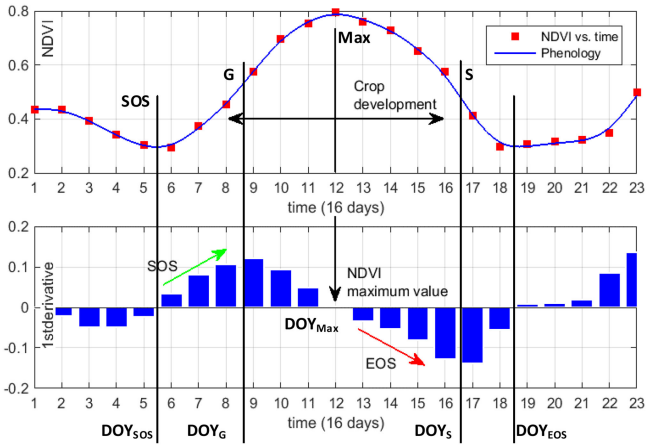

Fig. 3. NDVI time series was interpolated (above) and then analyzed (below) searching for phenological signature and transition dates representative of crop development during the season; phenological function first derivative was used to exactly locate the current transition dates of the growing stages.

The crop mapping system preclassification analysis architecture is shown in Fig. 4. The first step is searching for vegetated areas within the AOI. A threshold was settled on the time series $\mathrm{NDVI_{max}}$ value to isolate vegetated areas and then proceed with phenological metrics extraction. The second test was designed to search the phenological function for points eligible to be SOS or EOS stages. The third test was designed to catch the exact location of the LGP of the pixel under examination: This made possible the distinction between winter and summer crops. From this point on, the MDTs proceed separately for crop identification.

Crop classification MDT algorithm is illustrated in Fig. 5: Pixel geographic location selects the crop calendar to be used as reference phenology into the classification process. A series of decisional rules was designed to be exploited at each binary node of the DT. Fitted NDVI values at reference transition dates are evaluated. The purpose is to verify the matching between the

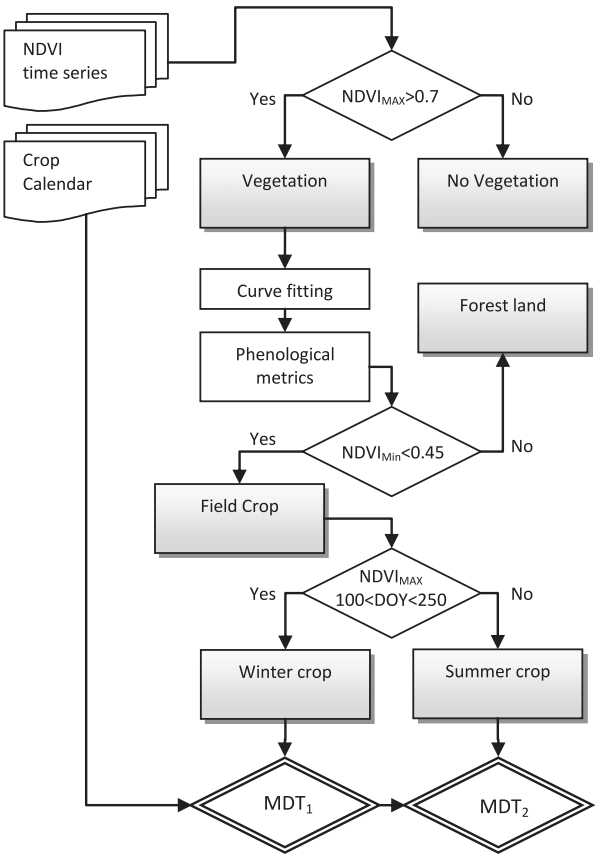

Fig. 4. Flowchart: The preclassification analysis leads to isolate winter and summer crops out of the original dataset. Winter and summer crops are going to be processed by the MDTs to carry out the classification at crop species level.

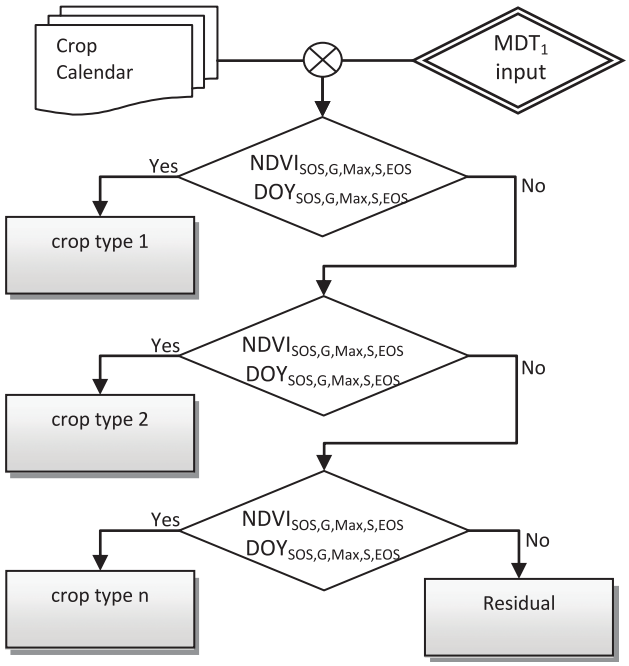

Fig. 5. Flowchart: MDT overall architecture. Transition dates pointed by the NDVI time series are compared with reference transition dates from crop calendars.

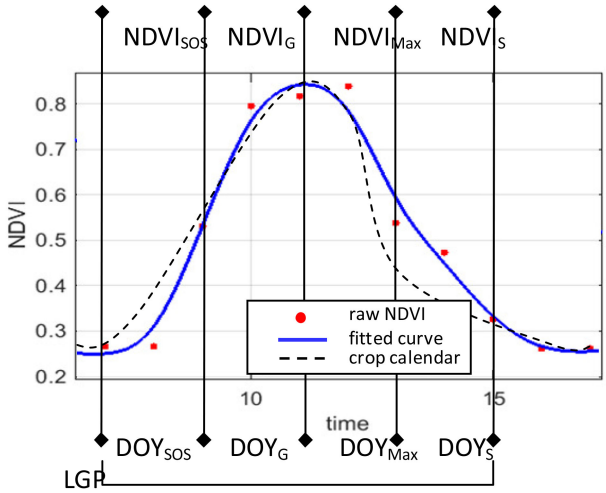

Fig. 6. NDVI values at the transition dates describing the key points of the phenological development of a crop species; the phenological curve has been retrieved from the NDVI raw values and the compared with reference phenology from FAO crop calendar.

reference transition dates and the ones retrieved from the current phenology: If the matching is verified, the pixel under analysis is consequently classified, otherwise the analysis moves on trying to verify the next possible match within the season temporal window. An example of a possible match between the reference phenology and the one retrieved from NDVI time series is shown in Fig. 6. Transition dates for the reference phenology are defined from the selected crop calendar; at the same time the analysis of pixel phenology allows us to retrieve current transition dates: the possible matching between the two sets of dates leads to crop species identification. Residual unclassified pixels can be the results of: 1) LGP mismatching or 2) unidentified crop species. In the first case, the reference LGP (based on historical analysis) is not representative of the LGP for the current season: This misalignment is mainly related with climate change. In the second case, the crop species reference phenology is not available.

B. Yield Estimation

The FAO AquaCrop model is a field-scale crop growth model that aims to simulate plant functions under specific stress conditions, within a modeled climatic/environmental scenario. The system architecture is shown in Fig. 7. Climatic and environmental data are previously processed and stored as well as crop species phenological and physiological information. The starting point is the study area selection made by the user using the interactive LULC geographic information system (GIS) product.

The climatic data (12-km spatial resolution, 19 pressure levels between 1000 and 350 mbar) and the digital terrain elevation model (SRTM) were resampled to the resolution of 250 m. Then, the climatic variables were interpolated along the atmospheric column to retrieve the variables values at the reference height of 2 m above the local ground [66].

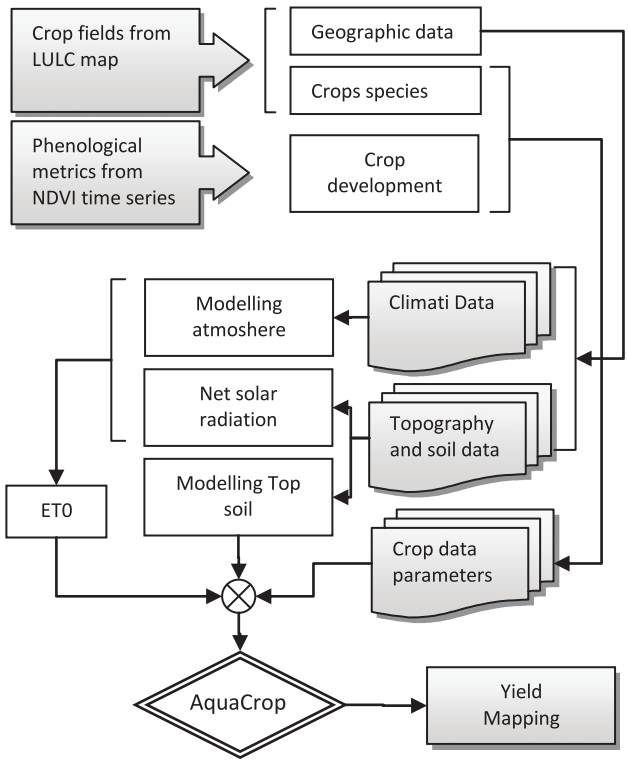

Fig. 7. Yield mapping system architecture. Crop fields and their related phenol are imported from the LULC map. Environmental data are loaded at the same time. The simulation scenario is developed and ready to run the crop productivity model.

The net solar radiation was computed using the FAO guidelines [67] and corrected for the effects of local solar illumination using topographic modeling [68].

The reference evapotranspiration (FAO-56 method) was calculated for the study areas from the available weather data using the Penman–Monteith equation [69].

From the GIS product, it is possible to select a specific crop field (or a wider area containing several fields): For each crop field, both geographic and crop specific data are processed.

Geographic data allow the system to retrieve local environmental information, while the crop species identification allow the system to load phenological information into the simulation model; critical stress factors and water requirements for the specific crop in analysis are also taken into account.

A MATLAB routine is in charge of automatically setting up all the required parameters throughout the simulation period, and to recall the AquaCrop routine as many time as necessary. The simulation was conducted using daily data over the simulation period. Simulations result were then validated using collected ground truth data.

The climatic data (daily average reference values at 2 m above the ground surface) used over the simulation period as result of climate modeling are shown in Table II.

The top soil characteristics over the testing sites (average values among the samples) are reported in Table III.

TABLE II
CLIMATIC DATA MONTHLY AVERAGE OVER THE SIMULATION PERIOD

Month	Rain (mm)	Et0 (mm)	Tmin °C	Tmax °C	CO2 ppm
March	3.0	3.03	14.3	16.7	400.55
April	1.2	3.31	14.6	16.1	400.55
May	1.7	2.93	14.6	16.4	400.55
June	0.8	2.89	13.2	15.2	400.55
July	0.8	2.75	13.1	14.9	400.55
August	1.5	2.93	12.4	20.0	400.55
September	5.2	3.02	13.4	14.6	400.55

TABLE III
TOP SOIL PROFILE

Depth m	Sat %	FC %	WP %	Ksat mm/day	CRa	CRb
0.14	42.1	32.7	21.9	50.4	-0.569	-1.397
0.14	45.1	39.4	27.5	11.0	-0.568	-2.298
0.30	45.2	40.5	28.7	6.2	-0.567	-2.638
0.22	45.5	35.8	18.3	64.3	-0.492	-0.142
0.20	39.3	21.3	13.4	129.5	-0.312	-0.318

Sat: saturation, FC: field capacity, WP: wilting point, K_{sat}: saturated hydraulic conductivity, CRa, CRb: 'a' and 'b' parameters for estimating capillary rise.

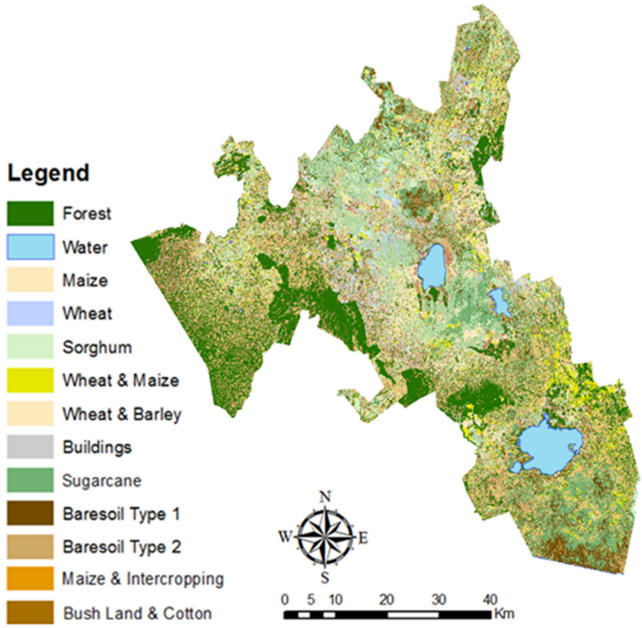

Fig. 8. LULC map resulting from the classification process, Nakuru County, Kenya, 2016.

IV. RESULTS AND DISCUSSION

A. Crop Mapping

The multitemporal phenology-based approach to the classification problem has been applied to Nakuru County. Classification results are shown in Fig. 8. Ground truth data have been collected during field campaign in 2016 and used to validate

TABLE IV
CONFUSION MATRIX

	a	b	c	d	e	f	g	h	i	l	m	n	o	PA
a	2449	70	0	13	0	10	0	0	0	0	0	0	0	96.34%
b	112	977	0	1	0	0	0	0	0	0	0	146	44	76.32%
c	0	7	9704	26	0	0	0	0	0	0	91	57	0	98.16%
d	0	0	0	56	0	0	0	0	0	0	0	0	0	100.00%
e	0	0	0	0	84359	0	0	0	0	0	0	0	0	100.00%
f	119	23	0	0	0	1123	0	0	0	5	0	0	0	88.42%
g	0	20	0	0	0	0	9	0	0	0	0	0	0	31.03%
h	0	0	0	0	0	0	0	611	0	0	0	0	0	100.00%
i	0	0	0	0	0	0	0	0	298	0	0	0	0	100.00%
l	28	19	0	0	0	9	0	0	0	768	0	0	0	93.20%
m	0	17	0	0	0	0	34	0	0	0	2347	185	9	90.54%
n	0	110	0	0	0	0	0	0	0	0	40	2996	17	94.72%
o	0	10	0	0	0	0	37	0	0	0	0	0	578	92.48%
UA	90.43 %	77.92 %	100 %	58.33 %	100 %	98.33 %	11.25 %	100 %	100 %	99.35 %	94.71 %	88.53 %	89.19 %	
														Overall Accuracy: 98.82 %

a: sugarcane, b: maize, c: forest, d: buildings, e: water, f: bush land & cotton, g: sorghum, h: bare soil type1, i: bare soil type2, l: maize and intercropping, m: biseasonal wheat and maize, n: biseasonal wheat and barley, o: wheat. PA: producer accuracy, UA: user accuracy.

TABLE V
CONFUSION MATRIX

	Sugarcane	Maize	Bush land & Cotton	Sorghum	Maize & Intercropping	Wheat & Maize	Wheat & Barley	Wheat	PA
Sugarcane	2449	70	10	0	0	0	0	0	96.83%
Maize	112	977	0	0	0	0	146	44	76.38%
Bush land & Cotton	119	23	1123	0	5	0	0	0	88.42%
Sorghum	0	20	0	9	0	0	0	0	31.03%
Maize & Intercropping	28	19	9	0	768	0	0	0	93.20%
Wheat & Maize	0	17	0	34	0	2347	185	9	90.54%
Wheat & barley	0	110	0	0	0	40	2996	17	94.72%
Wheat	0	10	0	37	0	0	0	578	92.48%
UA	90.43%	78.41%	98.33%	11.25%	99.35%	98.32%	90.05%	89.19%	
									Overall accuracy: 91.35 %

PA: producer accuracy, UA: user accuracy.

the developed LULC map. The first confusion matrix, concerning the assessment of the overall accuracy of the classification procedure of all the land-cover classes, is shown in Table IV. The significantly high score of the overall accuracy (98.82%) is mainly related to the effectiveness of the phenology-based approach, in classifying tropical forests, bare soils, buildings, and water basins (classes *c*, *h*, *i*, *d*, *e*).

The confusion matrix presented in Table V is representative of the validation process carried out exclusively with agricultural classes. Classes *m*, *n,* and *f* are intended to be representative of a biseasonal (first and second season) field exploitation: The crop rotation cycle for those classes is consistent with crop calendar. Class *l* is an experimental class intended to be representative of the widely used technique of intercropping maize with beans; we take advantage of the residual phenology after the harvest of maize to discriminate this class from classes *b* and *m*.

The overall accuracy score of 91.35%, with a significant drop down in accuracy related with sorghum classification that is often mistaken with maize. The classification failure concerning sorghum can be partially explained as a consequence of an insufficient validation due to the lack of ground truth data for this species. The classification error reduction leads to a significant improvement of the classification performances. The limited number of validation sites could result in overestimating the efficiency of our method. Nonetheless the algorithm has proved to be effective in discriminating different crop types. Moreover, phenological information has proved to be little affected by cloud contamination over period and the classification algorithm was able to properly classify also pixels affected by a significant cloud contamination during the year of interest.

B. Yield Estimation

The simulation model evaluates the expected crop yield starting from the biomass total amount estimation; in doing this, it takes into account several possible environmental stress factors as well as water availability, cultivation-management strategies, and possible supporting activities. Each simulation must be calibrated according to crop species, climatic data, and soil type. Test sites for maize and wheat crops have been previously selected on the ground. Their exact location is reported in Table I. Fig. 9 shows the sample sites localization using Google Earth imagery and the LULC map previously developed.

Fig. 9. Sample sites from Table I localization from Google Earth imagery (on the left) and from the LULC map previously developed (on the right).

TABLE VI
SIMULATION RESULTS

	s.f. 1	s.f. 2	s.f. 3	s.f. 4	s.f. 5	s.f. 6	s.f. 7	s.f. 8	s.f. 9	s.f.10	s.f. 11	s.f.12	s.f.13	s.f.14	s.f.15	s.f.16	s.f.17	s.f.18	s.f.19	s.f.20	s.f.21	s.f.22	s.f.23
Year	2016	2016	2016	2016	2016	2016	2016	2016	2016	2016	2016	2016	2016	2016	2016	2016	2016	2016	2016	2016	2016	2016	2016
Irri	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Infilt	190	190	176	201	190	174	190	190	190	201	201	201	333	338	368	338	333	338	338	355	320	360	355
E	166	164	164	174	164	170	166	164	166	164	164	164	255	256	253	257	255	257	257	255	259	253	255
Tr	125	126	109	126	126	109	126	125	125	126	126	126	155	155	143	155	155	155	155	143	155	143	143
TrW	125	126	109	126	126	109	126	125	125	126	126	126	155	155	143	155	155	155	155	143	155	143	143
Cycle	93	94	93	92	92	93	93	95	91	94	94	93	112	112	113	112	113	112	112	113	112	113	113
TempStr	21	21	24	22	20	24	21	20	20	22	22	22	0	0	0	0	0	0	0	0	0	0	0
ExpStr	33	32	26	33	33	26	33	33	32	32	32	32	32	32	30	33	32	33	33	30	34	30	32
StoStr	20	20	21	18	20	20	20	19	21	20	18	18	34	34	39	31	34	33	33	34	29	34	34
BioMass	9.86	10.11	7.67	9.37	10.10	7.53	8.41	10.27	10.44	8.28	8.28	8.32	7.87	7.94	7.53	8.34	7.80	8.10	8.10	7.73	8.96	7.67	7.75
Brelative	52	54	45	50	54	46	50	55	52	51	51	54	31	32	30	31	32	32	30	33	31	30	30
HI	17.5	18.4	17.6	18.1	18.4	17.4	17.6	18.1	17.8	18.1	18.1	17.9	27.3	27.5	22.8	27.7	27.1	23.4	23.4	27.3	27.9	27.5	23.1
Yield	1.72	1.86	1.35	1.69	1.86	1.31	1.48	1.86	1.86	1.50	1.50	1.49	2.15	2.17	1.71	2.31	2.11	1.89	1.89	2.11	2.50	2.11	1.79
WPet	0.72	0.78	0.63	0.67	0.78	0.60	0.64	0.78	0.78	0.65	0.65	0.65	0.68	0.69	0.60	0.71	0.66	0.64	0.64	0.66	0.72	0.66	0.62

Year, Start year of simulation run; **Irri**, Water applied by irrigation (mm); **Infilt**, Infiltrated water in soil profile (mm); **E**, Soil evaporation (mm); **Tr**, Total transpiration of crop (mm); **TrW**, Crop transpiration by weed infestation (mm); **Cycle**, Length of crop cycle (day); **TempStr**, Temperature stress (%); **ExpStr**, Leaf expansion stress (%); **StoStr**, Stomatal stress (%); **Biomass**, Cumulative biomass produced (ton/ha); **Brelative** Relative biomass (%); **HI**, Harvest Index adjusted for inadequate photosynthesis and water stress (%); **Yield**, Yield (ton/ha); **WPet**, ET Water Productivity for yield part (kg yield produced per m3 water evapotranspired); **s.f.**, sample field.

The simulations results are shown in Table VI: Stress factors such as soil salinization and soil fertility deficit have not been taken into account; on the other hand, thermal stress factors as well as high suffering conditions of the vegetation (mostly localized at foliar and stomata level) are decisive in determining the HI drastic reduction, from an expected value close to 48%, down to an actual value of 18%–23%.

Observed data over the study area have provided the necessary reference to make a proper validation of the simulation process. The difference between estimation and observation of corn and wheat crops for the study areas (Fig. 10) is satisfactory, with R values of 0.69 and 0.72 and MAE equal to 119 kg/ha and 491 kg/ha, respectively, for the two species. Aquacrop tends to overestimate the biomass production in response to elevated CO_2 values (>360 ppm) [70], this explains the wheat productivity overestimation. Although the general tendency of AquaCrop model to overestimate maize production under water stress condition has been reported by [71], under more severe conditions

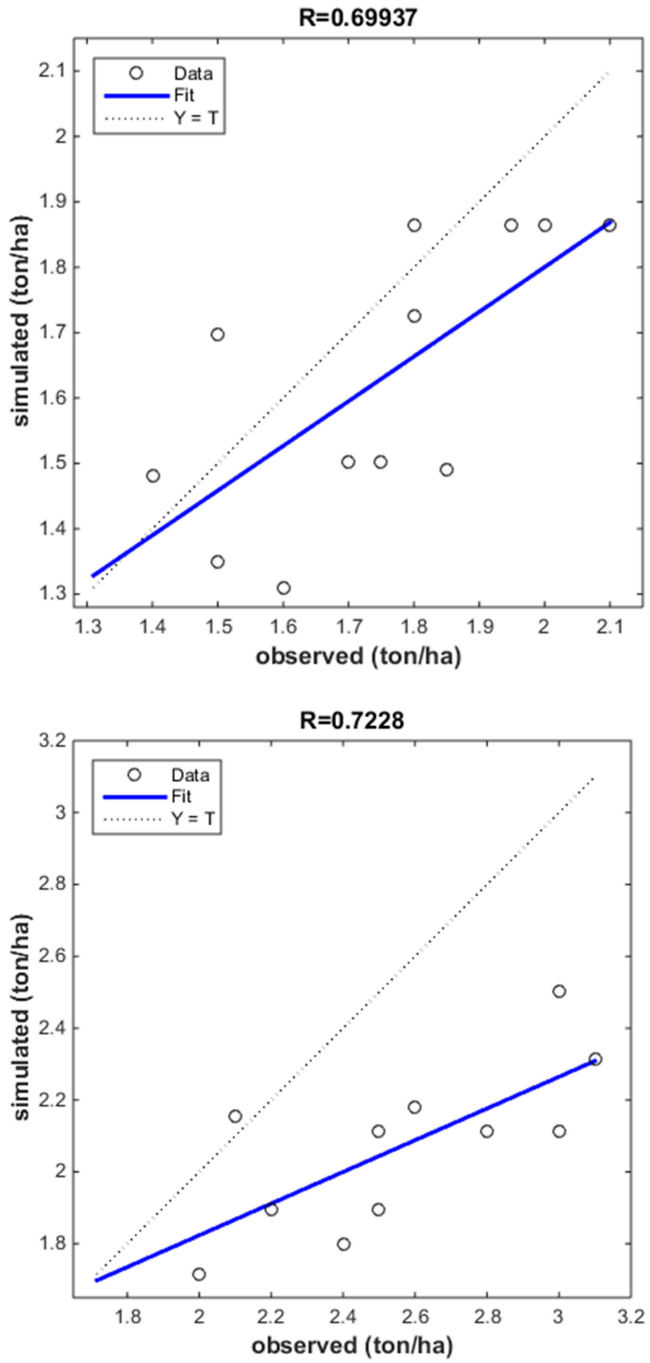

Fig. 10. Comparison between simulated and observed data for maize (above) and wheat (below).

(below the 33% of the expected full irrigation) the model underestimates the biomass through the whole season [72].

Model limitations are linked to the accuracy of the input data. A model requires information about soil composition, weather, and management practices, but often those data are not available or accessible. When this state is not met, this causes uncertainties in the results. Uncertainties are also related with model calibration. Users of this model should also refer to the condition in which the model was developed and tested. The use of remote sensing data can, at this point, make up for the lack of spatial information and select the correct temporal window in which to perform the simulations according to crop species.

V. Conclusion

This paper attempted to explore the combined use of phenological information, DT algorithm, and simulation model to crop mapping and yield assessment at regional level. A 15-m spatial resolution NDVI time series was developed in order to retrieve vegetation phenological signatures and metrics. Phenological data were used both by the classification process and by the simulation model.

Phenological data proved to be quite effective in discriminating crops at species level and in supporting crop simulation model by adding reliable information to correct and assess the reference LGP. The difference in phenology between crop species sometimes has less magnitude than interannual or regional variability: Strict rules related with growing stage dates, applied on phenological variables, result in high accuracy classification for a specific year or region but could be ineffective for another year or region. Toward the development of an automatic classifier, universal rules need to be established by the processing of a large amount of data instead of training different classifiers for individual years [64], [73]. Furthermore, the image acquisition time is more important than the overall number of available images. Even with a large number of available images, phenological stages of relevance can be poorly described due to missing data: This is eventually due to a discrepancy between the image acquisition scheduling and the temporal window in which the phenological stage could be catch and properly described. This is most likely for phenological stages (and the associated decision rules) that last only for a short time. The mapping system offers room for improvement, although higher complexity may lead to lower the transferability. Furthermore, characteristic crop sequences of previous years can give information about the probability of following crop types, although the actual crop rotation could differ from the expected one based on expert knowledge.

The yield mapping system relies upon a dedicated crop mapping service devoted to agricultural areas discrimination, identification, and representation. A specific study area can be selected and environmental as well as climatic data are loaded into the simulation process. Due to the simulation model calibration requirements, the monitoring system is really site specific and a previous and accurate knowledge related to crop characteristics, typical phenological development, and stress factors is required. The AquaCrop simulation model can be a valuable tool in the planning and management of agricultural activities; the model requires a limited number of parameters, largely intuitive and explicit, easily collectible, even according to automatic procedures, in many cases already available [74], [75]. During the validation procedure, it was observed that model performances depends, as expected, on the water stress levels experienced by the crops during the growing cycle. Model performances could become unsatisfactory in severely stressed environment. Observations made about a general season under/overestimation are

quite similar to the ones reported by [76]–[79]. The model performance requires a broader validation according to the different possible calibrations over a wider availability of sample sites and agricultural species. When properly calibrated, AquaCrop reliability implies that it can be used in developing new and site-specific strategies for field and water resources management techniques improvement.

The main advantages from incorporating remote sensing data into crop model are the better representation of crop fields spatial information (very important for several applications including site-specific agriculture and food security) and the more accurate description of crop actual development along the various stages of the growing season. A future development of the paper foresees the use of satellite-based estimation of Et0, LAI, EWT (related to water productivity), and cloud cover (to correct for the real isolation period). Considering the advantages and the wide applicability of the combination of crop models with remote sensing data and products, the collaboration of the two will increase especially by the means of using of automatic procedure.

REFERENCES

[1] Federal Research Division, *Kenya Country Profile.* Washington, DC, USA: United States Library of Congress, 2007.

[2] Mwanda, C.O. Engineering Division, Ministry of Agriculture. "A note on weed control in Machakos District, Kenya," 2000, Accessed: May 5, 2008. [Online]. Available: http://atnesa.org.

[3] A. Di Gregorio, and L. J. M. Jansen, "Part I - Technical document on the Africover land cover classification scheme," in *FAO Africover Land Cover Classification*, 1997, pp. 4–33; 63-76, *FAO Remote Sensing Centre series, N. 70*, Roma. W.-K. Chen, *Linear Networks and Systems*. Belmont, CA, USA: Wadsworth, 1993, pp. 123–135.

[4] O. Dubovyk, G. Menz, C. Conrad, J. Lamers, A. Lee, and A. Khamzina., Spatial targeting of land rehabilitation: A relational analysis of cropland productivity decline in arid Uzbekistan. *Erdkunde*, vol. 67, no. 2, pp. 167–181, 2013, doi: 10.3112/erdkunde.2013.02.05.

[5] C. Justice and P. Hiernaux, "Monitoring the grasslands of the Sahel using NOAA AVHRR data: Niger 1983," *Int. J. Remote Sens.*, vol. 7, pp. 1475–1498, 1986.

[6] C. Justice and J. R. G. Townshend, "Analysis of the phenology of global vegetation using meteorological satellite data," *Int. J. Remote Sens.*, vol. 6, pp. 1271–1318, 1985.

[7] C. Justice, B. N. Holben, and M. D. Gwyne, "Monitoring east African vegetation using AVHRR data," *Int. J. Remote Sens.*, vol. 7, pp. 1453–1474, 1986.

[8] A. Vrieling, K. M. de Beurs and M. E. Brown, "Variability of African farming systems from phenological analysis of NDVI time series," *Climatic Change*, vol. 109, no. 3–4, pp. 455–477, 2011.

[9] J. I. Fisher and J. F. Mustard, "Cross-scalar satellite phenology from ground, Landsat and MODIS data," *Remote Sens. Environ.*, vol. 109, pp. 261–273, 2007.

[10] G. Hmimina *et al.*, "Evaluation of the potential of MODIS satellite data to predict vegetation phenology in different biomes: An investigation using ground-based NDVI measurements," *Remote Sens. Environ.*, vol. 132, pp. 145–158, 2013.

[11] P. Teluguntla *et al.*, "Spectral matching techniques (SMTs) and automated cropland classification algorithms (ACCAs) for mapping croplands of Australia using MODIS 250-m time-series (2000–2015) data," *Int. J. Digital Earth*, vol. 10, no. 9, pp. 944–977, 2017.

[12] C. Justice *et al.*, "The moderate resolution imaging spectroradiometer (MODIS): Land remote sensing for global change research," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, no. 4, pp. 1228–1249, Jul. 1998.

[13] P. S. Thenkabail, A. D. Ward, J. G. Lyon, and C. J. Maerry, "Thematic mapper vegetation indices for determining soybean and corn growth parameters," *Photogrammetric Eng. Remote Sens.*, vol. 60, pp. 437–442, 1994.

[14] M. A. Vieira, A. R. Formaggio, C. D. Rennó, C. Atzberger, D. A. Aguiar, and M. P. Mello, "Object based image analysis and data mining applied to a remotely sensed Landsat time-series to map sugarcane over large areas," *Remote Sens. Environ.*, vol. 123, pp. 553–562, 2012.

[15] F. Gao *et al.*, "Fusing Landsat and MODIS data for vegetation monitoring," *IEEE Geosci. Remote Sens. Magaz.*, vol. 3, no. 3, pp. 47–60, Sep. 2015.

[16] H. E. Beck, T. R. McVicar, A. I. J. M. van Dijk, J. Schellekens, R. A. M. de Jeu, and L.A. Bruijnzeel, "Global evaluation of four AVHRR–NDVI data sets: Inter-comparison and assessment against Landsat imagery," *Remote Sens. Environ.*, vol. 115, no. 10, pp. 2547–2563, 2011.

[17] C. Boryan, Z. Yang, R. Mueller, and M. Craig, "Monitoring US Agriculture: The US department of agriculture, national agricultural statistics service, Cropland Data Layer Program," *Geocarto Int.*, vol. 26, no. 5, pp. 341–358, 2011.

[18] Y. Bai, M. Feng, H. Jiang, J. Wang, Y. Zhu, and Y. Liu, "Assessing consistency of five global land cover data sets in China," *Remote Sens.*, vol. 6, pp. 8739–8759, 2014.

[19] R. Luciani, G. Laneve, M. Jahjah, and C. Mito, "A multi-temporal phenology based classification approach for crop monitoring in the framework of the SBAM project," *South African J. Geomatics*, to be published.

[20] B. L. Turner, E. F. Lambin, and A. Reenberg, "The emergence of land change science for global environmental change and sustainability," in *Proc. Nat. Acad. Sci. USA*, 2007, vol. 104, pp. 20666–20671.

[21] T. R. Loveland, *et al.*, "Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data," *Int. J. Remote Sens.*, vol. 21, pp. 1303–1330, 2010.

[22] M. Friedl *et al.*, "Global land cover mapping from MODIS: Algorithms and early results," *Remote Sens. Environ.*, 83, pp. 287–302, 2002.

[23] P. Bicheron *et al.*, "GLOBCOVER: Products Description and Validation Report," MEDIAS, Toulouse, France, Tech. Rep., 2008.

[24] M. C. Hansen *et al.*, "High-resolution global maps of 21st-century forest cover change," *Science*, vol. 342, pp. 850–853, 2013.

[25] N. Joshi *et al.*, "A review of the application of optical and radar remote sensing data fusion to land use mapping and monitoring," *Remote Sens.*, vol. 8, pp. 70–92, 2016.

[26] N. Joshi *et al.*, "Mapping dynamics of deforestation and forest degradation in tropical forests using radar satellite data," *Environ. Res. Lett.*, vol. 10, p. 034014, 2015.

[27] C. M. Ryan, N. J. Berry, and N. Joshi, "Quantifying the causes of deforestation and degradation and creating transparent REDD+ baselines: A method and case study from central Mozambique," *Appl. Geogr.*, vol. 53, pp. 45–54, 2014.

[28] M. Brolly and I. H. Woodhouse, "A "matchstick" model of microwave backscatter from a forest," *Ecol. Model.*, vol. 237–238, pp. 74–87, 2012.

[29] C. Haub, K. Quilitz, D. Lindemann, and E. Leskinen, "Surveying European landscape dynamics," *J. Photogrammetry, Remote Sens. Geoinf. Sci.*, vol. 5, pp. 473–482, 2013.

[30] K. R. McCloy, "Development and evaluation of phenological change indices derived from time series of image data," *Remote Sens.*, vol. 2, pp. 2442–2473, 2010.

[31] B. C. Reed, "Trend analysis of time-series phenology of North America derived from satellite data," *GISci. Remote Sens.*, vol. 43, no. 1, pp. 24–38, 2006.

[32] J. Tang *et al.*, "Emerging opportunities and challenges in phenology: A review," *Ecosphere*, vol. 7, no. 8, 2016, Art. no. e0436.

[33] C. Wang, J. Li, Q. Liu, B. Zhong, S. Wu, and C. Xia, "Analysis of differences in phenology extracted from the enhanced vegetation index and the leaf area index," *Sensors*, vol. 17, no. 9, pp. 1982–1995, 2017.

[34] J. Xue and B. Su, "Significant remote sensing vegetation indices: A review of development and applications," *J. Sensors*, 2017, 2017, Art. no. 1353691,.

[35] S. Twomlow *et al.*, "Building adaptive capacity to cope with increasing vulnerability due to climatic change in Africa—A new approach," *Phys. Chem. Earth Part B*, vol. 33, pp. 780–787, 2008.

[36] S. R. Evett and J. A. Tolk, "Introduction: Can water use efficiency be modeled," *Agron. J.*, vol. 101, pp. 423–425, 2009.

[37] T. R. Sinclair and N. G. Seligman, "Crop modeling: From infancy to maturity," *Agron. J.*, vol. 88, pp. 698–703, 1996.

[38] B. K. Bhattacharya and P. S. Sastry, "Comparative evaluation of three-crop growth models for the simulation of soil water balance in oilseed Brassica," *Agr. Water Manage.*, vol. 42, pp. 29–46, 1999.

[39] R. Jaetzold and H. Schmidt, *Farm Management Handbook of Kenya*," vol. II/A-B, Ministry of agriculture, Nairobi, Kenya, in cooperation with the German Agricultural Team (GAT), German Agency for Technical Cooperation (GTZ), 1983.

[40] ERA Interim Public Datasets, *European Centre for Medium-Range Weather Forecast*. [Online]. Available: https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era-interim
[41] N. H. Batjes, "ISRIC-WISE global data set of derived soil properties on a 0.5 by 0.5 degree grid (Version 2.0)." International Soil Reference and Information Centre (ISRIC), Wageningen, 2002. [Online]. Available: http:\\www.isric.org.
[42] FAO. "Report on agro-ecological zones Project. Vol 1. Methodology and results for Africa." FAO, Rome, Tech. Rep., World Soil Resources Report 48, 1978.
[43] US Geological Survey 2004, *Shuttle Radar Topography Mission, 1 Arc Second scene SRTM_u03_n008e004, Unfilled Unfinished 2.0*, Global Land Cover Facility, University of Maryland, College Park, MD, USA, 2000.
[44] FAO, *FAOSTAT, Food and Agriculture Organization of the United Nations*, FAO, Rome, Italy , 2016.
[45] USDA Foreign Agricultural Service, USDA foreign agricultural service market and trade data, Washington, DC, USA, 2018. [Online]. Available: https://data.nal.usda.gov/dataset/usda-foreign-agricultural-service-market-and-trade-data
[46] F. Palsson, J. R. Svenisson, J. A. Benediktsson, and H. Aanaes, "Classification of pansharpened urban satellite images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 1, pp. 281–297, Feb. 2012.
[47] B. A. Johnson, H. Scheyvens, and B. R. Shivakoti, "An ensemble pansharpening approach for finer-scale mapping of sugarcane with Landsat 8 imagery," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 33, pp. 218–225, 2014.
[48] B. Johnson, "Effects of pansharpening on vegetation indices," *ISPRS Int. J. Geo-Inf.*, vol. 3, pp. 507–522, 2014.
[49] G. Laneve, R. Luciani, and M. Jahjah, "Developing a classification method for periodically updating agricultural maps in Kenya," in *Proc. Int. Geosci. Remote Sens. Symp.,* Beijing, Jul. 10–15, 2016, pp. 3543–3546, doi: 10.1109/IGARSS.2016.7729917.
[50] M. A. Friedl and C. Brodley, "Decision tree classification of land cover from remotely sensed data," *Remote Sens. Environ.*, vol. 61, pp. 399–409, 1997, doi: 10.1016/S0034-4257(97)00049-7.
[51] R. L. Lawrence and A. Wright, "Rule-based classification systems using classification and regression tree (CART) analysis," *Photogrammetric Eng. Remote Sens.*, vol. 67, no. 10, pp. 1137–1142, 2001.
[52] M. K. Ghose, R. Pradhan, and S. S. Ghose, "Decision tree classification of remotely sensed satellite data using spectral separability matrix," *Int. J. Adv. Comput. Sci. Appl.*, vol. 1, no. 5, 2010, doi: 10.14569/IJACSA.2010.010516.
[53] E. C. Brown de Colstoun, M. H. Story, C. Thompson, K. Commisso, T. G. Smith, and J. R. Irons, "National park vegetation mapping using multitemporal Landsat 7 data and a decision tree Classifier," *Remote Sens. Environ.*, vol. 85, pp. 316–327, 2003.
[54] M. A. Friedl and C. Brodley, "Decision tree classification of land cover from remotely sensed data," *Remote Sens. Environ.*, vol. 61, pp. 399–409, 1997.
[55] J. M. Peña-Barragán, M. Ngugi, R. E. Plant, and J. Six, "Object-based crop identification using multiple vegetation indices, textural features and crop phenology," *Remote Sens. Environ.*, vol. 115, pp. 1301–1316, 2011.
[56] P. Steduto, T. C. Hsiao, D. Raes, and E. Fereres, "AquaCrop-The FAO crop model for predicting yield response to water: I. Concepts and underlying principles," *Agron. J.*, vol. 101, pp. 426–437, 2009.
[57] T. C. Hsiao, L. K. Heng, P. Steduto, B. Rojas-Lara, D. Raes, and E. Fereres, "AquaCrop-The FAO crop model to simulate yield response to water: III. Parameterization and testing for maize," *Agron. J.*, vol. 101, pp. 448–459, 2009.
[58] J. Doorenbos and A. H Kassam, "Yield response to water," FAO, Rome, Italy, Tech. Rep. FAO Irrigation and Drainage Paper no. 33, pp. 193, 1979.
[59] T. Suepa, J. Qi, S. Lawawirojwong, and J. P. Messina, "Understanding spatio-temporal variation of vegetation phenology and rainfall seasonality in the monsoon Southeast Asia," *Environmental Res.*, vol. 147, pp. 621–629, 2016.
[60] P. Craven and G. Wahba, "Smoothing noisy data with spline functions," *Numer. Math.*, vol. 31, pp. 377–403, 1979.
[61] H. J. Woltring, "A Fortran package for generalized, cross-validatory spline smoothing and differentiation," *Adv. Eng. Softw.*, vol. 8, no. 2, pp. 104–113, 1986.
[62] G. Anthony Reina, "Woltring's generalized cross-variance (GCV) natural B-spline filter," *A Matlab MEX interface for the GCVSPL and SPLDER packages based on Dwight Meglan's C Code*. San Diego, CA: The Neurosciences Institute, 1998.
[63] T. Sakamoto, M. Yokozawa, H. Toritani, M. Shibayama, N. Ishitsuka, and H. Ohno, "A crop phenology detection method using time-series MODIS data," *Remote Sens. Environ.*, vol. 96, pp. 366–374, 2005.
[64] L. Zhong, L. Hu, L. Yu, P. Gong, and G. S. Biging, "Automated mapping of soybean and corn using phenology," *ISPRS J. Photogramm. Remote Sens.*, vol. 119, pp. 151–164, 2016, doi: 10.1016/j.isprsjprs.2016.05.014.
[65] FAO, "Agro-ecological Assessment for National Planning: The Example of Kenya," FAO, Rome, Italy, 1993.
[66] L. Zotarelli, M. D. Dukes, C. C. Romero, K. W. Migliaccio, and K. T. Morgan, "Step by step calculation of the Penman-Monteith evapotranspiration (FAO-56 Method)," Agricultural and Biological Engineering Department, UF/IFAS Extension, Gainesville, FL, USA, Tech. Rep. AE459, 2015.
[67] R. G. Allen, L. S. Pereira, D. Raes, and M. Smith, "Crop evapotranspiration—Guidelines for computing crop water requirements—FAO Irrigation and drainage paper 56," FAO, Rome, Italy, 1998.
[68] M. J. Canty, *Image analysis, Classification and Change Detection in Remote Sensing.* Boca Raton, FL, USA: CRC Press, 2006.
[69] T. A. Howell and S. R. Evett, "The Penman-Monteith method," Section 3 in Evapotranspiration: Determination of Consumptive Use in Water Rights Proceedings, Continuing Legal Education in Colorado, Inc. Denver, CO, 2004.
[70] E. Vanuytrecht, D. Raes, and P. Willems, "Considering sink strength to model crop production under elevated atmospheric CO_2," *Agr. Forest Meteorol*. vol. 151, 1753–1762, 2011.
[71] L. K. Heng, T. Hsiao, S. Evett, T. Howell, and P. Steduto, "Validating the FAO AquaCrop model for irrigated and water deficient field maize," *Agron. J.*, vol. 101, pp. 488–498, 2009.
[72] G. E. Greaves and Y.-M. Wang, "Assessment of FAO aquacrop model for simulating maize growth and productivity under deficit irrigation in a tropical environment," *Water*, vol. 8, p. 557, 2016.
[73] A. Baraldi, "Satellite image automatic MapperTM (SIAMTM)—A Turnkey software executable for automatic near real-time multi-sensor multi-resolution spectral rule-based preliminary classification of spaceborne multi-spectral images," *Recent Pat. Space Technol.*, vol. 1, pp. 81–106, 2011.
[74] M. S. Mkhabela and P. R. Bullock, "Performance of the fao aquacrop model for wheat grain yield and soil moisture simulation in western Canada," *Agr. Water Manage.*, vol. 110, pp. 16–24, 2012.
[75] D. A. Kasampalis, T. K. Alexandridis, C. Deva, A. Challinor, D. Moshou, and G. Zalidis, "Contribution of remote sensing on crop models: A review," *J. Imaging*, vol. 4, p. 52, 2018.
[76] N. Katerji, P. Campi, and M Mastrorilli, "Productivity, evapotranspiration, and water use efficiency of corn and tomato crops simulated by AquaCrop under contrasting water stress conditions in the Mediterranean region," *Agr. Water Manage.*, vol. 130, pp. 14–26, 2013.
[77] X.-L. Jin *et al.*, "Assessment of the AquaCrop model for use in simulation of irrigated winter wheat canopy cover, biomass, and grain yield in the North China plain," *PLoS One*, vol. 9, 2014, Art. no. e86938.
[78] A. Araya, S. Habtub, K. M. Hadguc, A. Kebedea, and T. Dejene, "Test of AquaCrop model in simulating biomass and yield of water deficient and irrigated barley (*Hordeum vulgare*)," *Agr. Water Manage.*, vol. 97:pp. 1838–1846, 2010a.
[79] G. Ali, M. Khalifa, M. Lyes, and M. Hafsi, 2017. "Evaluation of the FAO Aqua-Crop model for Durum wheat (*Triticum durum Desf.*) on the eastern Algeria under semi-arid conditions," *Indian J. Agricultural Res.*51, doi: 10.18805/ijare.v51i04.8430.

Roberto Luciani was born in Rome, Italy, in 1984. He received the B.A.Sc. degree in aerospace engineering and the M.A.Sc. degree in astronautical engineering from Sapienza University, Rome, Italy, and the Ph.D. degree in energy and environment from the Department of Astronautical, Electric and Energetic Engineer, Sapienza University, in 2018.

Starting from his Master's thesis, his research involves applying remote sensing, image processing, and spatial analyses to assess the impacts of natural and anthropogenic stressors on the marine and terrestrial ecosystems.

Giovanni Laneve (M'02) was born in Viggiano, Italy, in 1960. He received the Laurea degree in aeronautical engineering from the Università di Napoli, Naples, Italy, in 1985, and the Laurea degree in aerospace engineering from the Università di Roma "La Sapienza," Rome, Italy, in 1988.

From 1987 to 1991, he was a consultant with the Centro di Ricerca Progetto San Marco, Rome, Italy, where he was involved in the San Marco 5 satellite mission control and data analysis. Since 1991, he has been an Assistant Professor with the School of Aerospace Engineering, Università di Roma. "La Sapienza." He has authored more than 160 scientific papers. Since 1988, he has been teaching the course of "Aerospace Systems for Remote Sensing" with the School of Aerospace Engineering, Università di Roma, "La Sapienza." His research interests include aeronomy, satellite thermal control, mission design, new algorithms for the exploitation of satellite images, satellite remote sensing applications for fire management, applications of satellite data for the African regions, and studies on environmental and disaster monitoring.

Munzer Jahjah was born in Syria in 1965. He received the B.Sc. degree in electrical engineering in 1989, two years Post Lauream Specialization School in town-planning techniques for metropolitan areas in 1999, and the Ph.D. degree in aerospace engineering from the University of Rome "La Sapienza," Rome, Italy, in 2003.

Between 1991 and 1996, he worked in applied sciences with the Data Processing and Production Remote Sensing Department, Syria. Between 2003 and 2009, he was as a consultant with the San Marco Project Research Centre (CRPSM), Rome, Italy. Between 2009 and 2013, he was an employee with Department of Astronautical, Electrical and Energetic Engineering, Sapienza University, Rome, Italy. Since 2014, he has been with Italian Space Agency, Rome, Italy, as program manager for earth observation program and space operations.