1   **Remote sensing based yield monitoring: application to winter wheat in**

2   **United States and Ukraine**

3   B. Franch[1,2], E.F. Vermote[2], S. Skakun[1,2], J.C. Roger[1,2], I. Becker-Reshef[1], E. Murphy[1,2],

4   C. Justice[1]

5

6   [1]University of Maryland, Dept. of Geographical Sciences

7   [2]Terrestrial Information Systems Laboratory, NASA Goddard Space Flight Center

8

9   **ABSTRACT**

10   Accurate and timely crop yield forecasts are critical for making informed agricultural

11   policies and investments, as well as increasing market efficiency and stability. Earth

12   observation data from space can contribute to agricultural monitoring, including crop yield

13   assessment and forecasting. In this study, we present a new crop yield model based on the

14   Difference Vegetation Index (DVI) extracted from Moderate Resolution Imaging

15   Spectroradiometer (MODIS) data at 1km resolution and the un-mixing of DVI at coarse

16   resolution to a pure wheat signal (100% of wheat within the pixel). The model was applied

17   to estimate the national and subnational winter wheat yield in the United States and Ukraine

18   from 2001 to 2017. The model at the subnational level shows very good performance for

19   both countries with a coefficient of determination higher than 0.7 and a root mean square

20   error (RMSE) of lower than 0.6 t/ha (15-18%). At the national level for the United States

21   (US) and Ukraine the model provides a strong coefficient of determination of 0.77 and

22   0.86, respectively, which demonstrates good performance at this scale. The model was also

23   able to capture low winter wheat yields during years with extreme weather events, for

24 example 2002 in US and 2003 in Ukraine. The RMSE of the model for the US at the

25 national scale is 0.16 t/ha (5.5%) while for Ukraine it is 0.27 t/ha (8.4%).

26

27 **1. INTRODUCTION**

28 Agriculture faces major challenges in this century. According to the Food and Agriculture

29 Organization of the United Nations (FAO), global food production will need to increase by

30 70% by 2050 in order to meet the growing global demand [1]. However, despite

31 tremendous improvements in technology and crop yield potential, food production remains

32 highly dependent on climate [2]. Since Earth Observation (EO) satellites became available,

33 it was recognized that the technology held considerable promise for agricultural monitoring

34 [3]. The NASA Large Area Crop Inventory Experiment (LACIE) and Agriculture and

35 Resources Inventory Surveys Through Aerospace Remote Sensing (AgriSTARS)

36 programs made significant advances in crop monitoring but were seriously constrained by

37 satellite data availability. With recent advances in sensing technology, observation

38 frequency, data products, computational analytics models and information technology, we

39 are now better able to forecast and estimate crop production. Within-season crop yield

40 assessment and forecasting is one of the major components of crop production monitoring.

41 Timely and accurate information on crop yields at global, national, and regional scales is

42 extremely important for informing a range of agricultural and food security decisions. For

43 example, at the global and national level, timely information about crop development and

44 production prospects in food exporting countries can help to enable actions to mitigate the

45 effects of any food supply crisis (URL 1).

46 Observations from the NASA Earth Observing System (EOS) MODIS sensors have several

47 of the key qualities needed for global crop yield monitoring such as global coverage at

48 coarse spatial resolution (250m), an excellent temporal resolution of two images every day

49 that combined, minimize the data gaps associated with clouds and a suite of validated

50 products. A good temporal resolution is required when working on agriculture applications

51 given that the crop phenology and conditions (water supply, pests, environmental…) can

52 change very quickly. Additionally, the MODIS dataset of 17+ years allows us to build a

53 strong empirical model based on a large number of statistics and covering a wide range of

54 crop conditions relative to other available satellite data. Generally, quantitative crop yield

55 forecast regression models are based on EO-derived vegetation indices (e.g., maximum

56 Normalized Difference Vegetation Index (NDVI) for the season [4] or temporal NDVI

57 integration [5]–[9], senescence rate [10] or combining land surface temperature (LST) and

58 NDVI [11]).

59 However, the main disadvantage of coarse to moderate resolution sensors is the spatial

60 resolution that often mixes in a given pixel, signals from different land cover types and

61 crops. Stratifying a region into different crop types (commonly termed as crop masking) is

62 an important step in developing EO-based yield models [12]. Such masks enable the

63 isolation of the remotely sensed, crop-specific signal throughout the growing season,

64 reducing the noise on the signal from other land cover or crop types. In the United States

65 (US), the US Department of Agriculture (USDA) generates a yearly national Cropland

66 Data Layer (CDL) since 2007 [13] and Canada provides yearly national Annual Crop

67 Inventory Maps (ACIM) since 2009 [14]. However, these masks are provided at the end of

68 the growing season and there are no available crop type masks for other countries. As a

69 consequence, generally EO-based crop yield models use static cropland or crop type masks.

70 Next, we briefly describe some examples of studies found in the literature that use coarse

71 resolution EO data to monitor crop yields, how they deal with the signal mixing and which

72 masks they use as reference.

73 [4] developed an empirical but generalized crop yield model based on MODIS Climate

74 Modeling Grid (CMG, 0.05 degree spatial resolution) data and applied it to Kansas and

75 Ukraine. The model, that was calibrated in Kansas, is based on the relationship between

76 the NDVI value at the peak for the season and the final yield value, corrected for the

77 average purity or the percentage of crop of the 5% purest pixels within the area studied.

78 They used static crop type masks by aggregating the USDA CDL of 2007 to 0.05 degrees

79 spatial resolution and developed their own mask for Ukraine. Later, [15] improved the

80 temporal aspect of the model by including Growing Degree Days (GDD) information. The

81 model was applied satisfactorily (errors lower than 10%) to monitor wheat yield at the

82 national level over the US, Ukraine and China. It also showed good performance when

83 applied to the AVHRR Long Term Data Record (LTDR) [16]. However, this model was

84 designed to work primarily at the national level.

85 [17] developed an empirical model to forecast wheat yield over Europe. They used SPOT-

86 VEGETATION products of NDVI and Fraction of Absorbed Photosynthetically Active

87 Radiation (FAPAR) at 1 km spatial resolution with errors ranging from 10-17% depending

88 on the location. They divided Europe into different agro-climatic zones and aggregated

89 those pixels using as a reference the arable land fraction from the Corine Land Cover

90 (CLC, [18]) and estimated the weighted average of pixels with arable land fraction land

91 higher than 50%. Later, [19] analyzed the correlation between EO indicators (FAPAR from

92   SPOT-VEGETATION at 1km spatial resolution) and wheat, barley and grain maize yields

93   in Europe. In this case they divided Europe into different sub-national areas depending on

94   the available statistics for each country. In this work, they also analyzed the weighted

95   average of pixels within the 90th percentile arable lands fraction for each administrative

96   unit, using the CLC as a reference. They found that crops growing under moderately or

97   strongly water-limited conditions present a high inter-annual variability of leaf area

98   development and senescence that influences the final yield, which is observable by EO

99   systems. However, in agro-climatic areas where crop growth is not frequently water-

100   limited (e.g. irrigated crops), yield variations are explained by a complex interaction of

101   different factors not necessarily related to the crop leaf area dynamics, and, therefore,

102   weather indicators are needed to complement the FAPAR time series.

103   [20] assessed the feasibility and relative efficiency of using NDVI MODIS data to forecast

104   winter wheat yield at *oblast* (sub-national) level in Ukraine. They spatially aggregated each

105   oblast using a cropland map as a reference (Rainfed croplands class) that was extracted

106   from the ESA GlobCover map at the 300 m resolution. They concluded that NDVI values

107   2-3 months prior to harvest allowed reliable yield forecasts at the *oblast* level with RMSE

108   ranging from 0.5 to 0.8 t/ha.

109   [21] analyzed the correlation between the Evaporative Stress Index (ESI) from MODIS at

110   1km spatial resolution and yield for major crops grown in Brazil. They aggregated those

111   pixels classified as majority agricultural land using a land use map of Brazil, generated by

112   the Brazilian Geographical and Statistical Institute (IBGE). Their results showed better

113   correlation than other indices such as LAI for most crops and regions. Additionally, in the

114   case of soybeans, the yields showed earlier response (10 to 25 days) to extreme events than

115   LAI.

116   In this study, we present a new empirical model based on remote sensing data at coarse

117   spatial resolution (MODIS aggregated to 1km). This model is based on the wheat signal

118   un-mixing from the signal of other surfaces surrounding the wheat fields by using yearly

119   crop type masks. Additionally, based on the analysis of the un-mixed time series of the

120   wheat we use three different regressors that we calibrate for each administrative unit

121   against the historical yield statistics. The main goal of developing this model is to make it

122   sensitive to extreme weather events and applicable at the subnational level. This model is

123   based on the Difference Vegetation Index (DVI) [22] that is estimated by the simple

124   difference between surface reflectance in the red and the Near Infrared (NIR) spectral

125   bands. The following aspects are incorporated into the proposed model:

126   1. We use BRDF-corrected daily 1 km MODIS data derived from both Terra and Aqua

127   satellites.

128   2. We un-mix wheat surface reflectance's to estimate the signal as if the MODIS

129   pixels were 100% wheat. We use a satellite-derived amplitude of the peak DVI

130   along with the time period that the peak remains (Length) and Evaporative Fraction

131   (EF) to correlate with winter wheat yields. The EF is a parameter related to the

132   evapotranspiration and is further detailed in section 3.2.

133   In the next section (2.) we describe the study area and the data considered. Then, we

134   describe the method proposed and how we derive the different input parameters (Section

135   3.). Finally, we show the results both at the subnational and the national levels for the two

136   countries considered (Section 4.).

137

2. STUDY AREA AND DATA DESCRIPTION

2.1. Study area

The study is carried out over the US and Ukraine from 2001 until 2017. The US is one of

the main producers and exporters of wheat globally. In 2016 the US was the #1 wheat

exporting country, shipping 14.8% of global wheat exports (URL 2). Wheat is produced in

almost every state in the United States and winter wheat varieties dominate US production,

representing between 70-80% of the total wheat production. The winter wheat is planted

in the Fall and harvested during June-July. Generally, wheat is rain-fed and just 7% of the

national production is irrigated. The main wheat class is Hard Red Winter Wheat, which is

grown primarily in the Great Plains, with Kansas being the largest producing state.

Ukraine is another critical player in the global wheat market. In 2016, Ukraine ranked #6

by exporting 7.2% of global wheat exports (URL 2). Wheat is grown all across the country,

although the central and southern regions are the key growing areas (Forest-Steppe and

Steppe zones in Figure 1). About 95% of Ukraine's wheat production is winter wheat,

planted in the Fall and harvested during June-July of the following year. Generally, wheat

is rain-fed and is not irrigated. Ukraine produces mostly Hard Red Winter Wheat.

Tables 1 and 2 show the primary characteristics of the main climatic zones in the US and

Ukraine where most of the winter wheat fields are planted and the average yields for each

area using as a reference 2016.

Table 1. Main climate zones in the US (classification from Vocke & Ali (2013) and yield values

from NASS statistics).

Zone	States	Climate conditions	Major Type of winter wheat	Average yield (2016), t/ha
North Central	MT, ND, SD, MN	Humid continental	Hard Red	3.6
Southern Plains	OK, TX	Temperate climate	Hard Red	2.4
Central Plains	KS, CO	Temperate climate	Hard Red	3.5
Northern Plains	NE, WY	Humid continental	Hard Red	3.0
Northwest	WA, OR, ID	Oceanic	Soft White	5.0

160

161   Table 2. Main climate zones in Ukraine.

Zone	Conditions	Average yield (2016), t/ha
Plane-Polissya	Mixed forest zone in the north, 26% of the entire Ukrainian territory	4.36
Forest-Steppe	Mainly cultivated area in the center (34%)	4.88
Steppe	Intensive cultivated area (40%)	3.64

162
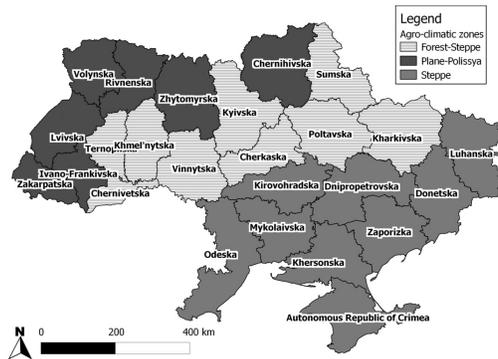
163

164   Figure 1. Agro-climatic zones of Ukraine

165

166

167   **2.2. Crop statistics on winter wheat**

168   In each country we run the wheat yield model for each Administrative Unit (AU). Thus,

169   we work at different spatial scales depending on the data availability. In the case of the US,

170   we work at the county level (average area of 258,000 ha), while in the case of Ukraine we

171   work at the *oblast* level (average area of 2,414,000 ha).

172   For the US, we use the official archive of county-level statistics on yield, area harvested,

173   and production that are available from the USDA National Agricultural Statistics Service

174   (NASS) Quick Stats database (http://www.nass.usda.gov/Quick_Stats/). The NASS crop

175   statistics are based on data obtained from multiple frame-based sample surveys of farm

176   operators, objective yield surveys, agribusinesses, shippers, processors and commercial

177   storage firms.

178   For Ukraine, *oblast*-level crop statistics are obtained from the State Statistics Service of

179   Ukraine (SSS) (http://www.ukrstat.gov.ua) for winter wheat area harvested and yield.

180   These official statistics are based on farm surveys collected from all the agricultural

181   enterprises (large-scale farms that produce commodities exclusively for sale).

182

183   **2.3. Crop type masks**

184   For the US, a national winter wheat mask is available from the CDL produced by NASS

185   from 2008 to 2016. The CDL is a rasterized land cover map using field level training data

186   from extensive ground surveys, farmer reports provided to the US Farm Service Agency

187 (FSA), and remotely sensed data from Landsat Thematic Mapper (TM), Landsat Enhanced

188 Thematic Mapper (ETM+) and Advanced Wide Field Sensor (AWiFS). These data are

189 used in a decision tree classifier in order to produce a land cover classification that

190 distinguishes between different crop types, including winter wheat [13], [24]. In this study,

191 yearly CDL layers are used to identify winter wheat growing areas by estimating the

192 percentage of 30m CDL pixels within each MODIS 1km pixel. At 1km spatial resolution

193 MODIS pixels mix the signal from wheat with the signal from other surfaces surrounding

194 the wheat fields. In order to isolate the wheat signal through the un-mixing process (see

195 section 3.3), we used the CDL masks to identify not only the wheat but also the other

196 surfaces and how much percentage of each class is observed by each pixel. Therefore,

197 looking at the main land cover classes surrounding the wheat fields in the US, we consider

198 a total of eight different cover classes that are included in the un-mixing method: winter

199 wheat, spring wheat, corn, soybean, alfalfa, potato, grassland and forest. Before 2008 the

200 CDL was not produced at the national scale and has gaps for some states. To generate

201 yearly winter wheat masks for US from 2001 to 2007 and for Ukraine for the 2001–2016

202 period, we used MODIS-derived 8-day composited NDVI values at 250 m spatial

203 resolution [25] and an automatic phenology-based approach for winter wheat mapping

204 [26]. This approach showed an accuracy of >90% for US and Ukraine, and good

205 correspondence to official statistics with an average coefficient of determination >0.85.

206 Additionally for Ukraine, we extracted static grassland and forest masks from the 30 m

207 land cover map for Ukraine for 2010 [27], [28]. The reported classification accuracies are

208 97% and 90% for forest and grassland, respectively. Note that these classes are just

209 considered when they are surrounding the wheat fields, that is, when their signal is mixed

210 with wheat in order to un-mix it. Thus, by using static masks of grassland and forests in

211 Ukraine we are assuming that these surfaces surrounding the wheat fields have not changed

212 considerably during the time period analyzed.

213

214 **2.4. MODIS surface reflectance time-series**

215 This study is based on MODIS daily surface reflectance Collection 6 data

216 (M{OY}D09GQ) distributed by the Land Processes Distributed Active Archive Center (LP

217 DAAC, https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table) which

218 are gridded in the sinusoidal projection at 250m resolution. Additionally, we used the

219 product M{OY}D09GA to extract the geometry of observation/illumination of each image.

220 Since the nominal 250 m MODIS resolution decreases for the off-nadir observations and

221 due to inaccurate registration [29], we re-scaled the 250m surface reflectance to 1km spatial

222 resolution to mitigate that effect by aggregating 4x4 pixels.

223

224 **3. METHODS**

225 **3.1. General workflow**

226 The diagram in Figure 2 shows the workflow followed in this study. The model is based

227 on three different regressors (blue boxes in the diagram): the amplitude of the peak, the

228 time period that the peak remains for and the average of the Evaporative Fraction (EF)

229 during the thirty-day period following the peak. In order to determine these parameters, the

230 inputs needed (yellow boxes in the diagram) are the MODIS Surface Reflectance product,

231 the MODIS Land Surface Temperature product, the CDL crop mask in the US and official

232   statistics for model calibration and validation. In the following subsections we describe

233   more detail about the methods (green boxes in the diagram).

234

235
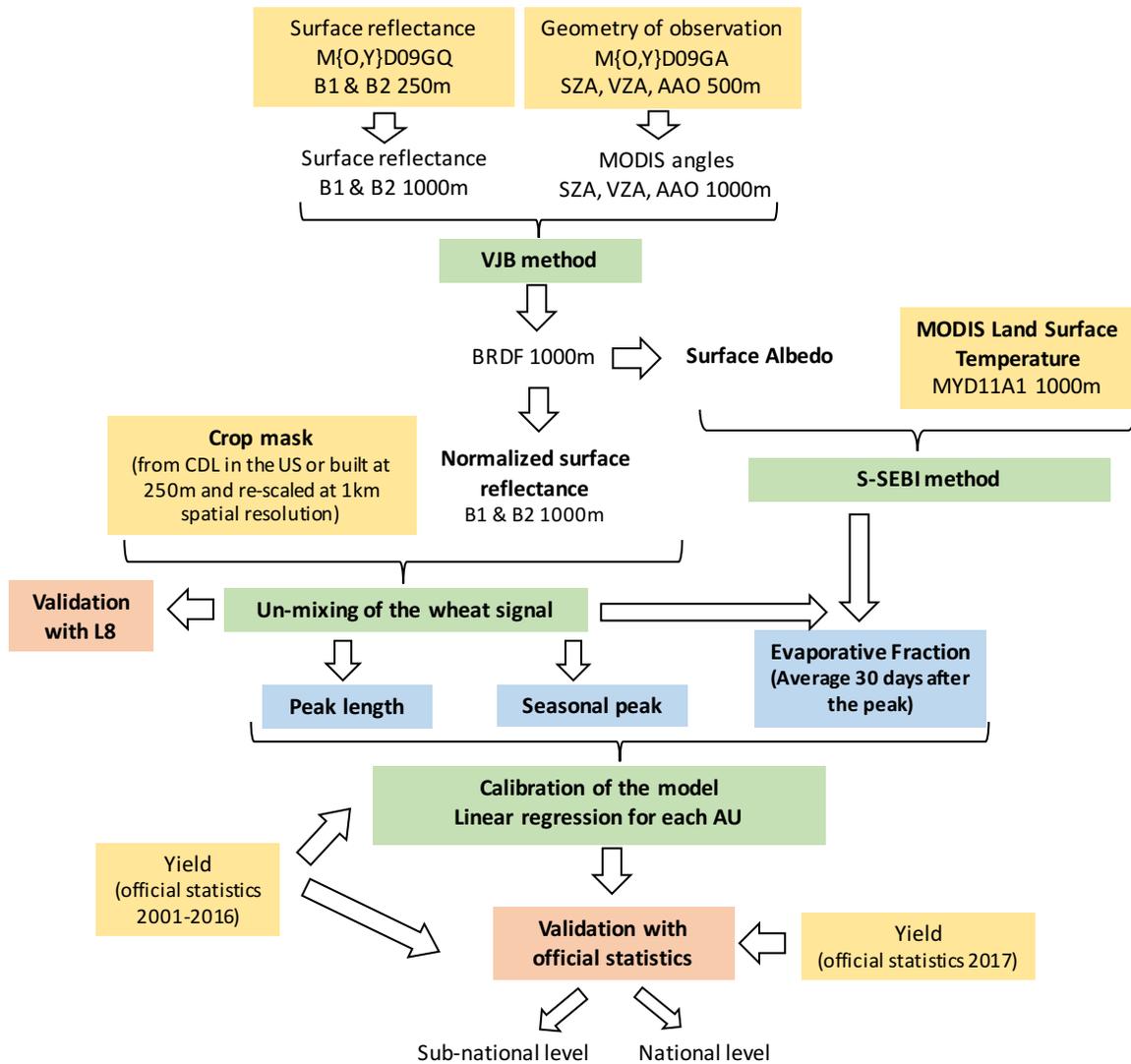
236

237   Figure 2. The workflow followed in this study.

238

239   **3.2.   Surface reflectance normalization and albedo estimation**

240   The wide swath MODIS sensor allows for near global coverage of the Earth every day.

241   However, it has a 16-day repeat cycle, which means that every day the geometry of

242   observation is different and can include View Zenith Angles (θ_v) of up to 65 degrees. As a

243   consequence, the surface reflectance that is defined for a given geometry of observation-

244   illumination has different values every day. When working with normalized vegetation

245   indices such as the NDVI, the geometry effects are minimized, but that is not the case when

246   working with non-normalized parameters such as the DVI. In order to normalize the BRDF

247   effects on the surface reflectance, we used the VJB method [30], [31]. This method uses

248   longer compositing periods (5 years in [30]) than the MCD43 product (16 days) [32] which

249   reduces the noise in the normalized reflectance time series [29]. In this study, we derive

250   the nadir BRDF parameters at 1km spatial resolution using the most recent five years (2012

251   to 2016). By using the daily surface reflectance (from both Aqua and Terra) and its angular

252   conditions during the five-year period considered, we derive the variables that define the

253   BRDF shape (V and R in Eq. 1) using the approach used by [31]. Then we apply Eq 1 for

254   deriving the normalized surface reflectance (ρ^N) in MODIS band 1 and 2 (1) and we also

255   use the BRDF coefficients for estimating the surface albedo.

256

257
$$\rho^N\left(45,0,0\right) = \rho\left(\theta_s,\theta_v,\phi\right)\frac{1+VF_1\left(45,0,0\right)+RF_2\left(45,0,0\right)}{1+VF_1\left(\theta_s,\theta_v,\phi\right)+RF_2\left(\theta_s,\theta_v,\phi\right)} \qquad (1)$$

258

259   where ρ is the directional surface reflectance, θ_s is the sun zenith angle, θ_v is the view zenith

260   angle, ϕ is the relative azimuth angle, F_1 is the volume scattering kernel, based on the Ross-

261   Thick function derived by [33] but corrected for the Hot-Spot process proposed by [34], F_2

262 is the geometric kernel, based on the Li-sparse model [35] but considering the reciprocal

263 form given by [36], *V* represents the volume BRDF parameter since it is linked to the

264 Volume kernel and *R* represents the roughness BRDF parameter since it is linked to the

265 geometric kernel. These parameters (V and R) represent the shape of the BRDF.

266 In order to estimate the surface albedo, that is needed to estimate the EF, we follow the

267 same methodology to derive the surface albedo as the official MCD43 MODIS product [32].

268 Downwelling flux may be written as the sum of a direct component and a diffuse component.

269 Black-sky albedo (α_{bs}) or directional-hemispherical reflectance (DHR) is defined as albedo in

270 the absence of a diffuse component and is a function of solar zenith angle. White-sky albedo

271 (α_{ws}) or bihemispherical reflectance is defined as albedo in the absence of a direct component

272 when the diffuse component is isotropic. It is independent of the geometry of illumination-

273 observation. Therefore, the albedo can be written by integrals of the BRDF model through the

274 black-sky albedo and the white-sky albedo.

275
$$\alpha_{bs}(\theta_i,\lambda)=\sum_k f_k(\lambda)h_k(\theta_i)$$
(2)

276
$$\alpha_{ws}(\lambda)=\sum_k f_k(\lambda)H_k$$
(3)

277 where $f_k(\lambda)$ are the BRDF model parameters and

278
$$h_k(\theta_i)=\int_0^{2\pi}\int_0^{\pi/2}k_k(\theta_i,\theta_v,\phi)\sin\theta_v\cos\theta_v d\theta_v d\phi$$
(4)

279
$$H_k=2\int_0^{\pi/2}h_k(\theta_i)\sin\theta_i\cos\theta_i d\theta_i$$
(5)

280 where $k_k(\theta_i, \theta_v, \phi)$ are the BRDF model functions of the observation geometry (F_1 and F_2 in

281 equation (1)).

282 The albedo, $\alpha(\theta,\lambda)$, under actual atmospheric conditions can also be modeled quite accurately

283 as an interpolation between the black-sky (direct beam) albedo and white-sky (completely

284 diffuse) albedo as a function of the fraction of diffuse skylight $S(\theta,\tau(\lambda))$, which, in turn, is a

285 function of optical depth τ [37], [38].

286 $$\alpha(\theta,\lambda) = \left[1 - S(\theta,\tau(\lambda))\right]\alpha_{bs}(\theta,\lambda) + S(\theta,\tau(\lambda))\alpha_{ws}(\lambda) \tag{6}$$

287 We simulate the diffuse skylight with 6S [39] using as input the MODIS Aerosol Optical

288 Thickness (AOT) at 550nm product, which is included in the Climate Modeling Grid

289 (CMG) product.

290

291 **3.3. Un-mixing of the wheat signal**

292 When working with medium to coarse spatial resolution sensors, the signal retrieved by

293 each pixel generally mixes different land cover classes. Therefore, crop yield modeling

294 requires availability of crop-specific masks, so that only the remotely sensed signal from

295 the crops in question will be considered. In order to work with the purest wheat signal

296 possible, in this work we present a method based on the un-mixing of the signal at 100%

297 wheat purity. For each AU and for a given date, the total DVI signal from each pixel, i, can

298 be written as the sum of the DVI signal from the wheat (DVI_{wheat}) multiplied by the

299 percentage of wheat within the pixel or wheat purity (Wpct) and the DVI from other

300 surfaces within the pixel (DVI_{others}) multiplied by the remaining percentage (1- Wpct) (1).

301 $$DVI_i = DVI_{i,wheat} \cdot Wpct_i + DVI_{i,others} \cdot (1 - Wpct_i) \tag{7}$$

302   Which has the structure of a linear regression function and can be solved with the Ordinary

303   Least Squares (OLS) method. Figure 3 shows an example of the linear regression between

304   the DVI_i and the wheat purity, $Wpct_i$, (black dots) for the Day of the Year (DOY) 79 of

305   2016 for Harper County (Kansas).
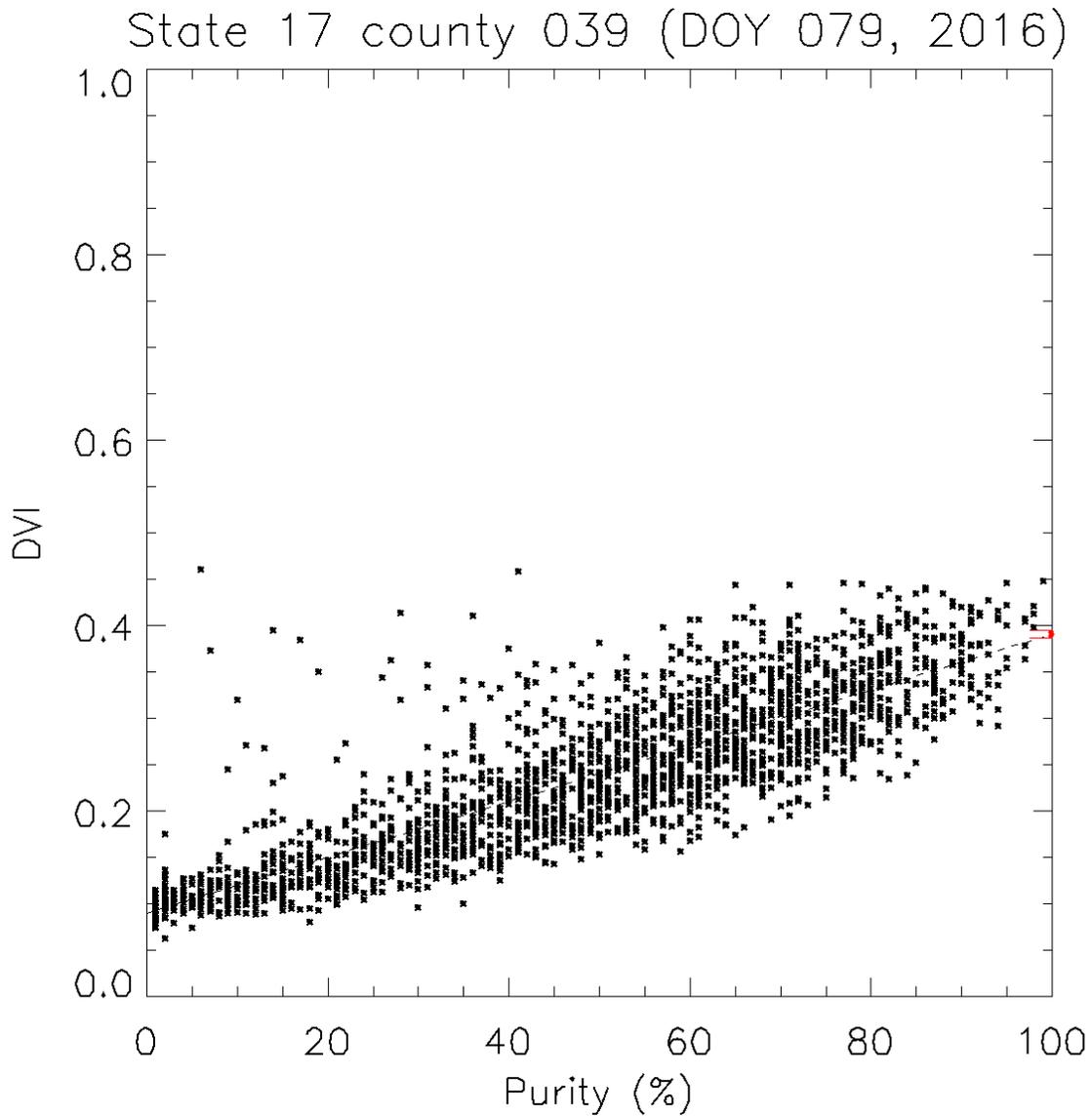
306

307   Figure 3. Linear regression of the DVI_i versus the wheat purity ($Wpct_i$) through the Harper County

308   for DOY 79 of 2016.

309

310   In the case of the US, where the CDL product provides masks for other crops, we can

311   take advantage of such information to improve the un-mixing of the signal. As described

312   in the previous section, after studying the main classes that generally surround the US

313   wheat fields, we consider a total of 8 different classes plus the "others" contribution. The

314   eight classes considered are: winter wheat (#1), spring wheat (#2), corn (#3), soybean (#4),

315   alfalfa (#5), potato (#6), grassland (#7) and forest (#8) with crop purities (percentage of a

316   given crop area within the pixel) Cpct (corn). Then, equation 7 is written as:

317

318    $$DVI_i = \sum_{Y=1}^{8} DVI_{i,Y} \cdot Ypct_i + DVI_{i,others} \cdot (1 - \sum_{Y=1}^{8} Ypct_i) \quad (8)$$

319

320   where Y is the class number, $Ypct_i$ is the purity of class Y (or percentage of class Y area

321   within the pixel). This equation also has the structure of a linear regression function.

322   This un-mixing method is applied to each AU and day, isolating in each case the DVI signal

323   coming from wheat. Based on the time series of the DVI_{wheat}, we estimate the amplitude of

324   the peak value of this variable and the time period that this peak remains (hereafter referred

325   as length), defined as the number of days when the DVI_{wheat} is higher than 80% of the

326   amplitude value.

327

328   **3.4.   Evaporative Fraction estimation**

329   Any change in the water resources of the crop takes some time to show an effect on the

330   biomass, that is directly linked to the vegetation indices. Therefore, any water stress

331   condition happening during or after the peak might not be accounted for when just using

332   the amplitude and length. However, these changes can be detected earlier when using

333   thermal indices [21]. One of the key factors that drive the land surface energy process is

334   evapotranspiration (ET). ET is defined as the flux of water evaporated at the earth-

335   atmosphere interface (from soil, water bodies and interception) and transpired by

336   vegetation through stomata in their leaves as a consequence of the photosynthetic

337   processes. There are several ET models and products [40]–[46]. However, they generally

338   rely on the estimation of several parameters such as net radiation, the sensible and latent

339   heat flux and the ground flux. In this study, we examine the use of the Evaporative Fraction

340   (EF) as one of the inputs to the proposed yield model. The EF describes the ratio of

341   evapotranspiration to available total energy (equation 9).

$$EF = \frac{LET}{LET + H} = \frac{LET}{R_n - G}$$

342   (9)

343   where LET is the latent heat flux, H is the sensible heat flux, R_n is the net radiation and G

344   is the soil heat flux. In this study, EF is derived based on the Simplified Surface Energy

345   Balance Index (S-SEBI) method developed by [44]. The surface temperature and

346   reflectance of areas with constant atmospheric forcing are correlated and the relationships

347   can be applied to determine the effective land surface properties [40], [47]. The S-SEBI

348   method uses as reference the LST vs albedo diagrams to determine the dry and the wet

349   conditions of the area considered. The dry conditions are defined by a baseline

350   representative of the maximum sensible heat with a reflectance dependent maximum

351   temperature. Meanwhile, the wet conditions are defined by a baseline of the maximum

352   potential evapotranspiration with a reflectance dependent minimum temperature, hence the

353   sensible and latent heat fluxes are partitioned according to the actual surface temperature.

354   The assumption of this method is that global radiation and atmospheric conditions tend to

355   be constant and that there are a sufficient number of wet and dry pixels to determine and

356   right wet and dry edge [44]. A detailed analysis on these baselines estimation uncertainty

357   can be found in [48]. The EF is derived using an approximation according to [44], [49],

358   [50]:

359

360    $$EF = \frac{T_H - T_S}{T_H - T_{LE}} = \frac{a_H \alpha + b_H - T_S}{(a_H - a_{LE})\alpha + (b_H - b_{LE})}$$ (10)

361

362   where T_H is land surface temperature for dry pixels; T_{LE} is land surface temperature for wet

363   pixels; T_s is land surface temperature; α is surface albedo; $a_H,$ a_{LE} are the slopes of the line

364   of, respectively, the high and low temperatures as a function of surface albebo; $b_H,$ b_{LE} are

365   the intercepts of the line of, respectively, the high and low temperatures as a function of

366   surface albedo.

367   In this study, the LST is extracted from the official MODIS (MYD11) LST product [51].

368   Note that the EF is estimated using the Aqua satellite whose overpass time is during the

369   afternoon, when higher water stress can be detected. Section 3.2 describes the method

370   followed to derive the surface albedo. We derive the LST vs albedo diagram parameters'

371   for each day at the state level in the US and at the *oblast* level for Ukraine. Note that in the

372   case of the US, the LST vs albedo diagrams are built at state level (instead of county level)

373   to ensure having enough pixels with wide variability of moisture conditions. Figure 4

374   shows an example of the LST-albedo diagram in Kansas. Then, for each pixel, Eq. 10 is

375   implemented, providing a pixel-based EF estimation. Finally, at the county level we apply

376   the same un-mixing procedure as the DVI in order to isolate the EF for the wheat. This

377   parameter is considered in the model by estimating its average during the 30 days after the

378   seasonal peak (grain filling period). We consider this period to account for any water stress

379   conditions happening during or after the peak and that will have a negative impact on the

380   final yield [52], [53].

381
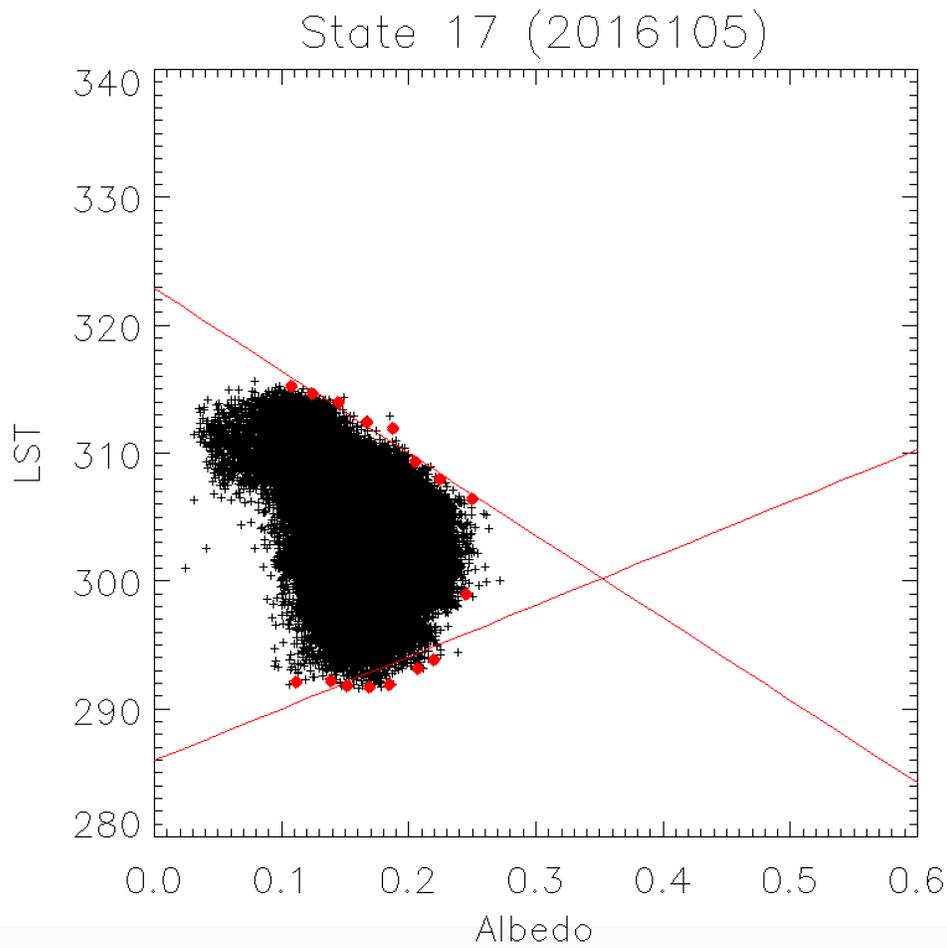
382

383   Figure 4. Example of a LST-albedo diagram in Kansas on April 15th, 2016. The red lines

384   represent the maximum sensible heat (dry boundary, top line) and the maximum potential

385   evapotranspiration (wet boundary, bottom line). Red dots represent the data considered to

386   derive the red lines.

387

388   **3.5.   Evaluation of the un-mixing method**

389   We test the performance of the un-mixing of the wheat signal by applying the un-mixing

390   method to all Landsat 8 OLI cloud-free data available from 2013 to 2016 in three different

391   counties in the US: Harper (Kansas, 17 images), Walla Walla (Washington, 23 images)

392   and Bingham (Idaho, 14 images). The validation of the method is just focused on Landsat

393   data to avoid any impact of spectral adjustment between two sensors or any error associated

394   with angular effects correction. Thus, first Landsat data are transformed to MODIS-like

395   images by upscaling the spatial resolution of Landsat to 1km and re-projecting them to the

396   sinusoidal projection. Then, using the same CDL crop masks that we use for MODIS, we

397   apply the un-mixing method. Finally, we validate the un-mixing of the wheat signal against

398   the average throughout the county of all the 100% purity wheat pixels at the original

399   Landsat spatial resolution.

400

401   **3.6.   Calibration of the yield model**

402   For each AU, a separate linear regression was built with different combinations of

403   regressors (Equations 11-16).

404     $yield = a_{10} + a_{11} \cdot A$ (11)

405     $yield = a_{20} + a_{21} \cdot L$ (12)

406     $yield = a_{30} + a_{31} \cdot EF$ (13)

407     $yield = a_{40} + a_{41} \cdot A + a_{42} \cdot L$ (14)

408     $yield = a_{50} + a_{51} \cdot A + a_{52} \cdot EF$ (15)

409     $yield = a_{60} + a_{61} \cdot A + a_{62} \cdot L + a_{63} \cdot EF$ (16)

410   A *p*-value for each regressor was calculated which provides a test of the null hypothesis

411   that a coefficient for this regressor is equal to zero, i.e. the regressor has no effect on the

412   dependent variable (crop yield in our case). A small *p*-value indicates that the

413   null hypothesis can be rejected, meaning that there is a statistically significant relationship

414   between the regressor and dependent variable. We used this statistic index to decide in each

415   AU which variables were significant (by using a threshold of 0.2 in p-value) and which

416   equation to consider (lowest *p*-value for each regressor).

417   There are several winter wheat varieties planted in the US (over 30,000) distributed over

418   different regions with diverse climates, soils, treated with different fertilizers and

419   susceptible to different diseases. Thus, specific modelling of such characteristics would be

420   complex, data demanding, and leading to highly parametrized models. Therefore, our

421   approach was to empirically calibrate the model at the county level assuming that

422   homogeneity in the above-mentioned parameters. Additionally, given that we use 1km

423   spatial resolution data whose wheat signal is un-mixed, we need to ensure that the wheat

424   purity is sufficient to invert the model. At low wheat purities, the wheat signal is low and

425   it is competing with the signals from other surfaces. In these circumstances, the un-mixing

426   can lead to high errors. Therefore, the model is applied to every county with wheat purity

427   values higher than 40%. After testing other purities, this 40% threshold showed the best

428   performance or balance between number of counties considered (using a higher threshold

429   reduces the number of counties where the model can be applied) and noise in the

430   extrapolated data (when using lower thresholds). During the calibration process, none of

431   the three regressors considered in the proposed model (amplitude, length and EF) was able

432   to capture the late frost that mostly affected Kansas and Oklahoma in 2007. This late frost

433   happened during early April and is the only event with these characteristics that occurred

434   during the 17 years analyzed. As a consequence, there are not enough data to study the

435   inclusion of another variable in the model capable to respond to these extreme conditions.

436   Following the approach of [4], every county in these two states during 2007 is excluded

437   from the calibration and the validation processes.

438   In Ukraine we follow the same methodology as in the US, by calibrating the crop yield

439   model individually each AU (*oblast*).

440

441   **3.7.   Validation of the yield model**

442   Each country is validated both at national and subnational level from 2001 to 2017. Note

443   that the period 2001 to 2016 is used both for calibration and validation, while 2017 is just

444   used for validation.

445   For the Ukraine, the model can be applied to every oblast. Therefore, the national number

446   match the addition of the oblast-level statistics. However, this is not the case for the US,

447   where there are some counties excluded from the study because the maximum purity of the

448   wheat signal is lower than 40% in any 1 km pixel due to small field size. Additionally, in

449   the official county-level statistics there are some counties with no statistics or some

450   temporal gaps in other counties whose data is not publicly available due to privacy issues

451   but are accounted at national level.

452

453   **4. RESULTS**

454   **4.1. Evaluation of the un-mixing method**

455   Figure 5 shows the evaluation of the un-mixing method using Landsat 8 as a reference. The

456   results show a good agreement of the extrapolated signal with a high coefficient of

457   determination (>0.96) and errors of around 0.02. As an example of the impact and utility

458 of the un-mixing approach, Figure 6 plots the average of every Landsat 1km pixel with any

459 wheat against the same x-axis as in Figure 5. Given that these coarse resolution pixels mix

460 different surfaces' DVI, this variable does not show any variability throughout the year not

461 allowing any phenological analysis.

462

463

State 13 County 006

RMSE=0.028
Bias=−0.016
Stddev=0.023
y=0.949x
r2=0.96

Landsat 1km extrapolated

Landsat OLI
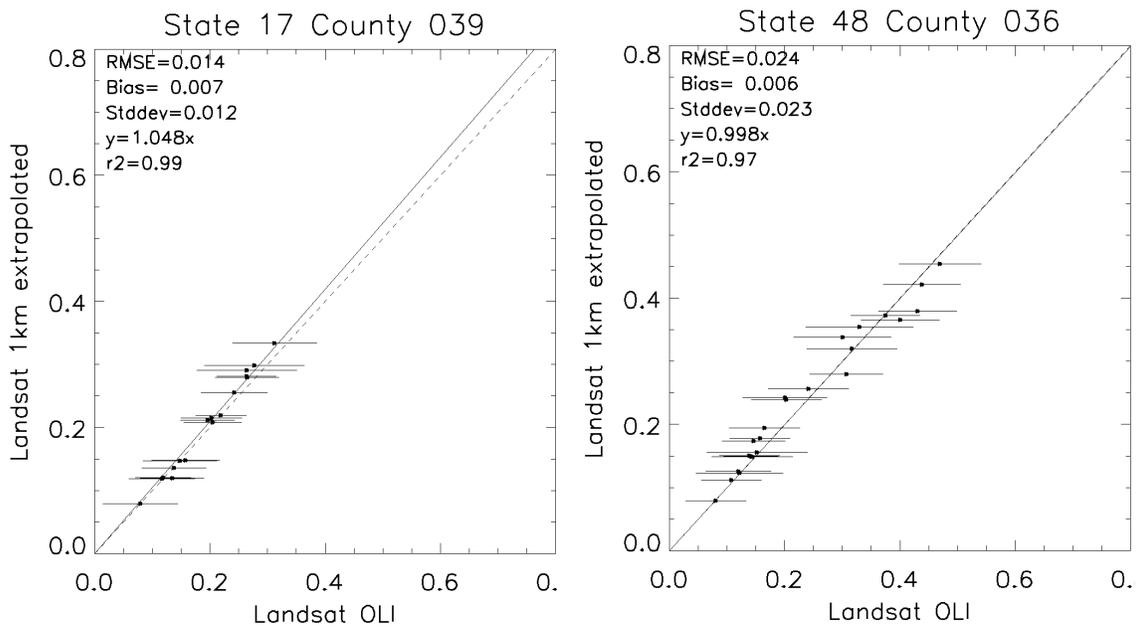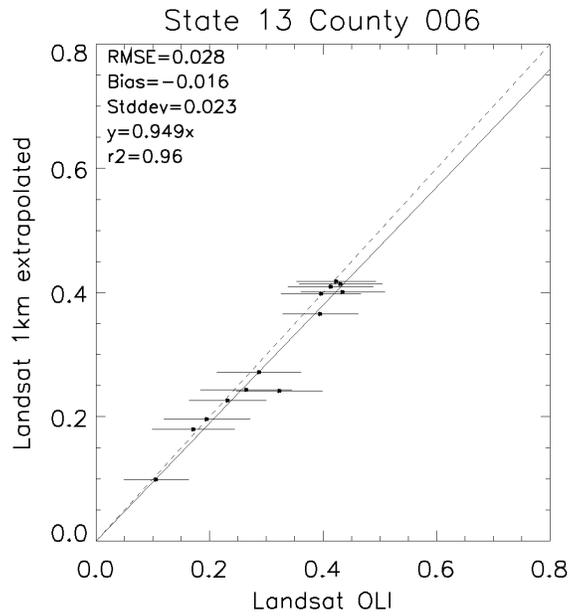
Figure 5. Evaluation of the un-mixing method using Landsat 8 as a reference for Harper County, Kansas (top left); Walla Walla County, Washington (top right); and Bingham County, Idaho (bottom). The error bars represent the standard deviation of the wheat pixels' average at the Landsat spatial resolution.

State 17 County 039 (NO extrapol)

RMSE=0.067
Bias=−0.013
Stddev=0.066
y=0.851x
r2=0.04

Landsat 1km NO extrapolated

Landsat OLI
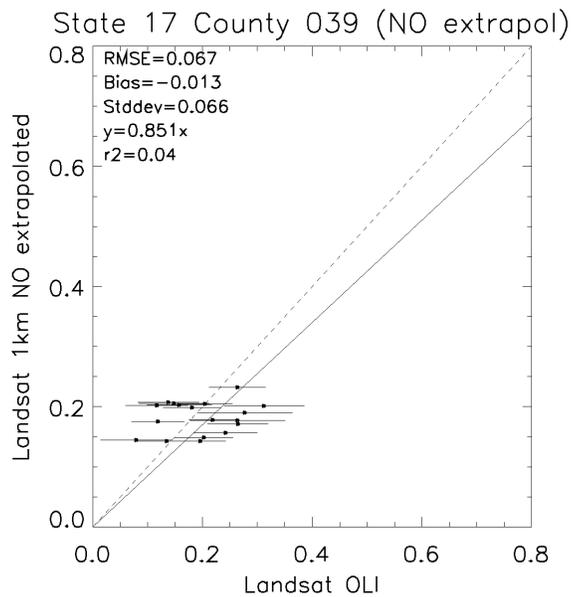
469   Figure 6. Example of the impact of not using the un-mixing method for Harper County, Kansas.

470   The error bars represent the standard deviation of the wheat pixels' average at the Landsat spatial

471   resolution.

472

473   **4.2.   Calibration of the yield model**

474   **4.2.1. The US**

475   Figure 7 shows the p-values associated with each input statistical significance in the

476   calibration equations of the model and the determination coefficient for each county.

477
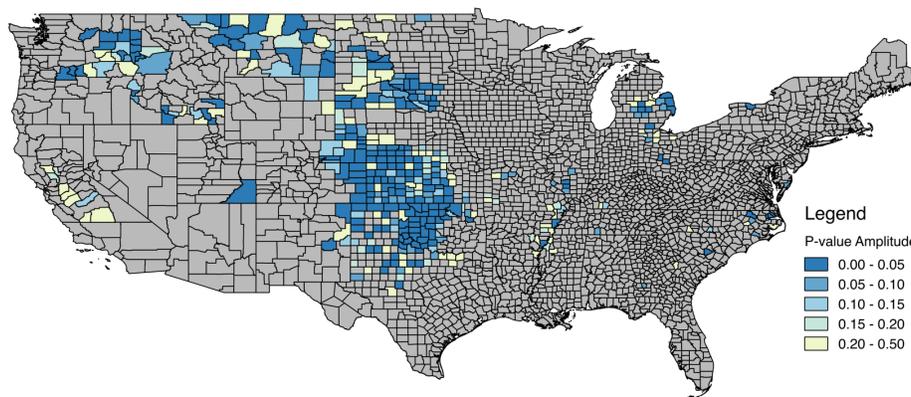
478

479   a)
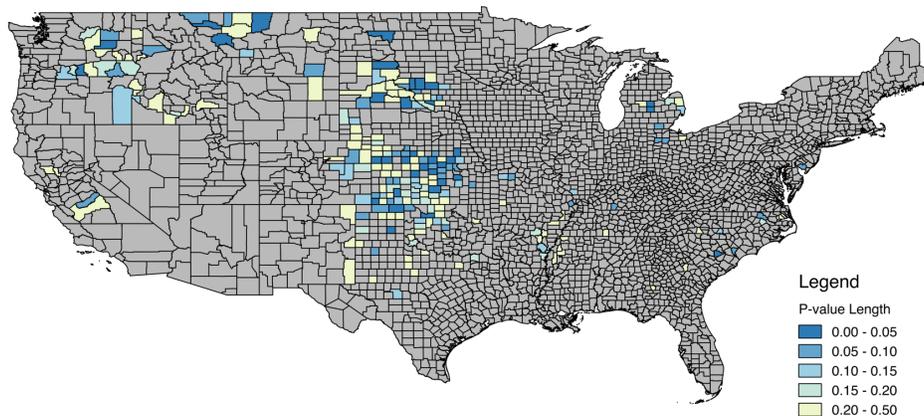
480

481   b)
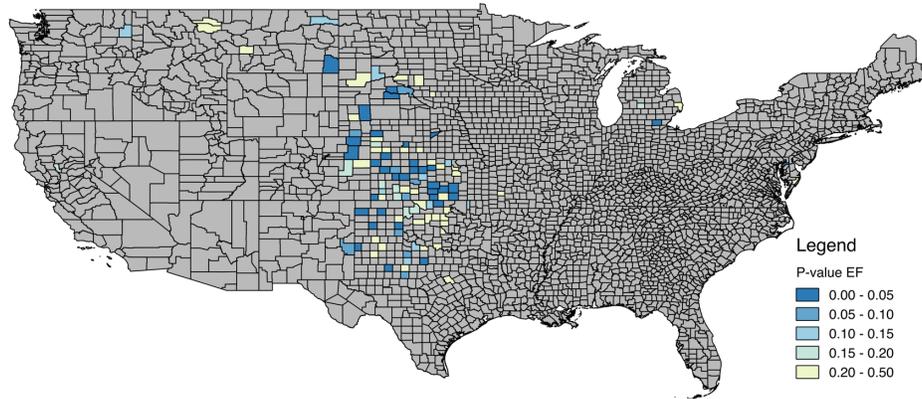
482

483   c)
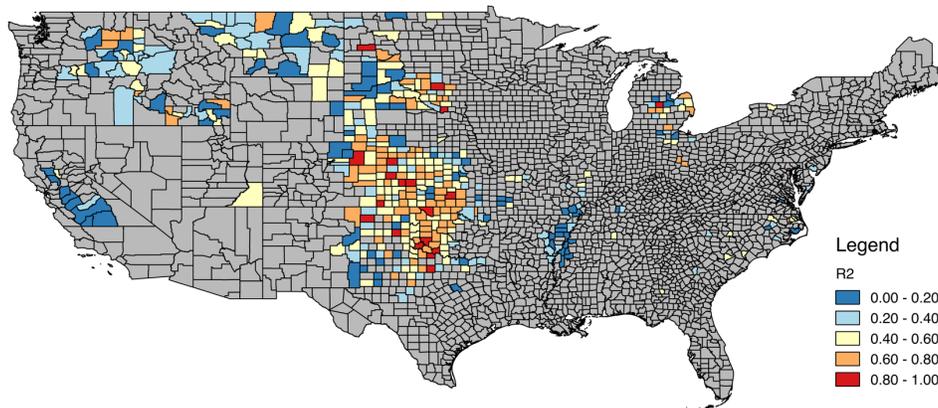
484

485   d)

486   Figure 7. Geographical distribution of *p*-values derived for a) the amplitude of the peak, b)

487   length of the peak and c) EF. Corresponding coefficients of determination for each county

488   is shown in subplot (d).

489

490   The amplitude is used in most counties and shows generally p-values lower than 0.05 in

491   the most important producing counties along the Great Plains, the North-West, North

492   Central regions and Michigan. However, it shows values higher than 0.10 in California.

493   The length is the second most used regressor and generally shows values lower than 0.05

494   in Kansas and South Dakota. Finally, the EF is the least used regressor and it is mostly

495   used in the Great Plains area where most counties show low p-values. The coefficient of

496   determination shows that the model works very well (r2>0.6) along the Great Plains and

497   some norther counties. However, the coefficient of determination is low in California (2%

498   of national production) and the counties around the Mississippi (3% of national

499   production). California winter wheat is one of the few states that plants most classes of

500   winter wheat such as hard red, hard white, soft white or durum. This heterogeneity of

501   classes maybe the reason of the low coefficients of determination in this area. Regarding

502   Mississippi and Arkansas counties, they generally show low purities (slightly above 40%).

503   This impacts the accuracy of the un-mixing given that the wheat signal is competing with

504   other surfaces whose signal will be stronger.

505

4.2.2. Ukraine

507   Figure 8 shows the p-value of each regressor considered in the model and the determination

508   coefficient in each *oblast*.

509

a)   b)

c)   d)
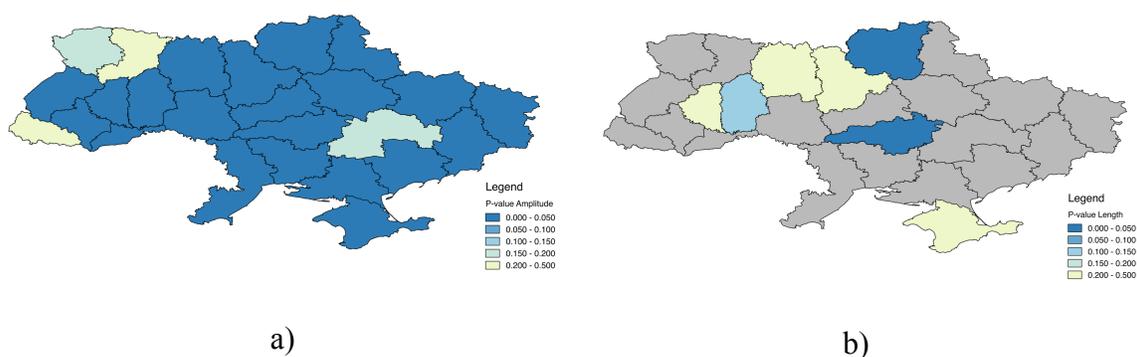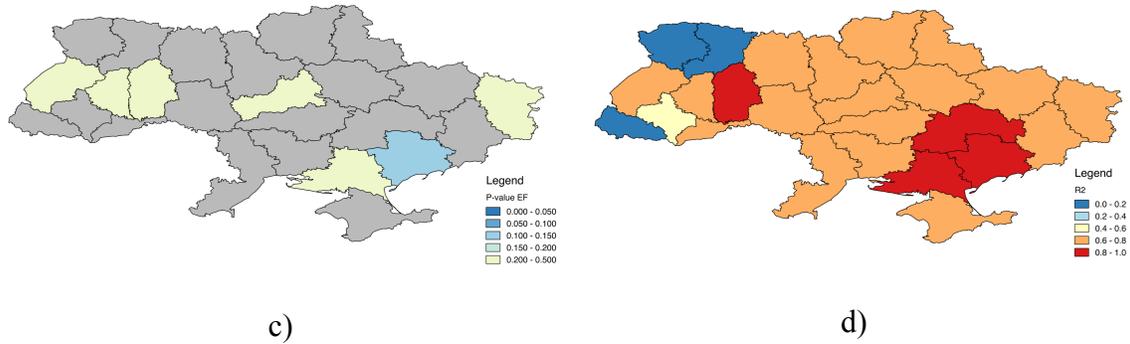
Figure 8. Geographical distribution of *p*-values derived for a) the amplitude of the peak, b) the length of the peak and c) EF. Corresponding coefficients of determination for each *oblast* is shown in subplot (d).

The amplitude shows really good values (<0.05) across the country and just shows p-values higher than 0.05 in four *oblasts*. Both the length and the EF are barely used in Ukraine. The length shows low p-values (<0.05) in two oblasts (located in the center and north respectively) while the other *oblasts* show higher p-values (>0.10). Average EF p-values are generally relatively high (>0.10). The coefficient of determination shows very good correlation (r2>0.4) for most oblasts, especially in the southern area where most of the wheat production is located.

4.3. Validation of the yield model

4.3.1. The US

Figure 9 (top left) shows the county level calibration validation of the model in the US. Each point in the plot represents a given county and a given year during the time-period from 2001 to 2016. The statistics show a high correlation with a coefficient of determination of 0.86 and an error of 0.4 t/ha (15.36%). Figure 9 (top right) shows the

528   county level calibration validation excluding Kansas and Oklahoma in 2007. Though the

529   global statistics barely improve (14.52%), it removes from the plot the overestimation of

530   the low yields caused by the late frost that the model does not capture. Validation of the

531   model using data from the year of 2017 (the only year not used in the calibration) shows

532   performance consistent with the one derived for 2001 to 2016 validation with an error of

533   15.67% and coefficient of determination of 0.86. These results show that the model does

534   not exhibit over-fitting. Figure 9 (bottom left) shows the county level leave-one-out cross-

535   validation (LOOCV), in which iteratively a year is retained for validation, while other years

536   are used for model's calibration. Analogously, Figure 9 (bottom left) shows the county

537   level LOOCV excluding Kansas and Oklahoma in 2007. Figure 10 (left) shows the

538   evaluation of the model by adding up all the county-level statistics where we can apply the

539   model (all data in Figure 9 left) to the national scale and comparing it to the addition of the

540   official county-level statistics of the same counties considered in the model to national-

541   level. Note that at the national level we do not analyze the performance of 2017 separately.

542   The model shows good results for every year except 2007, when the yield is overestimated.

543   Figure 10 (right) shows the national validation excluding Kansas and Oklahoma for 2007

544   since the model is not able to capture the late frost, which improves the validation statistics.

545
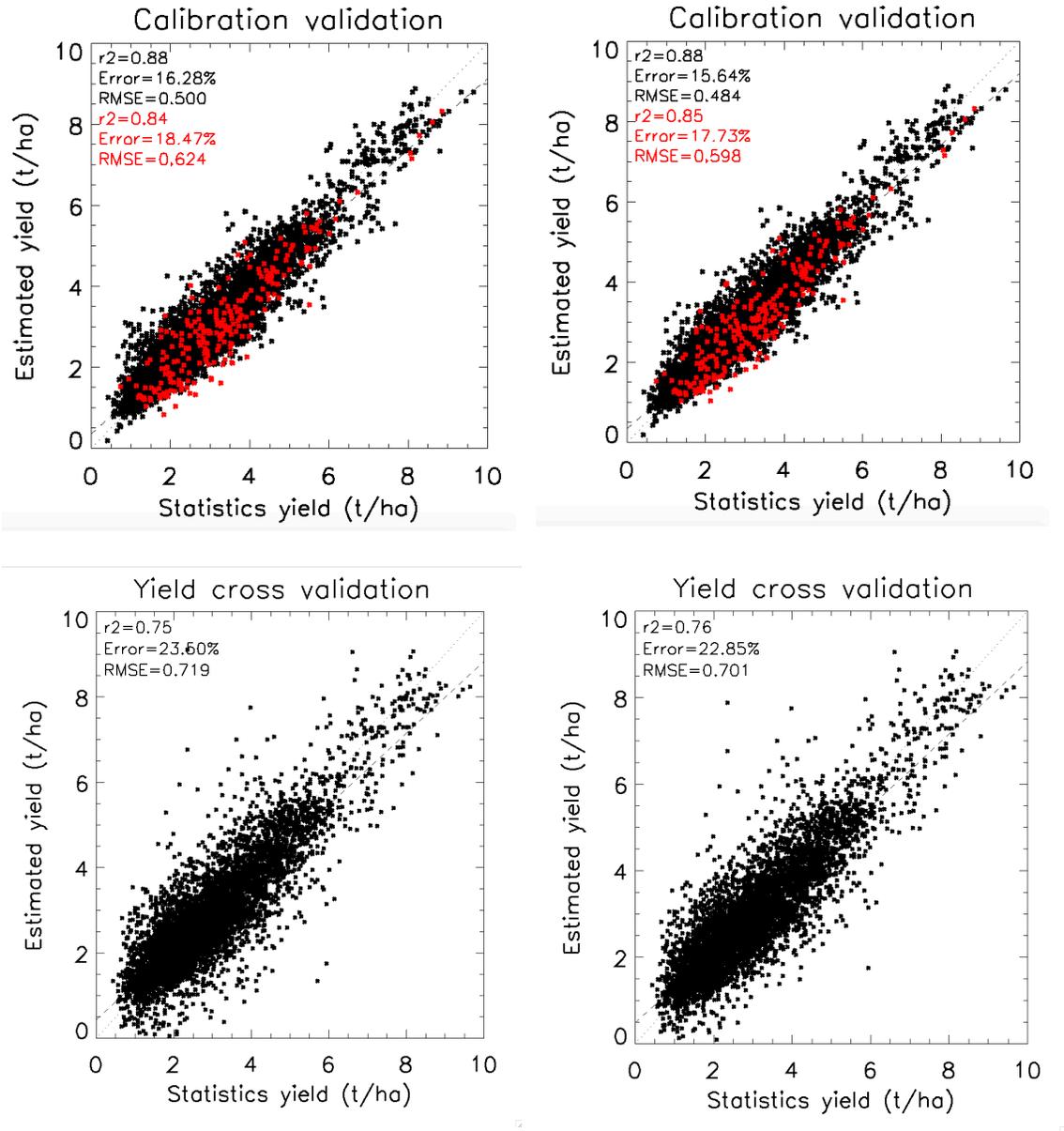
546

547   Figure 9. County level validation in the US (top) from 2001 to 2016 (left) and excluding

548   Kansas and Oklahoma in 2007 (right). Each black dot represents a county and a year from

549   2001 to 2016 while each red dot represents a county in 2017. Bottom: LOOCV at county

550   scale from 2001 to 2017 (left) and excluding Kansas and Oklahoma in 2007 (right).

551
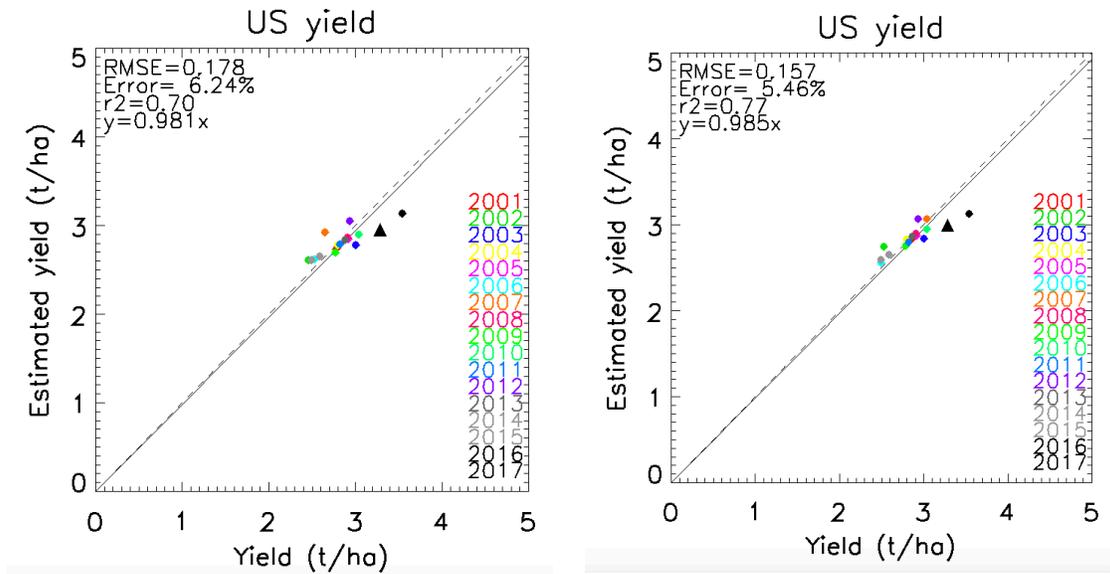
Figure 10. National level aggregation of the county level statistics (left). Excluding

Kansas and Oklahoma in 2007 (right). 2017 is the black triangle.

4.3.2. Ukraine

Figure 11 (left) shows the *oblast* level validation of the model. The statistics show very

similar numbers to the subnational validation in the US, with a little lower coefficient of

determination of 0.73 but a very similar error (to the US) of 0.48 t/ha (14%). Note that in

Ukraine there are much fewer *oblasts* than there are counties in the US and they present a

lower variability in yield as compared to the US (the yield at *oblast* level can get up to 6

t/ha while in the US counties it can get up to 10 t/ha). Validation of the model using data

from the year of 2017 (the only year not used in the calibration) shows a general tendency

of underestimation but performance statistics are similar with the ones derived for 2001 to

2016 validation. Figure 11 (right) shows the oblast level LOOCV. Figure 12 shows the

national validation. The results show an excellent coefficient of determination of 0.85,

567   being able to capture the very low yields in 2003. As in the case of the oblast validation,

568   2017 is underestimated by the model. The RMSE in Ukraine at the national level is 0.27

569   t/ha (8%).

570
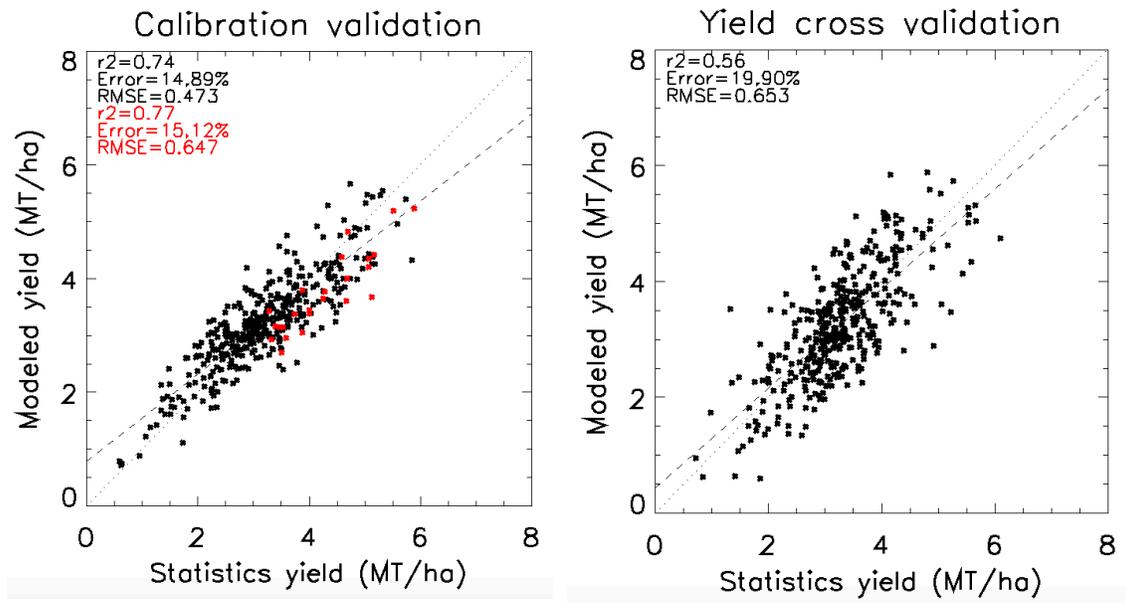
571

572   Figure 11. *Oblast* level calibration validation in Ukraine (left). Each black dot represents

573   an *oblast* and a year from 2001 to 2016 while each red dot represents an *oblast* in 2017.

574   LOOCV at oblast level in Ukraine (right).
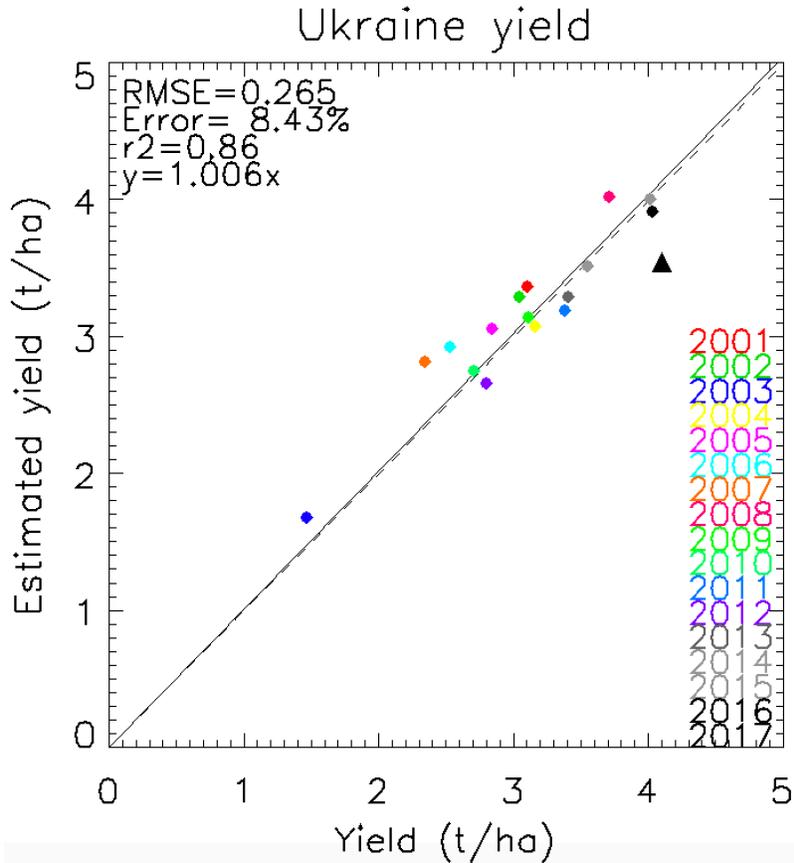
Figure 12. National level validation in Ukraine. 2017 is the black triangle.

5. DISCUSSION

Most crop yield models based on remote sensing coarse resolution data use static cropland

or crop type masks to aggregate spatially the remotely sensed variables as discussed in the

introduction section. When working with static crop type masks, coarsening the spatial

resolution of the EO timeseries data can result in a constant per-pixel proportion over

multiple seasons [54]. Besides, cropland static masks are more stable but do not isolate

different crops. The studies working on these masks generally look at a special time of the

growing season when the crop considered shows a particular pattern of the variable studied

that allows its isolation from the other crops in the cropland mask. However, for summer

587   crops, whose phenology is very similar the mix of different crops can lead to wrong

588   conclusions. In this work, we have developed a model to un-mix the wheat signal from

589   other surfaces surrounding the fields by using yearly crop type masks. This method

590   assumes that the DVI_{wheat} and the DVI_{others} remain constant for an AU for a given date.

591   Therefore, we assume that within the AU the climate is similar, there are no significant

592   elevation differences and the vegetation types in the area are in the same phenological

593   stage. This is a big assumption specially for AUs with very different varieties of wheat and

594   treatments. However, the model works well at sub-national to national scale. Every model

595   has a trade-off and we acknowledge this limitation of coarse spatial resolution data. This

596   will be addressed by using moderate to high spatial resolution data, that will allow the yield

597   estimation at field level. However, the trade-off of such sensors (i.e. Landsat, that dates

598   back to the 70s) is the low temporal resolution with 16 days revisiting time during the

599   historical time series. With the availability of Sentinel 2 data from 2015 and its combination

600   with Landsat [55] will address this gap but by now we can use coarse resolution time series

601   to train the model and learn from it. Looking at the 100% wheat pixels average at the

602   Landsat level for the three different counties considered, the maximum DVI standard

603   deviation is 0.06 which means an error of 10-15% at the peak. Nevertheless, when

604   evaluating the method with Landsat data, the results show good agreement compared to

605   the county 30m wheat pixels average with good correlation and errors of 0.02. One

606   limitation of the method, related to the limitation of working with coarse to moderate

607   spatial resolution data, is that it can only be applied when there is at least one pixel with

608   wheat purity higher than 40% within the AU. As discussed above, by establishing this

609   threshold, we ensure that there is enough of a wheat signal to be un-mixed from the other

610   surfaces.

611   Our model is based on the study of different patterns of the wheat DVI evolution through

612   the growing season. In the introduction, we described some yield models that use other

613   vegetation indexes such as the NDVI as a reference. However, the MODIS Surface

614   Reflectance product is now well established [56] and widely used by the science

615   community as a reference. In this context and given the limitations of using the NDVI on

616   dense vegetation (e.g., [57]–[63]), much simpler vegetation indices that were originally

617   discarded as they did not minimize for atmospheric or BRDF effects, can now be used.

618   This is the case of the DVI that we use in this work. The good results obtained using this

619   index demonstrate that it is well correlated to crop yields and is responsive to high yield

620   values. As an example of the performance of this index compared to the NDVI, Figure 14

621   shows the subnational (top) and national (bottom) validation just using the peak value of

622   the DVI (left) compared to NDVI (right). At subnational level (Figure 14 top), the

623   difference during the calibration period (2001-2016) between using DVI and NDVI is not

624   significant given the high amount of data included in the statistics. However, the validation

625   of a single year (2017) using the peak DVI reduces the RMSE 0.1 t/ha (3%). The difference

626   is also noticed at national level (Figure 14 bottom), where the peak DVI shows better results

627   than the peak NDVI. Note that in this work we choose the DVI for simplicity, but other

628   indexes such as the Enhanced Vegetation Index (EVI) or the Leaf Area Index (LAI) may

629   provide equivalent results.
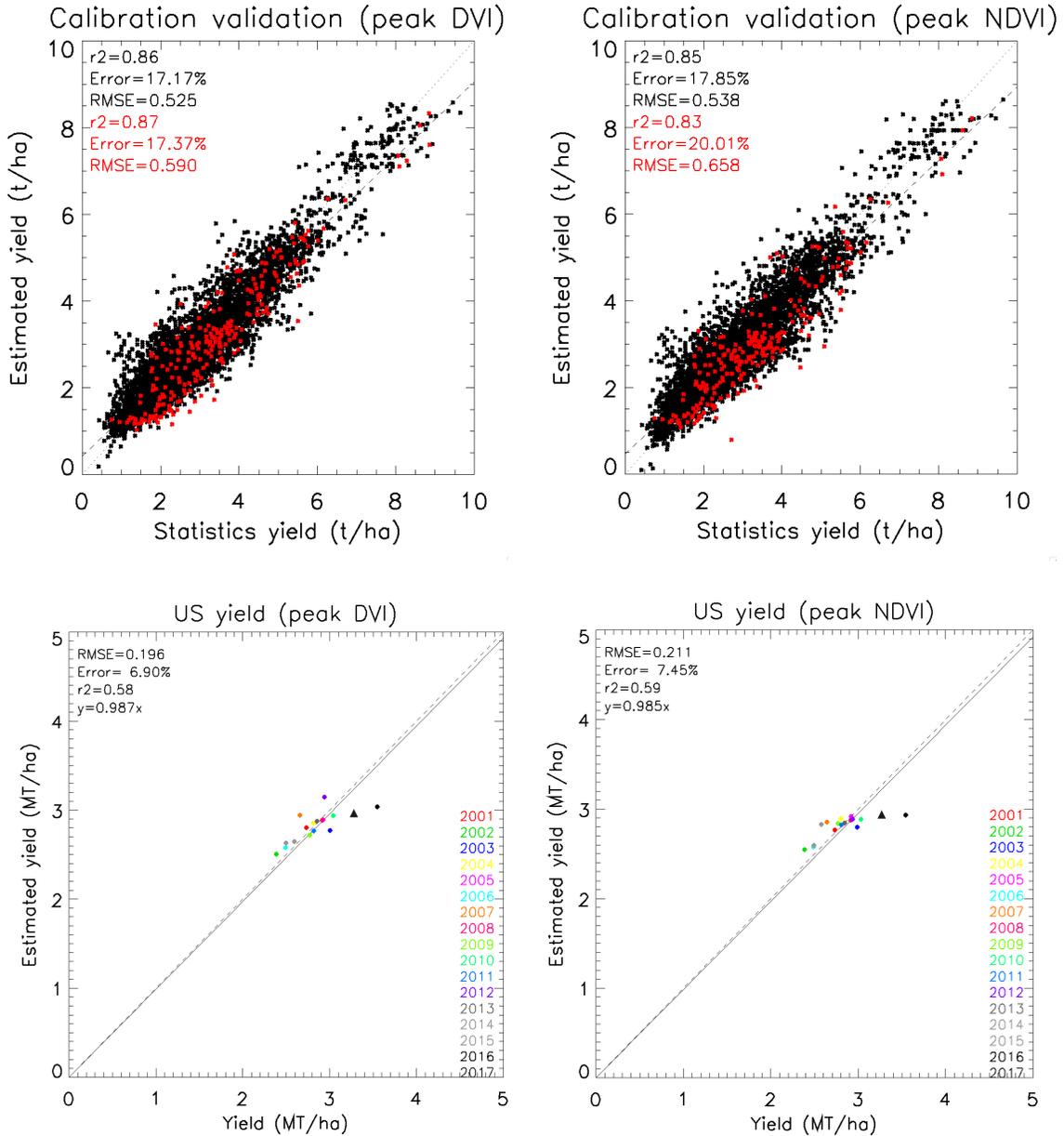
Figure 13. US county level (top) and national level (bottom) validation when using just the peak DVI (left) and peak NDVI (right).

Besides the un-mixing method and the use of the DVI, the main hypothesis of the model is the correlation of the yield with the DVI seasonal peak amplitude, the time that the peak remains and the average EF 30 days after the peak. The three regressors proposed are

636   focused on the reproduction stage of the wheat (DVI amplitude and length) and the grain

637   filling process (EF 30-days average). In the introduction, we referenced some EO-based

638   yield models whose main input variables might be the NDVI peak, its temporal integration,

639   its senescence rate, combining the NDVI with the LST or using other indexes such as

640   FAPAR or ESI. While the amplitude of the peak has been used in other studies (but using

641   NDVI instead), the length of the peak has not been analyzed previously (to our best

642   knowledge). Though the peak definition is a singular point, in some cases, the peak can

643   last a few days becoming a short plateau. During this time period, the plant remains in the

644   same phenological stage. Looking at the yield statistics we noticed that the yield is

645   positively correlated with the length of the peak. This is especially the case of 2016 (highest

646   production year) in some counties of Kansas, where the amplitude is very similar to other

647   years in the dataset while the length made a difference. The 30-days average EF, as

648   mentioned in the methods section, accounts for any water stress condition that happen

649   during and after the peak. Regarding the statistical significance of each regressor, the

650   amplitude of the DVI seasonal peak shows the lowest p-values of the three regressors both

651   in the US and Ukraine (Figures 5 and 6). Meanwhile, the other two regressors remain

652   significant in some regions of the US as discussed in the results section. However, in

653   Ukraine the length and the EF show less statistical significance throughout the different

654   *oblasts*. Besides, Table 4 shows the number of US counties where each calibration equation

655   is used. The most used equations are 11 (just using the amplitude) and 14 (amplitude and

656   length) that account for over 70% of the counties. However, the equations that include the

657   EF (13, 15 and 16) represent about the 25%. It is also interesting that the three regressors

658   are just considered in 14.2% of counties.

659

660   Table 4. Number of US counties where each equation is used.

Equation	Number of counties	%
11	167	38.8
12	27	6.3
13	10	2.3
14	147	34.2
15	18	4.2
16	61	14.2

661

662

663   When applying the method to the US, the proposed model was able to capture most of the

664   extreme events that occurred during the MODIS era. Figure 14 shows the county level

665   validation of two example of extreme events in the US. In 2012, the combination of hot

666   weather and dry conditions during May affected the final yield specially in the Great Plains

667   (Figure 14 left). Besides, during mid-May 2015 Texas and Oklahoma were hit by a flood

668   event that damaged the wheat fields (Figure 14 right).

669
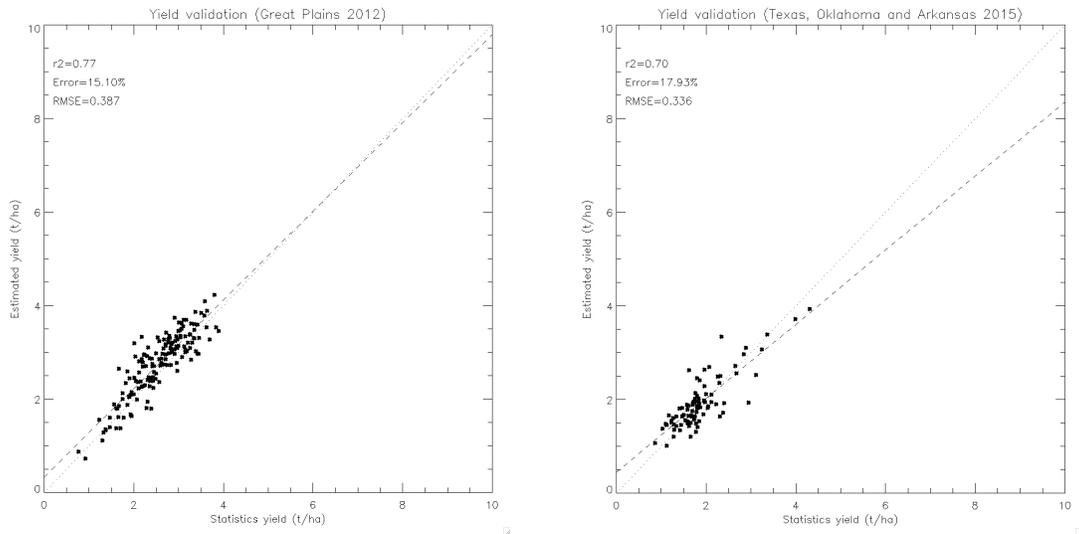
Figure 14. Examples of the model performance in areas hit by drought (left) and flooding (right).

The model was also able to capture high yields, such as 2016, when the rains in late April and mild temperatures in May led to the highest historical yields in the US. However, the model is currently not able to capture late frosts such as the one that impacted Kansas and Oklahoma during early April of 2007. However, it is able to capture the 2003 February-March frost that killed over 50% of the wheat crops in Ukraine. Winter wheat plants are tolerant to temperatures of −20 ◦C in the vegetative stages but suffer severe damage at much more moderate temperatures (−5 to −7 ◦C) during the reproductive stages (Fuller et al. , 2007). Thus, the impact of a frost depends on the development stage of the wheat [65]. In Ukraine, the frost occurred in earlier phenological stages of the plant and it affected the biomass by showing much lower peak amplitudes than other "normal" years. Additionally, the failure was result not just of the frost but also of winter kill, low snow cover and a May-June drought. However, the frost in the US happened later in the phenological stage of the plant and did not have any consequence on the amplitude or the length of the peak. Given

686   that the 2007 late frost is the only such occurrence during the MODIS era, we did not have

687   enough data to include a parameter responsive to this event to train the model. In the future,

688   we will explore the application of the model to AVHRR historical data looking for other

689   late frost events happening in the past in the US.

690   In Ukraine, the subnational statistics are divided into fewer cases than the US (29 *oblasts*

691   versus about 500 counties). The main advantage of having coarser sub-divisions is that we

692   can apply the model to every *oblast*, that there are statistics for all of them and the *oblast*-

693   level performance is comparable to the county-level (same error of 0.4 t/ha). However, this

694   lower number of subnational statistics leads to a higher error at the national level as

695   compared to the US.

696   Comparing the results of the proposed model to other methods in the literature to monitor

697   wheat in the US and Ukraine, the proposed model improves [4] and [15] errors both in the

698   US (7.5% to 5.5%) and Ukraine (12% to 8%) and increases significantly the coefficient of

699   determination in the US (0.24 to 0.77) and Ukraine (0.71 to 0.88). However, such studies

700   just analyzed data from 2001 to 2012. Compared to [20], Ukraine results at oblast level

701   (0.4 to 0.6 t/ha) are in line with their reported errors (0.5 to 0.8 t/ha) though their study just

702   included the time series from 2001 to 2012.

703   Going forward we will start to run this model in real time to forecast in season yields. This

704   will be based on the processing of in season wheat masks using [26] and the integration of

705   [15] model to forecast the peak amplitude. Regarding the length, we can explore the DVI

706   versus the GDD evolution and adjust it to a function to forecast the length. Finally, the 30-

707   days average EF can be related with different climatology variables that can be forecast by

708   weather models. Additionally, as new data becomes available, it will continue to then

709   improve the model.

710

711   **6. CONCLUSIONS**

712   This work presents a new EO-based empirical winter wheat yield model. It is based on the

713   un-mixing of the wheat signal by using yearly crop type masks. This wheat signal is used

714   to calibrate the model using as inputs the seasonal amplitude and length of the DVI peak

715   and the average of the EF 30 days after the peak. The model is calibrated in the US at the

716   county level and in Ukraine at the oblast level using historical statistics from 2001 to 2016.

717   In each administrative unit, a different calibration coefficient (based on all possible

718   combination of the three regressors) is selected depending on the statistical significance of

719   each variable. Finally, the model is validated not just during 2001 to 2016 but also

720   including 2017, which is not used in the calibration process. Overall, the presented model

721   shows a strong performance both at the national (r^2 higher than 0.7 and RMSE of 0.16 t/ha

722   in the US and 0.27 t/ha in Ukraine) and the subnational level (r^2 higher than 0.7 and a

723   RMSE lower than 0.6 t/ha). Our goal in the future is to extend the model to forecast the

724   yield and apply it to other crops and other countries.

725

730

731 **REFERENCES**

732 [1]   U. FAO, "How to feed the world in 2050," in *Rome: High-Level Expert Forum*,

733   2009.

734 [2]   C. Rosenzweig, A. Iglesias, X. B. Yang, P. R. Epstein, and E. Chivian, "Climate

735   Change and Extreme Weather Events; Implications for Food Production, Plant

736   Diseases, and Pests," *Glob. Chang. Hum. Heal.*, vol. 2, no. 2, pp. 90–104, 2001.

737 [3]   R. B. Macdonald, "A summary of the history of the development of automated

738   remote sensing for agricultural applications," *IEEE Transactions on Geoscience*

739   *and Remote Sensing*, vol. GE-22, no. 6. pp. 473–482, 1984.

740 [4]   I. Becker-Reshef, E. Vermote, M. Lindeman, and C. Justice, "A generalized

741   regression-based model for forecasting winter wheat yields in Kansas and Ukraine

742   using MODIS data," *Remote Sens. Environ.*, vol. 114, no. 6, pp. 1312–1323, 2010.

743 [5]   P. J. Pinter, R. D. Jackson, S. B. Idso, and R. J. Reginato, "Multidate spectral

744   reflectance as predictors of yield in water stressed wheat and barley," *Int. J.*

745   *Remote Sens.*, vol. 2, no. 1, pp. 43–48, Jan. 1981.

746 [6]   C. J. Tucker, C. Vanpraet, E. Boerwinkel, and A. Gaston, "Satellite remote sensing

747   of total dry matter production in the Senegalese Sahel," *Remote Sens. Environ.*,

748   vol. 13, no. 6, pp. 461–474, 1983.

749 [7]   S. N. Goward and D. G. Dye, "Evaluating North American net primary

750   productivity with satellite observations," *Adv. Sp. Res.*, vol. 7, no. 11, pp. 165–174,

751   1987.

752 [8]   P. C. Doraiswamy, T. R. Sinclair, S. Hollinger, B. Akhmedov, A. Stern, and J.

753   Prueger, "Application of MODIS derived parameters for regional crop yield

754   assessment," *Remote Sens. Environ.*, vol. 97, no. 2, pp. 192–202, 2005.

755   [9]   M. Meroni, E. Marinho, N. Sghaier, M. M. Verstrate, and O. Leo, "Remote

756   Sensing Based Yield Estimation in a Stochastic Framework — Case Study of

757   Durum Wheat in Tunisia," *Remote Sensing* , vol. 5, no. 2. 2013.

758   [10]   F. Baret and G. Guyot, "Monitoring of the ripening period of wheat canopies using

759   visible and near infra red radiometry [reflectance, vegetation index, senescence

760   rate, water plateau]," *Agronomie (France)*, vol. v. 6. 1986.

761   [11]   D. M. Johnson, "An assessment of pre- and within-season remotely sensed

762   variables for forecasting corn and soybean yields in the United States," *Remote*

763   *Sens. Environ.*, vol. 141, pp. 116–128, 2014.

764   [12]   J. H. Kastens, T. L. Kastens, D. L. A. Kastens, K. P. Price, E. A. Martinko, and R.-

765   Y. Lee, "Image masking for crop yield forecasting using AVHRR NDVI time

766   series imagery," *Remote Sens. Environ.*, vol. 99, no. 3, pp. 341–356, 2005.

767   [13]   D. M. Johnson and R. Mueller, "The 2009 Cropland Data Layer.," *Photogramm.*

768   *Eng. Remote Sens.*, vol. 76, no. 11, pp. 1201–1205, 2010.

769   [14]   T. Fisette *et al.*, "AAFC annual crop inventory," *2013 Second International*

770   *Conference on Agro-Geoinformatics (Agro-Geoinformatics)*. pp. 270–274, 2013.

771   [15]   B. Franch *et al.*, "Improving the timeliness of winter wheat production forecast in

772   the United States of America, Ukraine and China using MODIS data and NCAR

773   Growing Degree Day information," *Remote Sens. Environ.*, vol. 161, pp. 131–148,

774   2015.

775   [16]   B. Franch *et al.*, "A 30+ Year AVHRR Land Surface Reflectance Climate Data

776   Record and Its Application to Wheat Yield Monitoring," *Remote Sensing* , vol. 9,

777   no. 3. 2017.

778   [17] W. Kowalik, K. Dabrowska-Zielinska, M. Meroni, T. U. Raczka, and A. de Wit,

779   "Yield estimation using SPOT-VEGETATION products: A case study of wheat in

780   European countries," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 32, pp. 228–239, 2014.

781   [18] G. Büttner, J. Feranec, G. Jaffrain, L. Mari, G. Maucha, and T. Soukup, "The

782   CORINE land cover 2000 project," *EARSeL eProceedings*, vol. 3, no. 3, pp. 331–

783   346, 2004.

784   [19] R. López-Lozano *et al.*, "Towards regional grain yield forecasting with 1km-

785   resolution EO biophysical products: Strengths and limitations at pan-European

786   level," *Agric. For. Meteorol.*, vol. 206, pp. 12–32, 2015.

787   [20] F. Kogan *et al.*, "Winter wheat yield forecasting in Ukraine based on Earth

788   observation, meteorological data and biophysical models," *Int. J. Appl. Earth Obs.*

789   *Geoinf.*, vol. 23, pp. 192–203, 2013.

790   [21] M. C. Anderson *et al.*, "The Evaporative Stress Index as an indicator of

791   agricultural drought in Brazil: An assessment based on crop yield impacts,"

792   *Remote Sens. Environ.*, vol. 174, pp. 82–99, 2016.

793   [22] A. J. Richardson and C. L. Wiegand, "Distinguishing vegetation from soil

794   background information.," *Photogramm. Eng. Remote Sensing*, vol. 43, pp. 1541–

795   1552, 1977.

796   [23] G. Vocke and M. B. Ali, *US wheat production practices, costs, and yields:*

797   *Variations across regions*. United States Department of Agriculture, Economic

798   Research Service, 2013.

799   [24] C. Boryan, Z. Yang, R. Mueller, and M. Craig, "Monitoring US agriculture: the

800   US Department of Agriculture, National Agricultural Statistics Service, Cropland

801   Data Layer Program," *Geocarto Int.*, vol. 26, no. 5, pp. 341–358, Aug. 2011.

802   [25] E. Vermote, "MOD09A1 MODIS/Terra Surface Reflectance 8-Day L3 Global

803   500m SIN Grid V006," *NASA EOSDIS L. Process. DAAC*, vol. 10, 2015.

804   [26] S. Skakun *et al.*, "Early season large-area winter crop mapping using MODIS

805   NDVI data, growing degree days information and a Gaussian mixture model,"

806   *Remote Sens. Environ.*, vol. 195, pp. 244–258, 2017.

807   [27] M. Lavreniuk, N. Kussul, S. Skakun, A. Shelestov, and B. Yailymov, "Regional

808   retrospective high resolution land cover for Ukraine: Methodology and results," in

809   *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*,

810   2015, pp. 3965–3968.

811   [28] M. S. Lavreniuk *et al.*, "Large-Scale Classification of Land Cover Using

812   Retrospective Satellite Data," *Cybern. Syst. Anal.*, vol. 52, no. 1, pp. 127–138,

813   2016.

814   [29] F. M. Bréon, E. Vermote, E. F. Murphy, and B. Franch, "Measuring the

815   Directional Variations of Land Surface Reflectance From MODIS," *IEEE*

816   *Transactions on Geoscience and Remote Sensing*, vol. 53, no. 8. pp. 4638–4649,

817   2015.

818   [30] E. Vermote, C. O. Justice, and F.-M. Bréon, "Towards a generalized approach for

819   correction of the BRDF effect in MODIS directional reflectances," *IEEE Trans.*

820   *Geosci. Remote Sens.*, vol. 47, no. 3, pp. 898–908, 2009.

821   [31] B. Franch, E. F. Vermote, J. a. Sobrino, and Y. Julien, "Retrieval of Surface

822   Albedo on a Daily Basis: Application to MODIS Data," *IEEE Trans. Geosci.*

823   *Remote Sens.*, vol. 52, no. 12, pp. 7549–7558, 2014.

824   [32] C. B. Schaaf *et al.*, "First operational BRDF, albedo nadir reflectance products

825   from MODIS," *Remote Sens. Environ.*, vol. 83, no. 1–2, pp. 135–148, 2002.

826   [33] J.-L. Roujean, M. Leroy, and P.-Y. Deschamps, "A bidirectional reflectance model

827   of the Earth's surface for the correction of remote sensing data," *J. Geophys. Res.*,

828   vol. 97, no. D18, p. 20455, 1992.

829   [34] F. Maignan, F. M. Bréon, and R. Lacaze, "Bidirectional reflectance of Earth

830   targets: Evaluation of analytical models using a large set of spaceborne

831   measurements with emphasis on the Hot Spot," *Remote Sens. Environ.*, vol. 90, no.

832   2, pp. 210–220, 2004.

833   [35] X. Li and A. H. Strahler, "Geometric-optical bidirectional reflectance modeling of

834   the discrete crown vegetation canopy: effect of crown shape and mutual

835   shadowing," *IEEE Trans. Geosci. Remote Sens.*, vol. 30, no. 2, pp. 276–292, 1992.

836   [36] W. Lucht, "Expected retrieval accuracies of bidirectional reflectance and albedo

837   from EOS-MODIS and MISR angular sampling," *J. Geophys. Res.*, vol. 103, no.

838   D8, pp. 8763–8778, 1998.

839   [37] P. Lewis and M. Barnsley, "Influence of the sky radiance distribution on various

840   formulations of the earth surface albedo," *Proceedings of the Conference on*

841   *Physical Measurements and Signatures in Remote Sensing.* pp. 707–715, 1994.

842   [38] W. Lucht and P. Lewis, "Theoretical noise sensitivity of BRDF and albedo

843   retrieval from the EOS-MODIS and MISR sensors with respect to angular

844   sampling," *Int. J. Remote Sens.*, vol. 21, no. 1, pp. 81–98, 2000.

845   [39] E. Vermote, D. Tanre, J. L. Deuze, M. Herman, and J. J. Morcrette, "Second

846   Simulation of the Satellite Signal in the Solar Spectrum (6S). 6S User Guide

847   Version 2. Appendix III: Description of the subroutines," vol. 35, no. 3, 1997.

848   [40]   W. G. M. Bastiaanssen, M. Menenti, R. A. Feddes, and A. A. M. Holtslag, "A

849   remote sensing surface energy balance algorithm for land (SEBAL). 1.

850   Formulation," *J. Hydrol.*, vol. 212, pp. 198–212, 1998.

851   [41]   R. Allen, M. Tasumi, and R. Trezza, "Satellite-Based Energy Balance for Mapping

852   Evapotranspiration with Internalized Calibration (METRIC)—Model," *J. Irrig.*

853   *Drain. Eng.*, vol. 133, no. 4, pp. 380–394, Aug. 2007.

854   [42]   G. Ravazzani, C. Corbari, S. Morella, P. Gianoli, and M. Mancini, "Modified

855   Hargreaves-Samani Equation for the Assessment of Reference Evapotranspiration

856   in Alpine River Basins," *J. Irrig. Drain. Eng.*, vol. 138, no. 7, pp. 592–599, Jul.

857   2012.

858   [43]   M. Menenti and B. J. Choudhury, "Parameterization of land surface evaporation

859   by means of location dependent potential evaporation and surface temperature

860   range," *Proc. IAHS Conf. L. Surf. Process.*, no. 212, pp. 561−568, 1993.

861   [44]   G. J. Roerink, Z. Su, and M. Menenti, "S-SEBI: A simple remote sensing

862   algorithm to estimate the surface energy balance," *Phys. Chem. Earth, Part B*

863   *Hydrol. Ocean. Atmos.*, vol. 25, no. 2, pp. 147–157, 2000.

864   [45]   Z. Su, "The Surface Energy Balance System (SEBS) for estimation of turbulent

865   heat fluxes," *Hydrol. Earth Syst. Sci. Discuss.*, vol. 6, no. 1, pp. 85–100, 2002.

866   [46]   M. C. Anderson *et al.*, "Evaluation of Drought Indices Based on Thermal Remote

867   Sensing of Evapotranspiration over the Continental United States," *J. Clim.*, vol.

868   24, no. 8, pp. 2025–2044, Apr. 2011.

869   [47]   M. Menenti, W. Bastiaanssen, D. van Eick, and M. A. Abd el Karim, "Linear

870   relationships between surface reflectance and temperature and their application to

871   map actual evaporation of groundwater," *Adv. Sp. Res.*, vol. 9, no. 1, pp. 165–176,

872   1989.

873   [48]   M. Gómez, A. Olioso, J. A. Sobrino, and F. Jacob, "Retrieval of

874   evapotranspiration over the Alpilles/ReSeDA experimental site using airborne

875   POLDER sensor and a thermal camera," *Remote Sens. Environ.*, vol. 96, no. 3, pp.

876   399–408, 2005.

877   [49]   Z. Su, H. Pelgrum, and M. Menenti, "Aggregation effects of surface heterogeneity

878   in land surface processes," *Hydrol. Earth Syst. Sci.*, vol. 3, no. 4, pp. 549–563,

879   Dec. 1999.

880   [50]   W. W. Verstraeten, F. Veroustraete, and J. Feyen, "Estimating evapotranspiration

881   of European forests from NOAA-imagery at satellite overpass time: Towards an

882   operational processing chain for integrated optical and thermal sensor data

883   products," *Remote Sens. Environ.*, vol. 96, no. 2, pp. 256–276, 2005.

884   [51]   Z. Wan and J. Dozier, "A generalized split-window algorithm for retrieving land-

885   surface temperature from space," *IEEE Trans. Geosci. Remote Sens.*, vol. 34, no.

886   4, pp. 892–905, Jul. 1996.

887   [52]   R. A. Fischer and R. Maurer, "Drought resistance in spring wheat cultivars. I.

888   Grain yield responses," *Aust. J. Agric. Res.*, vol. 29, no. 5, pp. 897–912, 1978.

889   [53]   F. Giunta, R. Motzo, and M. Deidda, "Effect of drought on yield and yield

890   components of durum wheat and triticale in a Mediterranean environment," *F.

891   Crop. Res.*, vol. 33, no. 4, pp. 399–409, 1993.

892   [54]   I. Becker-Reshef *et al.*, "Prior Season Crop Type Masks for Winter Wheat Yield

893   Forecasting: A US Case Study," *Remote Sensing* , vol. 10, no. 10. 2018.

894   [55]   M. Claverie *et al.*, "The Harmonized Landsat and Sentinel-2 surface reflectance

895   data set," *Remote Sens. Environ.*, vol. 219, pp. 145–161, 2018.

896   [56]   E. F. Vermote, N. Z. El Saleous, and C. O. Justice, "Atmospheric correction of

897   MODIS data in the visible to middle infrared: first results," *Remote Sens. Environ.*,

898   vol. 83, no. 1–2, pp. 97–111, 2002.

899   [57]   A. A. Gitelson, "Wide Dynamic Range Vegetation Index for Remote

900   Quantification of Biophysical Characteristics of Vegetation," *J. Plant Physiol.*,

901   vol. 161, no. 2, pp. 165–173, 2004.

902   [58]   F. Baret and G. Guyot, "Potentials and limits of vegetation indices for LAI and

903   APAR assessment," *Remote Sens. Environ.*, vol. 35, no. 2, pp. 161–173, 1991.

904   [59]   A. A. Gitelson, Y. J. Kaufman, R. Stark, and D. Rundquist, "Novel algorithms for

905   remote estimation of vegetation fraction," *Remote Sens. Environ.*, vol. 80, no. 1,

906   pp. 76–87, 2002.

907   [60]   A. A. Gitelson, R. Stark, U. Grits, D. Rundquist, Y. Kaufman, and D. Derry,

908   "Vegetation and soil lines in visible spectral space: A concept and technique for

909   remote estimation of vegetation fraction," *Int. J. Remote Sens.*, vol. 23, no. 13, pp.

910   2537–2562, Jan. 2002.

911   [61]   C. Buschmann and E. Nagel, "In vivo spectroscopy and internal optics of leaves as

912   basis for remote sensing of vegetation," *Int. J. Remote Sens.*, vol. 14, no. 4, pp.

913   711–722, Mar. 1993.

914   [62]   P. J. Sellers *et al.*, "Remote sensing of the land surface for studies of global

915   change: Models — algorithms — experiments," *Remote Sens. Environ.*, vol. 51,

916   no. 1, pp. 3–26, 1995.

917   [63]   S. Skakun, E. Vermote, J.-C. Roger, and B. Franch, "Combined Use of Landsat-8

918   and Sentinel-2A Images for Winter Crop Mapping and Winter Wheat Yield

919   Assessment at Regional Scale," *AIMS Geosci.*, vol. 3, no. 2, pp. 163–186, 2017.

920   [64]   M. P. Fuller, A. M. Fuller, S. Kaniouras, J. Christophers, and T. Fredericks, "The

921   freezing characteristics of wheat at ear emergence," *Eur. J. Agron.*, vol. 26, no. 4,

922   pp. 435–441, 2007.

923   [65]   J. P. Shroyer, M. E. Mikesell, and G. M. Paulsen, *Spring freeze injury to Kansas*

924   *wheat*. Cooperative Extension Service, Kansas State University, 1995.

925   Vermote, E., 2015. MOD09Q1 MODIS/Terra Surface Reflectance 8-day L3 Global 250m

926   SIN Grid V006. NASA EOSDIS Land Processes DAAC.

927   http://dx.doi.org/10.5067/MODIS/MOD09Q1.006 (Accessed 26.02.2018).

928   URL 1. http://www.amis-outlook.org/

929   URL 2. http://www.worldstopexports.com/wheat-exports-country/