International Journal of Remote Sensing

ISSN: 0143-1161 (Print) 1366-5901 (Online) Journal homepage: www.tandfonline.com/journals/tres20

Crop yield prediction from remotely sensed vegetation indices and primary productivity in arid and semi-arid lands

Hadi H. Jaafar & Farah A. Ahmad

International Journal of Remote Sensing, 2015
Vol. 36, No. 18, 4570–4589, http://dx.doi.org/10.1080/01431161.2015.1084434

Crop yield prediction from remotely sensed vegetation indices and primary productivity in arid and semi-arid lands

Hadi H. Jaafar* and Farah A. Ahmad

Department of Agriculture, Faculty of Agriculture and Food Sciences, American University of Beirut, Beirut 110236, Lebanon

(*Received 21 January 2015; accepted 11 August 2015*)

Global demands for biomass and arable lands are expected to double in the next 35 years. Scarcity of water resources in arid and semi-arid areas poses a serious threat to their agricultural productivity and hence their food security. In this study, we examine whether crop yields can be predicted from remotely sensed vegetation indices and remotely sensed estimates of primary productivity. Spatial relationships between remotely sensed enhanced vegetation index (EVI), net photosynthesis (P_{Net}), and gross and net primary production (GPP and NPP, respectively) in irrigated semi-arid and arid agro-ecosystems since the beginning of the century are analysed. The conflict-affected country of Syria is selected as the case study. Relationships between EVI and crop yield are investigated in an effort to enhance food production estimates in affected areas outside governmental jurisdictions. Estimates of NPP derived from reported irrigated agriculture crop data in a semi-arid and an arid zone are compared to remotely sensed NPP in a geospatial environment. Results show that winter crop yields are correlated with spring GPP in semi-arid zones of the study area (R^2 = 0.85). Summer crop yield can be predicted from either cumulative summer EVI (R^2 = 0.77) or P_{Net} in most zones. Where fully irrigated fields are surrounded by hyper-arid landscape, summer P_{Net} was negative in all instances and EVI was inversely correlated with yield. NPP from crops was much higher (290 gC m^{-2} year^{-1}) in those regions than MOD17 NPP (70 gC m^{-2}), where 1.0 g of carbon is equivalent to 2.2 g of oven-dry organic matter (= 45% carbon by weight). The gap was less in semi-arid zones (2–39% difference). Overall crop-derived NPP for the period 2000–2013 was 322 *versus* 300 gC m^{-2} for that remotely sensed within the cropped zones of the political units. The results of this study are crucial to derive accurate estimates of irrigated agriculture productivity and to study the effect of the latter on net ecosystem carbon storage.

1. Introduction

Remote sensing of terrestrial biosphere productivity is a promising tool in the monitoring of crop production. It provides a basis for ecologists to study carbon cycle processes and to follow ecosystem canopy processes, photosynthesis, and transpiration (Bandaru et al. 2013). Spectral reflectance of terrestrial vegetation, detected by remote sensing (e.g. Moderate Resolution Imaging Spectroradiometer (MODIS)), is translated to real ecosystem variables, such as productivity and vegetation indices. The latter are valuable parameters in many applications: science, policy, and land management (Running et al. 2004). Remote sensing has been used in agriculture to analyse historical agricultural changes (Hole and Smith 2012; Lobell et al. 2002), to forecast production (Quarmby et al. 1993; Reeves, Zhao, and Running 2005), and further to develop agricultural control strategies.

*Corresponding author. Email: hj01@aub.edu.lb

During the last two decades, researchers have been active in many related areas: developing the theory and the algorithms of vegetation and productivity parameters (Zhao et al. 2005), validating their regular measurement (Fensholt, Sandholt, and Stisen 2006; Smith et al. 2014), estimating yield production (Quarmby et al. 1993; Reeves, Zhao, and Running 2005), and studying the relationships among different indices (Sims et al. 2008). Terrestrial net primary production (NPP), a measure of the vegetation growth, is one of the major resources of food and energy (Krausmann et al. 2013). Yield–NPP studies in the literature have focused on productive areas. Lobell et al. (2002) combined an inventory NPP approach with a parametric model that directly related light use efficiency (LUE) to NPP in order to refine crop NPP estimates in US agriculture for the period 1982–1998. While more than one-third of the world's food crops depends on irrigation (Bastiaanssen, Molden, and Makin 2000), remote-sensing data sets have identified agricultural land cover but have sometimes overlooked agricultural practices like crop type and irrigation (Monfreda, Ramankutty, and Foley 2008). In arid and semi-arid regions, irrigation is a necessity to conduct successful agriculture (Hole and Smith 2012). It has been argued that irrigated agriculture can locally either increase (by reducing water and nutrient limitations) or decrease NPP (by decreasing the growing season length through harvesting) (Long et al. 2006). Yield data can be converted to NPP, but the reverse process, which might be needed in cases where yield data are lacking, is not fully understood. Due to the lack of global crop-specific land cover data, remotely sensed estimates of NPP tend to be generic for cropland classification within the MODIS Landcover product. They also tend to overestimate carbon uptake by productive croplands (Bandaru et al. 2013). Vegetation indices have also been correlated with crop yields. Examples are the normalized difference vegetation index (NDVI) and the enhanced vegetation index (EVI). These have been evaluated *in situ* using radiometric measurements (Fensholt, Sandholt, and Stisen 2006). Empirical NDVI models were used frequently in the literature to predict crop yield, due to its simplicity and its link to photosynthesis activity (Reeves, Zhao, and Running 2005). Quarmby et al. (1993) estimated yield for wheat, cotton, rice, and maize crops to a high level of accuracy using a simple linear correlation with NDVI in northern Greece for the period 1986–1989. NDVI utilizes only the red and infrared bands and is subject to noise caused by underlying soil reflectance, especially in low-density vegetation canopies, and also to noise from atmospheric absorption. EVI utilizes the blue band for correcting for atmospheric aerosols. Regression estimates of yield tend to be controlled by spatio-temporal conditions. Reeves, Zhao, and Running (2005) assessed the estimation of wheat yield from MODIS gross primary production (GPP), using a conversion equation of carbon to yield rather than empirical regression models in two semi-arid regions of the USA. The results were significantly accurate at the state level, but less so at the county and climate district levels. Prediction of yield from remotely sensed estimates of productivity remain an active area of research.

The main aim of this work is to develop a methodology for remotely assessing the quantitative variation in crop yield at the level of regions and basins. This is important as the process of data collection within countries with an unstable security situation is seriously hindered by conflicts, and is therefore either non-existent or unreliable. It is also crucial for estimating crop water productivities and water mass balance at both the watershed and district level. In this study, four MODIS-derived parameters, NPP, GPP, net photosynthesis (P_{Net}), and EVI, are assessed in irrigated semi-arid and arid agro-ecosystems located in Syria, a conflict-affected country since 2011. The temporal variations of yield and the above parameters are analysed during summer and winter seasons, at governorate level, over the period 2000–2011 (pre-conflict years). The spatial correlation between productivity estimates and vegetation indices is examined in summer seasons for the major irrigated zones. MOD17 NPP is also compared to NPP derived from reported irrigated agriculture crop data. The main advantage of using remotely sensed yield-related parameters such as vegetation indices and estimates of primary productivity is highlighted in areas where access to the field is hindered by security situations (i.e. conflict zones), and in areas where yield reports are of questionable quality or sometimes falsified for political reasons.

2. Methodology

2.1. Study area

The study covers Syrian and Lebanese territories, including Syrian irrigated lands (Figure 1). The focus of the study is Syria during the period of recording, 2000–2013. Arid and semi-arid regions represent more than 70% of Syrian territory; the remaining 30% is characterized by a Mediterranean coastal (sub-humid) climate and is not included in this analysis. The MOD12Q1 product provides land-cover mapping (500 m resolution) according to the University of Maryland Department of Geography (UMD) classification (Friedl et al. 2010). The land cover of Syria is classified as mostly barren or sparsely vegetated (76.3%), open shrublands (13.3%), croplands (8.0%), urban built-up (0.87%), water (0.47%), and other classes (1.1%) (NASA LPDAAC 2014). In the last two decades, inland agricultural regions have been subject to agrarian reform. Many plans implemented for water resources management have contributed to a development in irrigation methods (Hole and Smith 2012). According to government websites, and within the last decade,

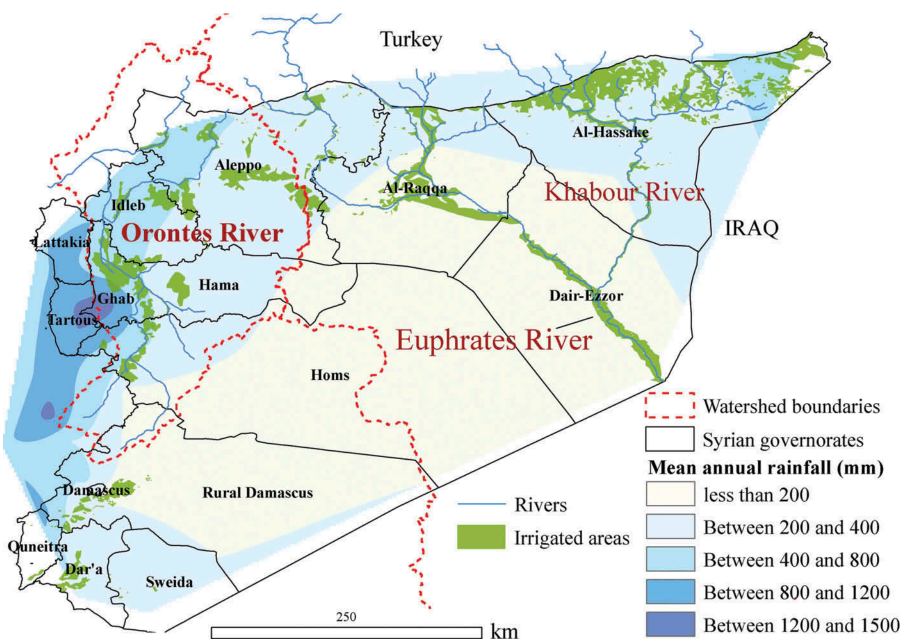


Figure 1. Study area map: Syrian governorates and Syrian irrigated lands (excluding coastal governorates and Lebanon).

about 30% of Syrian territories were cultivated, of which 24% were irrigated. This comprises the 8% croplands in the NASA LPDAAC product data set. Irrigated agriculture is prevalent in the fertile valleys of main rivers and tributaries (Euphrates, Khabour, and Orontes). Al-Hassake, Dair-Ezzor, and Al-Raqqa are the main governorates located in the arid and semi-arid zones. These governorates produce around one-third of Syrian crops and vegetables. Irrigated agriculture (1.3 million ha) accounts for more than 80% of the total production in those governorates. It is worth mentioning that the annual rainfall does not exceed 400 mm within the semi-arid regions (MOAAR 2014). Summer temperatures in the major agricultural zones range between 20°C at night and about 40°C during the day. Winter temperatures are in the range 3–12°C. Potential evaporation can reach more than 12 mm day^{-1} along the Euphrates and its tributary, the Khabour.

2.2. Data products

MODIS data sets, derived from both the Terra MODIS and Aqua MODIS instruments, provide land products, including GPP, annual NPP, daily net photosynthesis (P_{Net}), and EVI. NASA publishes these data sets from 2000 up to the present (NASA LPDAAC 2014). Daily GPP, daily P_{Net}, and annual NPP are derived from an algorithm using satellite-derived photosynthetically active radiation, meteorological data (NASA Data Assimilation Office (DAO)), estimated growth, and maintenance respiration (Zhao et al. 2005; Running et al. 2004). The meteorological data provided by DAO are derived from a global climate model, using collected data from all available surface weather observations globally every 3 hours. These data sets approximate canopy height conditions 10 m above the land surface (Running et al. 2004). We use GPP, P_{Net}, and NPP data sets, as improved by the Numerical Terra-dynamic Simulation Group (NTSG) at the University of Montana, that exclude cloud contamination pixels (NTSG 2014). The monthly GPP and P_{Net} data sets are extracted from the improved MOD17A2 products. The yearly GPP & the yearly NPP are derived from the improved MOD17A3 products. The monthly EVI used herein is derived from MOD13A3 product. MOD13A3 includes EVI on monthly basis, at 1 km spatial resolution. MOD13A3 EVI provides consistent spatial and temporal comparisons of vegetation conditions. Although NDVI has been widely used and applied, EVI is used herein as it minimizes canopy background variation and maintains sensitivity over dense vegetation conditions. NDVI saturates faster than EVI and hence may not be responsive to vegetation computed from atmospherically corrected bidirectional surface reflectance that has been masked for water, clouds, heavy aerosols, and cloud shadows. For a more detailed assessment, the reader is referred to Huete et al. (2002). The improved MOD17A2 is compiled from the following four monthly data sets – GPP, FPAR, PAR, and P_{Net} – at 1 km spatial resolution:

- GPP (gC m^{-2}): a cumulative composite of GPP values based on the radiation use efficiency concept (GPP = $\varepsilon \times$ (FPAR) $\times$ (PAR)); ε is the conversion efficiency, FPAR is the fraction of photosynthetically active radiation, and PAR is the incident radiation in photosynthetic wavelengths (Running et al. 2004)).
- P_{Net} (gC m^{-2}): the difference between GPP and maintenance respiration of leaves and fine roots.

The Syrian Ministry of Agriculture and Agrarian Reform (MOAAR) provides data on the production of winter and summer vegetables and crops, irrigated and non-irrigated, by all administrative units for the period of recording 2000–2012. Two years of records are

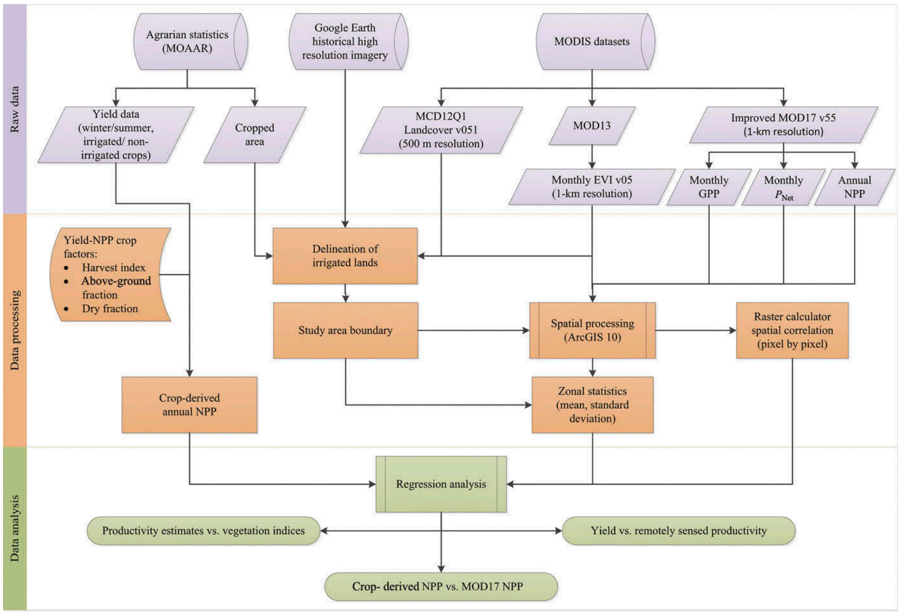


Figure 2. Flow chart showing methodology of data processing and analysis.

excluded from the analysis (2005 and 2006) due to missing/discrepant data. A flow chart of the methodology used in this work is illustrated in Figure 2.

2.3. Data processing

2.3.1. Delineation of summer irrigated lands

EVI was used as the basis of irrigated land delineation. The spatial mean of summer EVI for the period of recording (2000–2011) was calculated. By ‘summer’ we mean the period from 1 June to 31 August. The raster of mean summer EVI was compared to a series of high-resolution Google Earth imagery available for the summer seasons between 2000 and 2011. This comparison was based on visual analysis of field shapes and evidence of cultivation. Irrigated lands were extracted in a shape file from EVI pixels with values exceeding 0.12, which was found to correlate well with irrigated lands visible from the satellite imagery. The latter shape file was fine-tuned by manual delineation using high-resolution Google Earth imagery. The spatial mean of summer EVI within the delineated irrigated lands was found to have an average ranging between 0.14 and 0.22, and a maximum ranging between 0.24 and 0.45. Due to the sensitivity of EVI to topographic effect, summer mean EVI spatial raster was necessary but not sufficient to delineate summer irrigated lands (Matsushita et al. 2007). Relatively high EVI values were noted in forest areas on hilly terrain. Hence, the use of high-resolution Google Earth imagery along with topographic data proved to be indispensable for summer irrigated lands delineation.

The delineated irrigated lands were checked against two data sets: Land Cover Products (MOD12Q1) and statistical data from MOAAR. The delineated irrigated lands included an average of 80% of croplands and 16% of open shrublands. MOAAR provides statistics on irrigated lands in each political unit from 2000 to 2012. The areas of delineated irrigated lands were comparable to the average of total irrigated lands provided

Table 1. Comparison between irrigated lands reported by the Syrian Ministry of Agriculture and Agrarian Reform (MOAAR) and delineated irrigated lands.

Governorate	Irrigated area (ha) (MOAAR)	Delineated summer irrigated area (ha)	Difference from MOAAR (%)
Aleppo	190,806	200,672	5
Al-Hassake	403,418	470,010	17
Al-Raqqa	183,375	207,255	13
Dair-Ezzor	137,310	156,522	14
Damascus	65,823	55,510	−16
Dar'a	30,920	35,230	14
Hama	68,713	65,195	−5
Homs	53,386	48,300	−10
Idleb	54,572	42,383	−22
Sweida	1903	2053	8
Total	1,190,255	1,283,130	8

by the MOAAR over the period 2000–2011 in each political unit (Table 1). They were also in line with those from the Global Irrigated Area Map (Thenkabail et al. 2009). The delineated area of Ghab, a mainly agricultural zone located in the northwest of Syria, was found to be almost equivalent to the government-reported area (MOAAR).

2.3.2. *Geospatial processing*

The zonal statistics analysis of the data sets was performed using the spatial analysis tools in a geographic information system (ArcGIS Desktop Advanced™ 10.1). The above-mentioned data sets (monthly GPP, monthly P_{Net}, and monthly EVI) were downloaded for horizontal tile numbers 20 and 21 and vertical tile number 5, mosaicked, and clipped to cover the Syrian and Lebanese boundaries. In this study, the summer values of these data sets are of interest. Hence, the GPP, P_{Net}, and EVI values for June, July, and August were summed for each year. The zonal statistics for each study area were calculated. The mean summer GPP, mean summer P_{Net}, and the sum of EVI were obtained for each zone. Spatially standardized EVI, GPP, and P_{Net} were also derived for the year 2013 to evaluate the conflict impact, by subtracting the pixel temporal mean and dividing by the pixel temporal standard deviation (2000–2011).

2.3.3. *Spatial correlation between parameters*

The spatial correlation was investigated between (1) summer GPP and the sum of summer EVI, and between (2) summer P_{Net} and the sum of summer EVI. The spatial correlation r is based on Equation (1) (Abraham and Ledolter 2005):

$$r = \frac{1}{n-1}\sum_{i=1}^{n}\frac{X_i - \bar{X}}{S_x} \times \frac{Y_i - \bar{Y}}{S_y}, \qquad (1)$$

where X_i are the summer P_{Net} or summer GPP in year i, Y_i are the summer EVI values in year i, $\bar{X}$ is the mean of summer P_{Net} or summer GPP in n years, $\bar{Y}$ is the mean of summer EVI in n years; S_x is the standard deviation of summer P_{Net} or summer GPP in n years, and S_y is the standard deviation of summer EVI in n years. Pixel-by-pixel statistics were derived from multiple raster. The spatial mean and spatial standard deviations of these parameters were calculated for the period of recording 2001–2013. The spatial correlations between summer values of P_{Net} and EVI as well as P_{Net} and GPP were derived by applying Equation (1) on a pixel-by-pixel basis. Means and standard deviations of spatial statistics were calculated for the irrigated zones within each governorate.

2.3.4. Conversion of yield data to net primary production

NPP is the primary monitor of various ecological activities. It represents the total annual growth of vegetation on land (Smith et al. 2014). NPP measures the atmospheric carbon converted into plant biomass (Chapin et al. 2006; Running et al. 2004), and in agricultural lands can be calculated from agronomic data. Yield production was converted to carbon primary production using Equation (2) (Li et al. 2014; Bandaru et al. 2013; Monfreda, Ramankutty, and Foley 2008; Hicke and Lobell 2004; Prince et al. 2001):

$$\mathrm{NPP}_i = (\mathrm{EY}_i \times \mathrm{DF}_i \times C)/(\mathrm{HI}_i \times f_{\mathrm{AG}}), \tag{2}$$

where NPP_i is the net primary production of crop i, EY_i is the metric tonnes of economic yield per hectare of crop i, DF_i is the dry proportion of the economic yield i, C is the carbon content (0.45 gC (g dry matter)$^{-1}$), HI_i is the harvest index of crop i, and f_{AG} is the fraction of production allocated above-ground, directly related to the root:shoot ratio. The crops and their parameters used in this study are shown in Table 2.

The harvest index is a standard measure of the proportion of total above-ground biological yield allocated to the economic yield of the plant. It specifies the ratio of yield mass to above-ground biomass, varying between 0.23 and 1.0 depending on the type of yield. The root:shoot ratio indicates the ratio of below- to above-ground production. It is defined as the dry weight of root biomass divided by the dry weight of shoot biomass, ranging between 0.5 and 0.94. The crop-derived NPP was calculated for every crop in each agricultural unit based on Equation (2) and MOAAR statistical data for the period of recording 2000–2011. The sum of winter and summer crop NPP in each agricultural unit was compared to the mean of annual NPP (MOD17A3 products) in the delineated irrigated lands. The MOD17 NPP was produced as the difference between the accumulated daily P_{Net} (defined in Equation (3) as the difference between GPP and maintenance respiration of leaves and fine roots, R_{lr}), the estimated growth respiration required to produce leaves, fine roots, and new woody tissues, R_{g}, and the calculated live woody tissue maintenance respiration, R_{m}, as per Equation (4) (Zhao et al. 2005; Running et al. 2004):

$$P_{\mathrm{Net}} = (\mathrm{GPP}) - R_{\mathrm{lr}}, \tag{3}$$

$$\mathrm{NPP} = \sum (P_{\mathrm{Net}}) - R_{\mathrm{g}} - R_{\mathrm{m}}. \tag{4}$$

3. Results and discussion

3.1. Summer irrigated lands primary production and EVI

Cotton is the major crop in most arid and semi-arid governorates (Aleppo, Al-Raqqa, Dair-Ezzor, and Al-Hassake) and in the Ghab Valley (Orontes Basin). Vegetables are second, dominant in Homs, Hama, Idleb, and Dar'a. Maize is considered the second major

Table 2. Parameters used for converting yield to carbon production (Li et al. 2014; Bandaru et al. 2013; Monfreda, Ramankutty, and Foley 2008; Hicke and Lobell 2004; Prince et al. 2001).

Crop	Harvest index (HI)	Dry fraction (DF)	Above-ground fraction (f_{AG})
Barley	0.49	0.89	0.81
Broad beans	0.45	0.13	0.85
Cabbages/cauliflowers	0.45	0.08	0.85
Cantaloupes	0.45	0.10	0.85
Chickpeas	0.44	0.90	0.85
Cotton	0.55	0.92	0.86
Cucumbers	0.45	0.04	0.85
Dry beans	0.55	0.90	0.74
Eggplants	0.45	0.08	0.85
Garlic	0.45	0.13	0.85
Grazing maize	1.0	0.35	0.85
Green beans	0.45	0.10	0.85
Lentils	0.46	0.89	0.85
Maize	0.45	0.89	0.85
Oil seeds	0.52	0.73	0.80
Okra	0.45	0.10	0.85
Onions	0.45	0.13	0.85
Peanuts	0.40	0.92	0.80
Peppers	0.28	0.80	0.50
Potatoes	0.50	0.28	0.80
Sorghum	0.40	0.89	0.80
Sugar beet	0.40	0.12	0.80
Sunflower	0.39	0.94	0.94
Tobacco	0.28	0.80	0.80
Tomatoes	0.45	0.06	0.85
Watermelons	0.45	0.09	0.85
Wheat	0.39	0.89	0.81

crop in Aleppo, Al-Raqqa, and Dair-Ezzor. Figure 3 shows the temporal means (2000–2011) for cumulative summer EVI, cumulative summer P_{Net}, cumulative summer GPP, and their standardized values for the year 2013 over the study area. The standardized values are used to indicate changes due to the conflict (rainfall in 2013 was higher than average). EVI highlights the major irrigated lands. The cumulative summer EVI ranges between −0.14 for lands adjacent to water bodies and 1.35 for areas of heavy irrigation and full crop cover (Figure 3(*a*)). The highest P_{Net} (268 gC/m^2) is in the Ghab Valley in Northern Hama, and the lowest (negative) summer P_{Net} is in Dair-Ezzor and Al-Raqqa irrigated lands (rainfall <200 mm year^{-1}) on the banks of the Euphrates, which are surrounded by a hyper-arid landscape (−195 gC m^{-2}) (Figure 3(*c*)). Summer P_{Net} is almost zero in areas where annual rainfall is between 200 and 400 mm (Al-Hassake). Summer GPP is lowest in the arid area of Dair-Ezzor (Figure 3(*e*)). It appears to increase with rainfall (higher in semi-arid regions), possibly due to a lower vapour pressure deficit (VPD) in these regions. High VPD contributes to physiological water stress control in plants and a reduction in PAR conversion efficiency ε, hence limiting daily GPP (Running et al. 2004). Also, variations in photosynthetic LUE resulting from water stress caused by drought or lack of sufficient irrigation can affect MOD17 GPP and NPP products. LUE varies with vegetation type (Long et al. 2006). Moreover, the daily meteorological data used in the algorithm have a maximum resolution of 1 degree (approximately 108 km), while the MODIS product has a resolution of 1 km or better. This could be problematic in

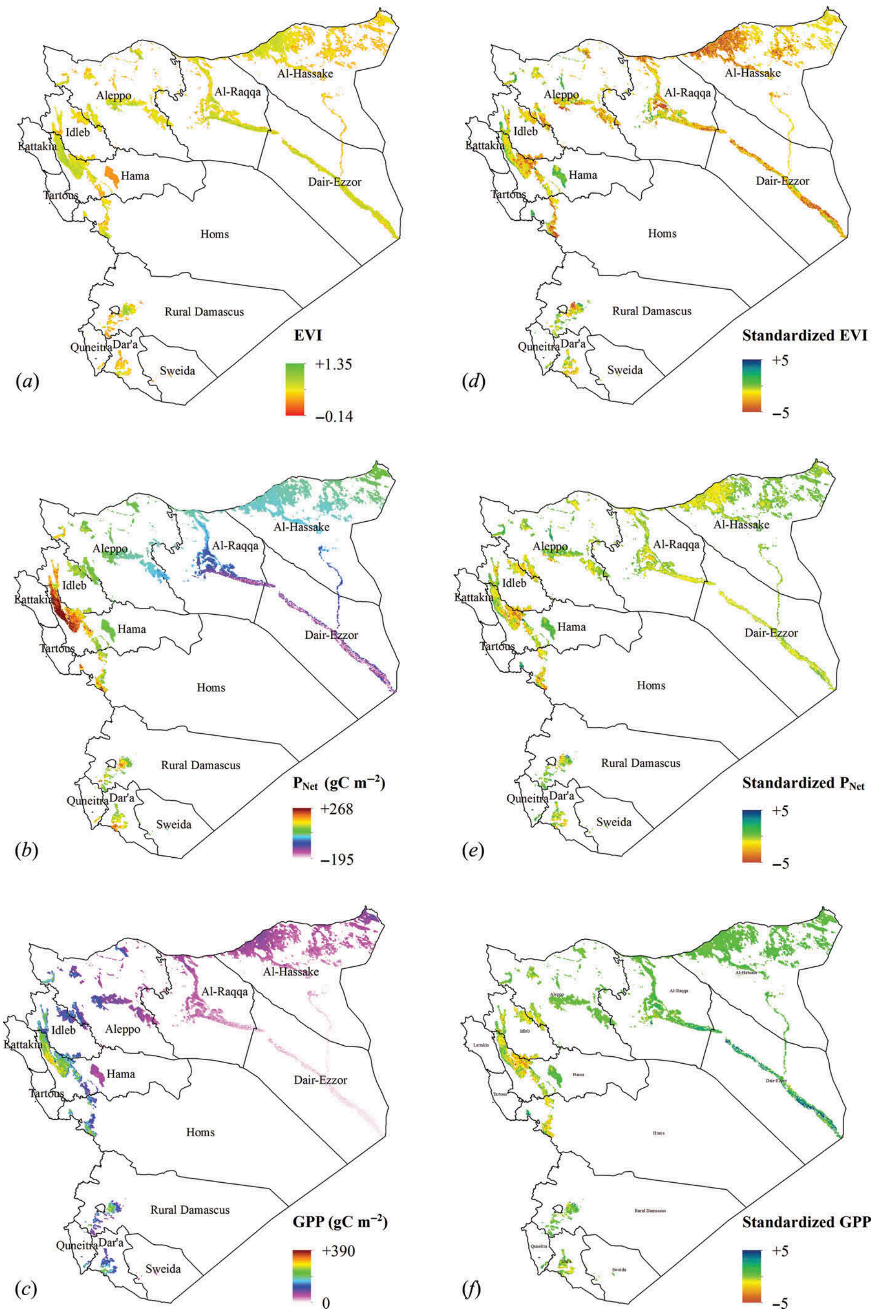


Figure 3. Means of (*a*) cumulative summer EVI (2001–2011), (*b*) summer P_{Net} (2000–2011) (gC m^{-2}), (*c*) summer GPP (2000–2011) (gC m^{-2}), and standardized values for the year 2013 of (*d*) cumulative summer EVI, (*e*) summer P_{Net}, and (*f*) summer GPP.

irrigated croplands surrounded by hyper-arid landscape, where averaging of the temperature in such a pixel does not represent the climatic conditions in the irrigated corridor around the Euphrates river. This average temperature is used within the MOD17 algorithm to calculate daily leaf maintenance respiration as well as fine root mass respiration, both of which increase exponentially with temperature. Overestimation of temperature will ultimately result in a higher respiration and consequently a lower P_{Net}, yielding a lower NPP than expected.

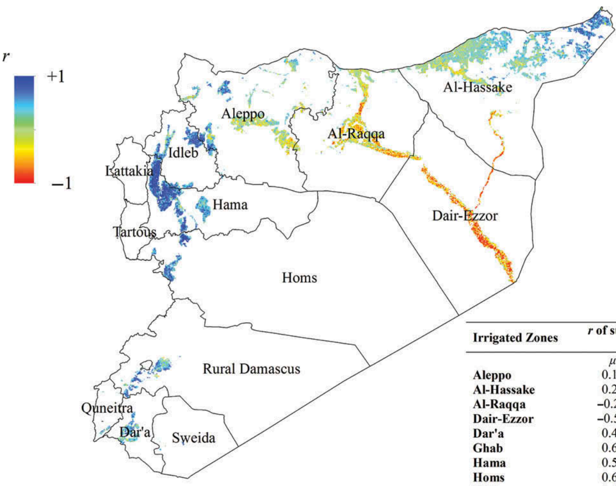
Spring GPP also is calculated to determine whether it is dependent on summer P_{Net}. Mean spring GPP varies between 34 gC m^{-2} in hyper-arid zones to 772 gC m^{-2} in heavily irrigated zones (figure not shown). Standardized summer EVI and standardized summer P_{Net} for the year 2013 are computed by spatially subtracting the respective pixel-by-pixel EVI and P_{Net} mean for the 2000–2011 and dividing by the pixels' standard deviation. Figure 3(*b*) and (*d*) shows that vegetation EVI dropped significantly in Al-Hassake, Ghab, Idleb, and Dair-Ezzor. The primary reasons for this decrease are the deficiency of water supplies (empty reservoirs, destruction of main water channels and pump stations), lack of fuel, lack of seeds and fertilizers, and the displacement of farmers (Swiss Agency for Development and Cooperation SADC 2014). While a drop in EVI and P_{Net} in semi-arid areas has been correlated with a drop in summer irrigated agricultural yields, the relationship is reversed in the arid region of Dair-Ezzor. This is verified by the negative mean P_{Net} in Dair-Ezzor (Figure 3(*d*)) and the higher standardized P_{Net} in the same area in 2013, indicating a decrease in agricultural production, as will be shown in the following sections. Dair-Ezzor irrigated lands are mainly composed of cotton (48%) and maize (29%). More data years are needed to verify whether this is sufficiently consistent to derive any conclusions. Smith et al. (2014) found that cereal, oil, and sugar crops reduce GPP and NPP, while some C4 perennial crops can minimize this impact. Standardized summer GPP for the year 2013 (Figure 3(*f*)) shows extreme low values in some regions of Hama bordering Idleb, and some parts of Damascus and environs (the scene to major battles in 2013).

3.2. Spatial correlation between EVI and primary productivity

The spatial correlation between summer EVI and summer P_{Net} and its statistics are shown in Figure 4. A high negative correlation is noted between summer EVI and summer P_{Net} in Dair-Ezzor and Al-Raqqa governorates, with a significant inverse relationship in Dair-Ezzor irrigated areas along the Euphrates river (mean of this negative correlation is −0.58). It will be shown later that the higher regression coefficient between EVI and P_{Net} yielded a higher spatial correlation (either positive or negative) between the two tested parameters.

3.3. Summer crop yield correlation with remotely sensed parameters

The time series variation of summer crops in major governorates is shown in Figure 5. In Dair-Ezzor and Al-Raqqa, on the banks of the Euphrates and Al-Khabour, respectively, summer crops could not be predicted from MOD17 summer P_{Net}. In Al-Raqqa, however, summer yield variation (MOAAR) is in accord with MOD17 P_{Net} since 2008. In Ghab and Hama, crop yields are highly correlated with P_{Net} with significant F-values ($p < 0.05$) of 12.04 and 8.55, respectively. The major crop planted in Ghab is cotton (85%). Lower r^2 values are found in Al-Hassake (on the banks of the Al-Khabour) and Dar'a irrigated lands (0.47 and 0.42, respectively).



Irrigated Zones	r of summer P_{Net} vs. total summer EVI	
	μ	σ
Aleppo	0.16	0.35
Al-Hassake	0.24	0.31
Al-Raqqa	−0.22	0.33
Dair-Ezzor	−0.58	0.29
Dar'a	0.40	0.24
Ghab	0.69	0.25
Hama	0.50	0.25
Homs	0.61	0.24
Idleb	0.61	0.22
Quneitra	0.35	0.22
Rural Damascus	0.49	0.27
Sweida	0.55	0.19

Figure 4. Spatial correlation r (2001–2013) between summer P_{Net} and total summer EVI; μ and σ are the mean and the standard deviation, respectively, of r for irrigated zones within the governorates.

In Hama, Al-Hassake, Ghab, and Dar'a, summer crops are slightly better predicted by mean summer GPP than by mean summer P_{Net}. Either summer P_{Net} or summer GPP can be used to predict crop yields in those governorates. High yield is correlated with high summer GPP and P_{Net} (Table 3). However, in Dair-Ezzor irrigated lands, high yield is correlated with low summer GPP. Other governorates did not show significant relationships, possibly due to discrepancies in land-cover classification in the MOD17 algorithm, indicating the need for better validation of the land-cover product.

Summer crop yields can be readily predicted from remotely sensed cumulative summer EVI, to a varying degree. Hama and Ghab irrigated lands show the highest correlation between crop yields and cumulative summer EVI (0.77 and 0.73, respectively). These areas are characterized by heavy agricultural activity, with a total production of 1.6 million t year^{-1} (MOAAR 2014). Major crops include cotton and vegetables. In 2013, 1 year after the Syrian conflict started, a variable drop in summer production is noted in major irrigated lands, most significantly in Homs, Dar'a, and Dair-Ezzor. Summer production decreased by 95% in Homs and to half in Dar'a. Dair-Ezzor production was around 200,000 t in summer 2012, while its average summer production was 300,000 t during the period 2000–2011. The drop was less significant in Ghab. In regard to Hama, no significant drop is noted as this was not a major combat zone in 2012. In Al-Raqqa and Dair-Ezzor, the major producing zones for Syria with summer agricultural production amounting to more than 700,000 t (MOAAR 2014), higher yields are correlated with low EVI values (Figure 6). It is worth noting that the latter two governorates are the only two where maize accounts for more than one quarter of the irrigated lands, second to cotton. These two areas are irrigated from Lake Assad (Tabaqa Dam) in Al-Raqqa fed by the Euphrates river, and the irrigated ecosystem is surrounded by a hyper-arid climate. The relationship is not evident in other governorates, possibly due mainly to the presence of fallow parcels with natural vegetation growth in between. Summer P_{Net} and summer GPP are correlated with the cumulative summer EVI in most governorates (Table 3). Ghab and Hama show the highest correlation.

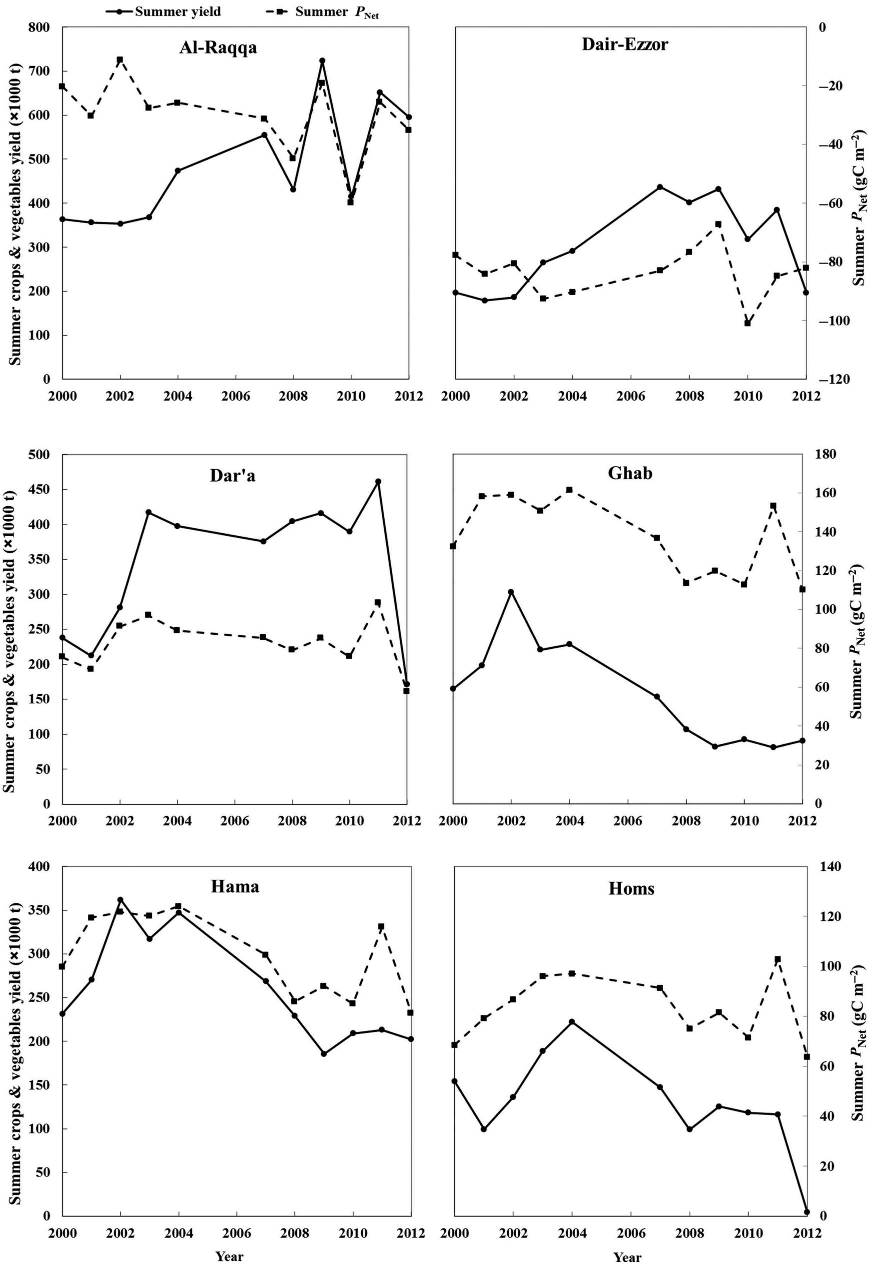


Figure 5. Time series of summer crop and vegetable yield (left) and summer P_{Net} (right) for the period of recording 2000–2012.

In Al-Raqqa, no significant correlation is evident. In Dair-Ezzor, where P_{Net} is negative in all instances, an increased EVI is highly correlated with a negative P_{Net} (R^2 = 0.72) but not significantly correlated with summer GPP. Mean summer P_{Net} is highly correlated with spring GPP in Lebanon in general, and also in the Litani and Orontes basins, but not in the arid zones of other Syrian governorates, indicating a lack of dependence of summer P_{Net} on

Table 3. Regression analysis for major agricultural Syrian governorates for the recording period 2000–2011.

			Summer yield *vs* summer GPP			Summer yield *vs* summer P_{Net}			Summer P_{Net} *vs* summer ΣEVI		
Governorate	Major summer crops		Slope	R^2	*F*-statistic	Slope	R^2	*F*-statistic	Slope	R^2	*F*-statistic
Al-Hassake	Cotton (85%)	Watermelon (7%)	12.06	0.60	11.91*	10.21	0.47	7.14*	0.01	0.08	0.99
Al-Raqqa	Cotton (65%)	Maize (22%)	−0.60	0.00	0.02	1.32	0.02	0.16	−0.01	0.08	0.96
Damascus	Alfalfa (36%)	Vegetables (29%)	0.79	0.22	2.20	0.62	0.06	0.48	0.02	0.29	4.45
Dar'a	Vegetables (49%)	Sesame (13%)	4.28	0.55	9.89*	5.18	0.42	5.77*	0.01	0.19	2.58
Dair-Ezzor	Cotton (48%)	Maize (29%)	−21.56	0.59	11.45*	2.58	0.06	0.48	−0.02	0.72	28.09*
Ghab	Cotton (52%)	Vegetables (12%)	2.08	0.57	10.69*	2.77	0.52	8.55*	0.02	0.83	52.59*
Hama	Vegetables (34%)	Cotton (22%)	2.34	0.69	17.86*	3.08	0.60	12.04*	0.02	0.79	42.53*
Homs	Vegetables (42%)	Sorghum (9%)	1.10	0.27	2.93	1.54	0.21	2.13	0.02	0.56	14.21*
Idleb	Vegetables (23%)	Cotton (22%)	−0.21	0.01	0.04	−0.11	0.00	0.01	0.02	0.53	12.42*
Sweida	Tomato (45%)	Watermelon (30%)	1.3	0.14	1.24	2.45	0.24	2.45	0.01	0.17	2.25

Note: Summer yield (×1000 t), summer P_{Net}, and summer GPP in (gC m^{-2}), and summer EVI (×10^4); *significant R^2 using *F*-statistic ($p < 0.05$).

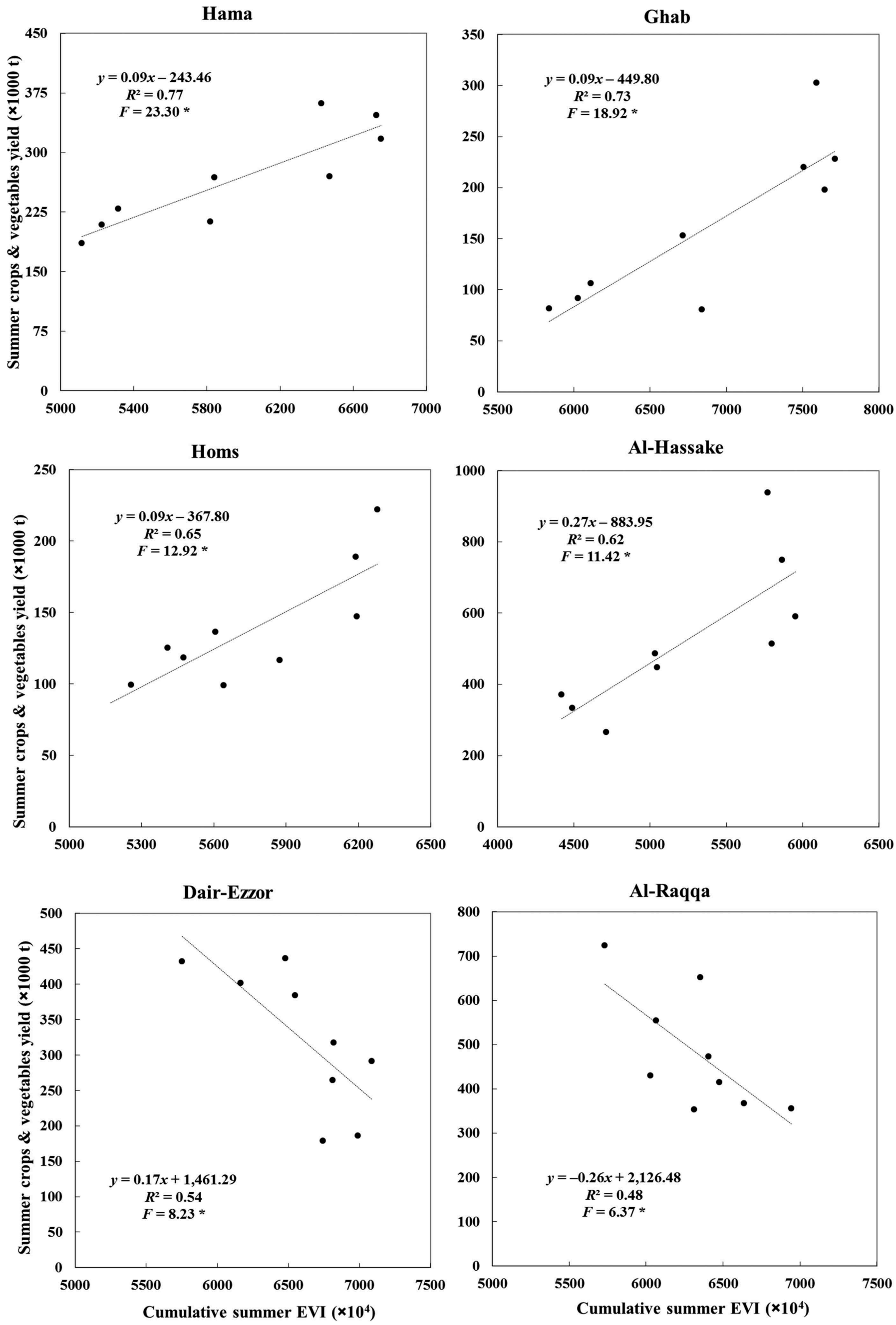


Figure 6. Regression between summer crop and vegetable yield and cumulative summer EVI for major agricultural Syrian governorates for 2001–2011 (graphs sorted by decreasing significance of R^2; *significant F-statistics at $p < 0.05$).

spring GPP. As spring GPP is related to winter crop yield, it is imperative that any summer EVI and P_{Net} in those governorates are directly related to agricultural activities and not to natural vegetation.

3.4. Winter production

Wheat and barley are the two major winter crops in Syria, covering more than 80% of the total winter cultivated area. Syria produces around 4.0 million t year^{-1} of wheat, with 35% produced in Al-Hassake (MOAAR 2014). Winter crops and vegetables and spring GPP follow a very similar pattern according to time series analysis for the period 2000–2012 and in most Syrian irrigated lands (Figure 7). As with summer production, winter production decreased in 2012. In Ghab, winter production reached its lowest value in

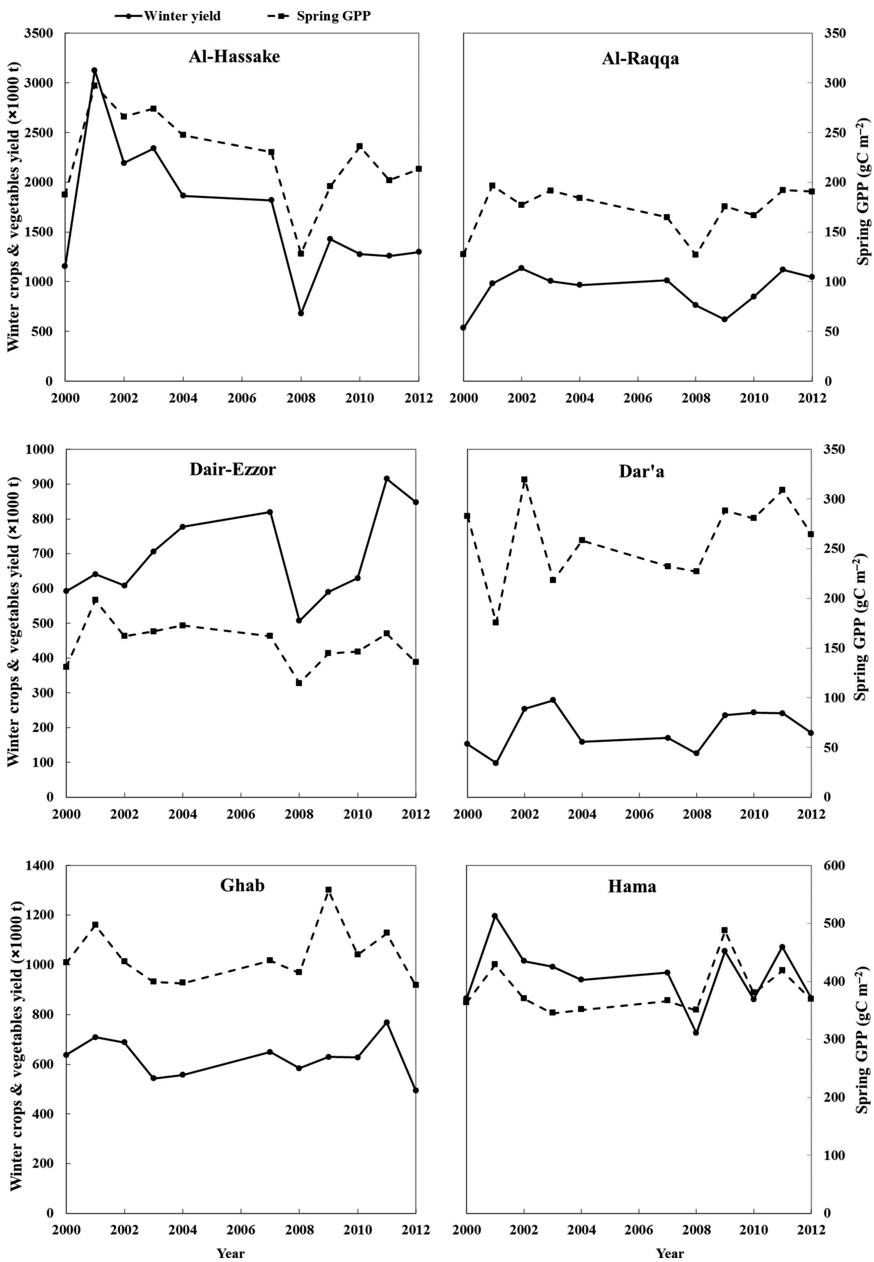


Figure 7. Time series of winter crop and vegetable yield (left) and spring GPP (right) for the period of recording 2000–2012.

the most recent decade (494,000 t). Spring GPP varies between 100–300 gC m^{-2} in arid zones; it increases to 600 gC m^{-2} in Ghab, Hama, Idleb, and Homs irrigated lands. Within the major agricultural areas of Syria, winter crop yields can be estimated from spring GPP values (Figure 8). In Dair-Ezzor (by the Euphrates), rural Damascus (Ghouta area), and Sweida, where fruit trees prevail, the correlations among winter and vegetables and spring GPP are weaker. Winter crops are best predicted from spring GPP in Al-Hassake and Dar'a (Figure 8). Spring GPP is correlated with rainfall in Al-Hassake, Al-Raqqa, and Dair-Ezzor. Winter crop production can be estimated mainly from annual rainfall in the semi-arid regions of Syria, but not in the hyper-arid irrigation region of Dair-Ezzor where no significant correlation occurs ($r^2 = 0.09$). Dair-Ezzor is mainly dependent on irrigated cultivation since rainfall rarely exceeds 150 mm $year^{-1}$. In other governorates, rain-fed cultivation comprises between 30% (Al-Hassake) and 80% (Sweida) of the total (MOAAR 2014). Non-irrigated crops and vegetables in Al-Raqqa are highly correlated with rainfall ($R^2 = 0.72$).

3.5. Comparison of annual MOD17 NPP to crop-derived annual NPP

In Idleb, Dar'a, and Hama (fruit, vegetables, cotton, and olives), crop-derived NPP and MOD17 NPP run parallel (Figure 9). The percentage difference in mean values (2001–2012) is 15 in Idleb, 18 in Dar'a, and 25 in Hama. In other governorates, crop-derived NPP is much higher than MOD17 NPP. In Dair-Ezzor, croplands produce annually between 243 and 333 gC m^{-2} while according to the MOD17 product, less than 110 gC m^{-2} is produced annually. In regard to the spatial correlation between EVI and P_{Net} (Figure 4) and also the data in Figure 9, it appears that in regions where EVI is highly correlated with P_{Net}, MODIS NPP estimates appear to be higher than crop-derived NPP. Along the Euphrates irrigated lands in Raqqa and Deir Ezzor, where cotton and maize are the major summer crops, the relationship is reversed. EVI is inversely related to crop yield and, at the same time, crop-derived NPP is consistently higher than the MOD17 estimate. The differences between crop-derived NPP and remotely sensed estimates can be attributed to several factors. The MOD17 NPP product is not crop specific, being based on one class of land cover (croplands). Indeed NPP from different crops will differ, and consequently the MODIS product does not allow for this distinction in its algorithm. Mixed pixels can be another limitation. A 1 km^2 pixel might be classified as cropland while in fact it could be a mixture of vegetation cover, contaminated with bare soil, or other types of land cover. Quantifying the degree of uncertainty imposed by the MODIS Landcover product assumptions on the MOD17 NPP estimate is yet to be determined. For a more detailed analysis, the reader is referred to Li et al. (2014). Another reason for this difference is that crop yield-derived NPP estimates are derived from the yield itself, which is usually a function of external inputs such as irrigation, fertilizer, cultural practices, crop health, and others. Not all of these inputs are directly reflected in measured indices such as leaf area index or modelled by biophysical factors such as temperature and LUE.

4. Conclusion

In this research we examined the possibility of predicting yields from remotely sensed vegetation-related indices (EVI) and productivity estimates (GPP, NPP, and P_{Net}). Through utilizing annual and seasonal cropland yield data over the last decade within Syria, and analysing their spatial relationship with remotely sensed estimates of EVI

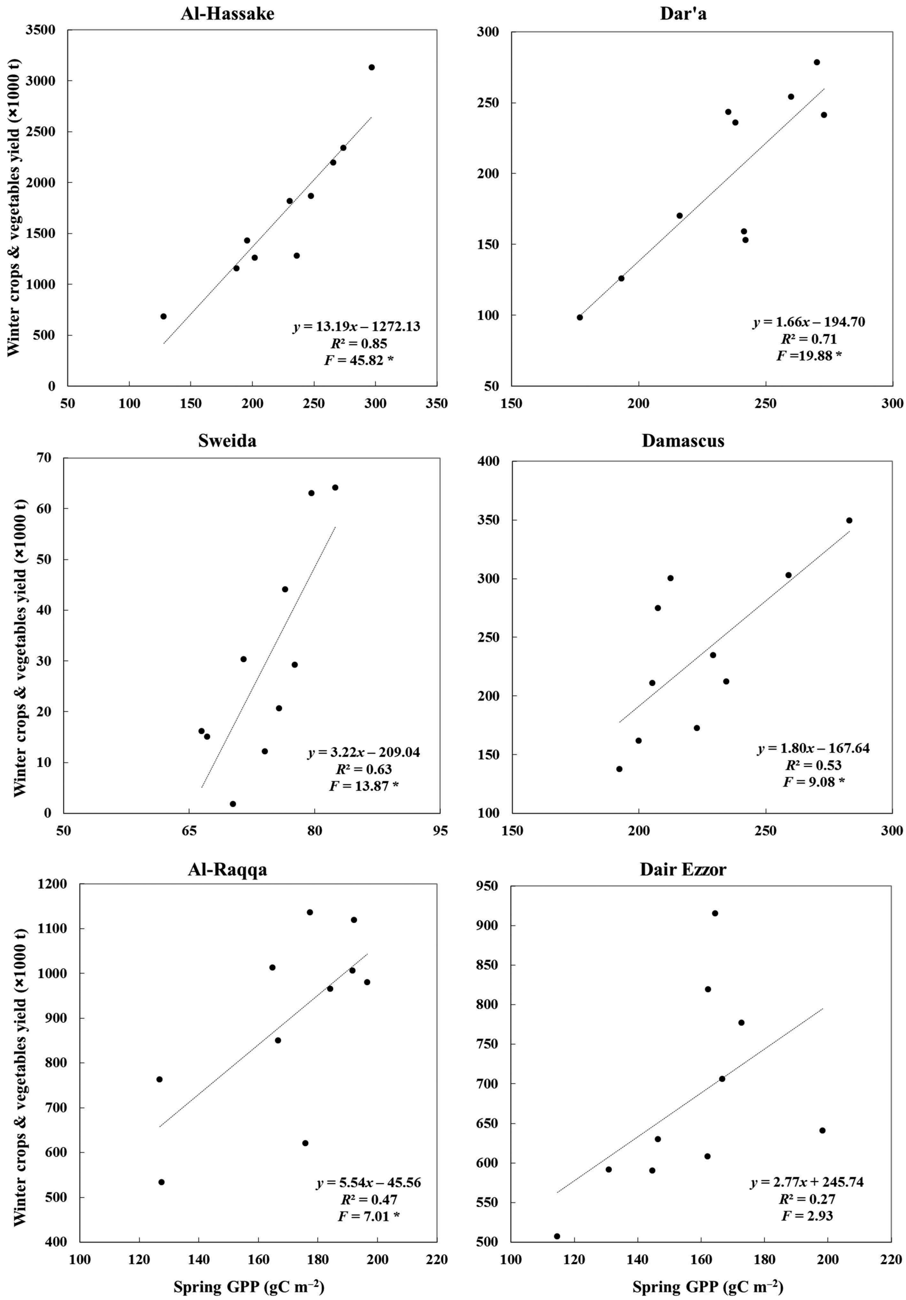


Figure 8. Regression between winter crop and vegetable yield and spring GPP for major agricultural Syrian governorates for 2001–2011 (graphs sorted by decreasing significance of R^2; *significant F-statistics at $p < 0.05$).

and terrestrial productivity, we assessed whether crop yields can be predicted from any of the above parameters. The aim was to provide quantitative data on the production of summer irrigated crops in conflict zones where access to data is hindered by the security situation and ongoing military activity. We analysed relationships between

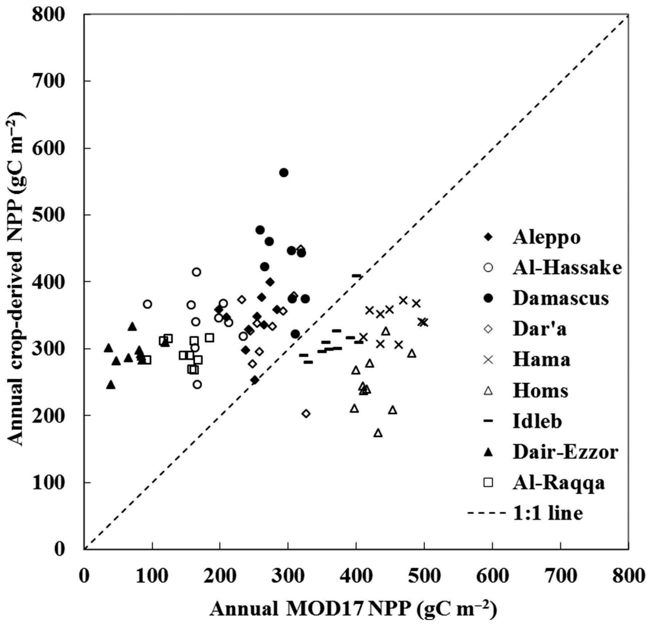


Figure 9. Comparison of crop-derived annual NPP and spatial mean of MOD17 annual NPP for major agricultural Syrian governorates for the period of recording 2001–2012. Remotely sensed estimation of NPP underestimates crop-derived NPP in arid regions (by the Euphrates and Khabour).

crop yields on the one hand, and EVI, GPP, NPP, and P_{Net} on the other. We show the advantages and limitations of using these relationships when used for crop yield prediction. We also tested the validity of the derived regression model by comparing its output to production data from the pre-conflict years. Our results show that summer crop yields can be predicted from remotely sensed EVI, but not as readily from P_{Net} and GPP of the MOD17 product. It is noted that cumulative summer EVI is inversely proportional to summer yields in areas where maize had been planted. A negative spatial correlation between EVI and summer P_{Net} is found for those irrigated areas within hyper-arid ecosystems. Winter crops (mainly wheat and barley) can be estimated from remotely sensed spring GPP. Annual MODIS NPP is higher than crop-derived NPP in areas where EVI and summer P_{Net} were positively correlated, and vice versa. We deduce that improvements to MODIS productivity estimation should be sought, especially by using a higher-resolution climatic data set than that currently used, by using crop-specific land-cover data at better resolution. Developing a relatively simple and inexpensive, spatially based model for the prediction of crop yield and forecasting food shortages could prove vital for humanitarian relief organizations. Such a model would be important in times of war and also in reducing the impacts of both natural and human-induced disasters.

Acknowledgements

The authors are thankful for the anonymous reviewers who improved the manuscript through useful comments.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was made possible by funds from the American University of Beirut Research Board (URB), as well as funds from ICARDA's Middle East Water and Livelihood Initiative (WLI) funded by USAID. The views expressed in this publication are those of the authors and do not necessarily reflect those of the funding agencies.

ORCID

Hadi H. Jaafar http://orcid.org/0000-0003-2612-3191

References

Abraham, B., and J. Ledolter. 2005. *Statistical Methods for Forecasting*. New York: Wiley Interscience.

Bandaru, V., T. O. West, D. M. Ricciuto, and R. C. Izaurralde. 2013. "Estimating Crop Net Primary Production Using National Inventory Data and Modis-Derived Parameters." *ISPRS Journal of Photogrammetry and Remote Sensing* 80: 61–71. doi:10.1016/j.isprsjprs.2013.03.005.

Bastiaanssen, W. G., D. J. Molden, and I. W. Makin. 2000. "Remote Sensing for Irrigated Agriculture: Examples from Research and Possible Applications." *Agricultural Water Management* 46: 137–155. doi:10.1016/S0378-3774(00)00080-9.

Chapin, F. S., G. M. Woodwell, J. T. Randerson, E. B. Rastetter, G. M. Lovett, D. D. Baldocchi, D. A. Clark, M. E. Harmon, D. S. Schimel, R. Valentini, C. Wirth, J. D. Aber, J. J. Cole, M. L. Goulden, J. W. Harden, M. Heimann, R. W. Howarth, P. A. Matson, A. D. McGuire, J. M. Melillo, H. A. Mooney, J. C. Neff, R. A. Houghton, M. L. Pace, M. G. Ryan, S. W. Running, O. E. Sala, W. H. Schlesinger, and E.-D. Schulze. 2006. "Reconciling Carbon-Cycle Concepts, Terminology, and Methods." *Ecosystems* 9: 1041–1050. doi:10.1007/s10021-005-0105-7.

Fensholt, R., I. Sandholt, and S. Stisen. 2006. "Evaluating MODIS, MERIS, and VEGETATION Vegetation Indices Using In Situ Measurements in a Semiarid Environment." *IEEE Transactions on Geoscience and Remote Sensing* 44 (7): 1774–1786. doi:10.1109/TGRS.2006.875940.

Friedl, M. A., D. Sulla-Menashe, B. Tan, A. Schneider, N. Ramankutty, A. Sibley, and X. Huang. 2010. "MODIS Collection 5 Global Land Cover: Algorithm Refinements and Characterization of New Datasets." *Remote Sensing of Environment* 114: 168–182. doi:10.1016/j.rse.2009.08.016.

Gibson, G. R., J. B. Campbell, and C. E. Zipper. 2012. "Spatio-Temporal Patterns of Cropland Land Cover Change in Iraq during Operation Iraqi Freedom and the End of Sanctions." In *War and Agriculture: Three Decades of Agricultural Land Use and Land Cover Change in Iraq*, edited by G. R. Gibson, 8–45. Unpublished doctoral dissertation, Virginia Tech University, Blacksburg.

Hicke, J. A., and D. B. Lobell. 2004. "Spatiotemporal Patterns of Cropland Area and Net Primary Production in the Central United States Estimated from USDA Agricultural Information." *Geophysical Research Letters* 31: 1–5.

Hole, F., and R. Smith. 2012. "Arid Land Agriculture in Northeastern Syria." In *Land Change Science*, edited by G. Gutman et al., 213–226. Springer Science + Business Media B.V. doi:10.1007/978-1-4020-2562-4.

Huete, A., K. Didan, T. Miura, E. P. Rodriguez, X. Gao, and L. G. Ferreira. 2002. "Overview of the Radiometric and Biophysical Performance of the MODIS Vegetation Indices." *Remote Sensing of Environment* 83: 195–213. doi:10.1016/S0034-4257(02)00096-2.

Krausmann, F., K.-H. Erb, S. Gingrich, H. Haberl, A. Bondeau, V. Gaube, C. Lauk, C. Plutzar, and T. D. Searchinger. 2013. "Global Human Appropriation of Net Primary Production Doubled in the 20th Century." *Proceedings of the National Academy of Sciences* 110 (25): 10324–10329. doi:10.1073/pnas.1211349110.

Li, Z., S. Liu, Z. Tan, N. B. Bliss, C. J. Young, T. O. West, and S. M. Ogle. 2014. "Comparing Cropland Net Primary Production Estimates from Inventory, a Satellite-Based Model, and a

Process-Based Model in the Midwest of the United States." *Ecological Modelling* 277: 1–12. doi:10.1016/j.ecolmodel.2014.01.012.

Lobell, D. B., J. A. Hicke, G. P. Asner, C. B. Field, C. J. Tucker, and S. O. Los. 2002. "Satellite Estimates of Productivity and Light Use Efficiency in United States Agriculture, 1982–98." *Global Change Biology* 8: 722–735. doi:10.1046/j.1365-2486.2002.00503.x.

Long, S. P., X.-G. Zhu, S. L. Naidu, and D. R. Ort. 2006. "Can Improvement in Photosynthesis Increase Crop Yields?" *Plant Cell Environ* 29 (3): 315–330. doi:10.1111/pce.2006.29.issue-3.

Matsushita, B., W. Yang, J. Chen, Y. Onda, and G. Qiu. 2007. "Sensitivity of the Enhanced Vegetation Index (EVI) and Normalized Difference Vegetation Index (NDVI) to Topographic Effects: A Case Study in High-density Cypress Forest." *Sensors* 7: 2636–2651. doi:10.3390/s7112636.

MOAAR (Ministry of Agriculture and Agrarian Reform in Syria). 2014. *Statistical datasets*. http://moaar.gov.sy

Monfreda, C., N. Ramankutty, and J. A. Foley. 2008. "Farming the Planet: 2. Geographic distribution of Crop Areas, Yields, Physiological Types, and Net Primary Production in the Year 2000." *Global Biogeochemical Cycles* 22: 1–19.

NASA LPDAAC (National Aeronautics and Space Administration Land Processes Distributed Active Archive Center). 2014. *MODIS Products*. Sioux Falls, SD: USGS/Earth Resources Observation and Science (EROS) Center. https://lpdaac.usgs.gov/products

NTSG (Numerical Terradynamic Simulation Group). 2014. *MODIS Products*. http://www.ntsg.umt.edu/project/mod17

Prince, S. D., J. Haskett, M. Steininger, H. Strand, and R. Wright. 2001. "Net Primary Production of US Midwest Croplands from Agricultural Harvest Yield Data." *Ecological Applications* 11: 1194–1205. doi:10.1890/1051-0761(2001)011[1194:NPPOUS]2.0.CO;2.

Quarmby, N. A., M. Milnes, T. L. Hindle, and N. Silleos. 1993. "The Use of Multi-Temporal NDVI Measurements from AVHRR Data for Crop Yield Estimation and Prediction." *International Journal of Remote Sensing* 14: 199–210. doi:10.1080/01431169308904332.

Ramankutty, N., A. T. Evan, C. Monfreda, and J. A. Foley. 2008. "Farming the Planet: 1. Geographic distribution of Global Agricultural Lands in the Year 2000." *Global Biogeochemical Cycles* 22: 1–19.

Reeves, M. C., M. Zhao, and W. Running. 2005. "Usefulness and Limits of MODIS GPP for Estimating Wheat Yield." *International Journal of Remote Sensing* 26 (7): 1403–1421. doi:10.1080/01431160512331326567.

Running, S. W., R. R. Nemani, F. A. Heinsch, M. Zhao, M. Reeves, and H. Hashimoto. 2004. "A Continuous Satellite-Derived Measure of Global Terrestrial Primary Production." *BioScience* 54: 547–560. doi:10.1641/0006-3568(2004)054[0547:ACSMOG]2.0.CO;2.

SADC (Swiss Agency for Development and Cooperation). 2014. *Syria: The impact of the conflict on population displacement, water and agriculture in the Orontes River basin*. http://graduateinstitute.ch/files/live/sites/iheid/files/shared/news/2014/Syria%20Feb%202014-2.pdf

Sims, D. A., A. F. Rahman, V. D. Cordova, B. Z. El-Masri, D. D. Baldocchi, P. V. Bolstad, L. Flanagan, A. Goldstein, D. Hollinger, and L. Misson. 2008. "A New Model of Gross Primary Productivity for North American Ecosystems Based Solely on the Enhanced Vegetation Index and Land Surface Temperature from MODIS." *Remote Sensing of Environment* 112: 1633–1646. doi:10.1016/j.rse.2007.08.004.

Smith, W. K., C. C. Cleveland, S. C. Reed, and S. W. Running. 2014. "Agricultural Conversion without External Water and Nutrient Inputs Reduces Terrestrial Vegetation Productivity." *Geophysical Research Letters* 41: 449–455. doi:10.1002/2013GL058857.

Thenkabail, P. S., C. M. Biradar, P. Noojipady, V. Dheeravath, Y. Li, M. Velpuri, M. Gumma, O. R. P. Gangalakunta, H. Turral, X. Cai, J. Vithanage, M. A. Schull, and R. Dutta 2009. "Global Irrigated Area Map (GIAM), Derived from Remote Sensing, for the End of the Last Millennium." *International Journal of Remote Sensing* 30 (14): 3679–3733. doi:10.1080/01431160802698919.

Zhao, M., F. A. Heinsch, R. R. Nemani, and S. W. Running. 2005. "Improvements of the MODIS Terrestrial Gross and Net Primary Production Global Dataset." *Remote Sensing of Environment* 95: 164–176.