Policy Research Working Paper 9306

Using Machine Learning to Assess Yield Impacts of Crop Rotation

Combining Satellite and Statistical Data for Ukraine

Klaus Deininger
Daniel Ayalew Ali
Nataliia Kussul
Mykola Lavreniuk
Oleg Nivievskyi

WORLD BANK GROUP
Development Economics
Development Research Group
June 2020

Abstract

To overcome the constraints for policy and practice posed by limited availability of data on crop rotation, this paper applies machine learning to freely available satellite imagery to identify the rotational practices of more than 7,000 villages in Ukraine. Rotation effects estimated based on combining these data with survey-based yield information point toward statistically significant and economically meaningful effects that differ from what has been reported in the literature, highlighting the value of this approach. Independently derived indices of vegetative development and soil water content produce similar results, not only supporting the robustness of the results, but also suggesting that the opportunities for spatial and temporal disaggregation inherent in such data offer tremendous unexploited opportunities for policy-relevant analysis.

This paper is a product of the Development Research Group, Development Economics. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://www.worldbank.org/prwp. The authors may be contacted at kdeininger@worldbank.org and dali1@worldbank.org.

Using Machine Learning to Assess Yield Impacts of Crop Rotation: Combining Satellite and Statistical Data for Ukraine[1]

Klaus Deininger[*], **Daniel Ayalew Ali**[*], **Nataliia Kussul**[§], **Mykola Lavreniuk**[†], **Oleg Nivievskyi**[#]

JEL Codes O13, Q12, Q15

Keywords: Machine learning, crop rotation, Ukraine, agricultural productivity

* Development Research Group, The World Bank
[§] Space Research Institute of National Academy of Sciences of Ukraine
and State Space Agency of Ukraine
[†] National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute"
[#] Kyiv School of Economics

[1] We thank Olha Halytsya, Siobhan Murray, Andrii Shelestov, and Anna Yailymova for useful inputs and insightful comments. The views presented in this paper are those of the authors and do not necessarily represent those of the World Bank, its Executive Directors, or the countries they represent or any of the other institutions they are associated with.

**Using machine learning to assess yield impacts of crop rotation:
Combining satellite and statistical data for Ukraine**

1. Introduction

While agronomic benefits from crop rotation are well documented, economic factors together with the ability to manage nutrients as well as pest, weed, and insect populations by applying synthetic fertilizers or pesticides, often prompt farmers to implement short rotations or plant the same crop in continuation. This raises issues for policy if there are externalities, either across space -e.g., from harmful pests migrating to neighboring fields or run-off from excessive fertilizer application contaminating rivers- or if, possibly due to insecure tenure, land users do not fully account for long-term effects of their actions on soil fertility, thus contributing to natural resource degradation.

Studies suggests that, in cases where such external effects are present, relatively modest incentives could prompt producers to adjust their behavior. Yet, studies to explore this issue are limited, mainly because of limitations in routine statistical data that fail to report on the crops grown in the same field in different years. Most studies are thus based on information from agronomic trials that is infrequent in nature, often removed from field realities, and not available for the production systems practices in many parts of the globe.

In this paper we explore whether use of crop cover information derived from freely available satellite imagery can fill this gap. We generate a crop map for all of Ukraine in three consecutive years and use the resulting data to estimate impacts of crop rotations on yields for the country's four main crops (maize, soybeans, cereals, and sunflower) at the village level. Results are statistically significant, economically meaningful, and very different from what is reported in the US-based literature. Contrary to the literature that is mostly focused on corn(maize)-soybean rotations, we find that continued soybean cultivation incurs only a limited yield penalty and that the magnitude of yield reductions from monocropping is often below that of unsuitable predecessor crops, mainly sunflower, that deplete the soil horizon of water or increase disease pressure.

To allay concerns about measurement error or replicability of this approach in settings where yield information is not at all or less regularly collected, we repeat the regressions using indices of vegetative development that can be derived directly from satellite imagery (EVI, LAI, FAPAR, and LSWI). Using crop masks for the crops considered here produces results that are very similar to those obtained for yields. As such indices can be constructed at levels of spatial and temporal resolution well beyond what is within reach of statistical systems, this offers tremendous potential for research.

Beyond providing substantive insights regarding the impact of rotational practices, our paper is related and contributes to a growing body of evidence on machine learning and the use of remotely sensed data for analysis of agricultural production relations. We document the potential of machine learning by showing that a modest effort of training data collection allows generation of a crop map for the entire agricultural area of Ukraine. Although field-level crop cover data at slightly coarser resolution are freely available for the United States in the form of the crop data layer (CDL), there are few, if any studies outside the United States that have achieved national coverage with field-level data on crop cover.

We also contribute to the evidence on use of remotely sensed data by showing that use of common indices of crops' vegetative development not only produces results consistent with those from combining crop cover data with statistical information on yields, but also affords insights (e.g., on soil water content) in their own right. The scope for disaggregation inherent in such data is a promising avenue for future research, including exploration of implications on profitability and implications for potential policy interventions to eliminate negative externalities resulting from rotational choices.

The rest of the paper is organized as follows. Section 2 briefly reviews available literature as well as the relevance of the topic within the context of Ukraine's agricultural development. Section 3 describes the data sources, provides descriptive statistics on the incidence of crop rotations and yields, and outlines our econometric approach. Section 4 reviews evidence on the impact of rotational practice on yield, commonly used vegetation indices, and soil water content. Section 5 concludes by drawing out implications for future research as well as policy.

2. Literature and context

Insights into the trade-offs involved in on-farm decision making regarding crop choice and sequence derived from detailed study of the productivity of different rotations can help identify economic effects of rotational choices, including whether such choices are likely to trigger external effects and, if yes, help identify suitable instruments to internalize such externalities. Yet, data limitations have implied that study of this topic remained extremely limited, especially outside the United States. The nature of Ukraine's transition from collectivized agriculture as well as recent policy reforms make such analysis particularly relevant.

2.1 The rationale for studying crop rotation effects

Field-level management of the soil's nutrients, physical structure, and weed as well as pest populations via rotation of crops has long been a key source of agricultural productivity, efficiency and resilience. The underlying principle is simple: as crops differ from each other in terms of their susceptibility to disease, the dynamics of their vegetative development, and their effects on the soil's net nutrient balance, sequencing them in ways that capitalize on positive synergies while avoiding negative interactions can increase output or help save input costs (Hennessy 2006). Historically, addition of new crops has been a key driver of productivity growth and the ability of the agricultural sector to release labor for non-agricultural development (Brunt 1999). More recently, synthetic fertilizers, pesticides, and herbicides allow producers to maximize profits by shortening rotations and reducing their diversity (Plourde *et al.* 2013).

Crop rotation affects outcomes via three channels. The first one is the reduction of pest and disease-pressure (Hilton *et al.* 2013) by preventing development and accumulation of crop-specific fungi, bacteria and insects that often survive in crop residues (Gentry *et al.* 2013). Proper rotational management can also reduce greenhouse gas emissions (Behnke *et al.* 2018) and capture soil carbon. A second mechanism is to enhance plants' ability to access nutrients or water through differences in root structure to improve capillarity and aeration of the soil, thus also reducing susceptibility to disease (Koch *et al.* 2018). The underlying dynamics of bacterial, fungal or insect populations and their interaction with climatic conditions can be captured in biological models (Regis Mauri 2019). A third element is legumes' ability to capture nitrogen (Archer *et al.* 2020) as well as crops' differential nutrient requirements (Bowles *et al.* 2020). Together with data on prices for inputs and outputs, technical coefficients based on the varietal choice and crop management options (e.g., till- and no-till practices), availability of on-farm resources -most importantly labor in different parts of the growing season- and their shadow prices, as well as any exogenous constraints (e.g., regulation) can then be used to formulate decision-rules for optimum rotational choices at the farm level as a programming model (Boyabatli *et al.* 2019).

Whether such choices are not only individually but also socially optimal requires more detailed analysis. If property rights to land are well defined and can be enforced and the effects of crop rotation remain plot-specific, privately and socially optimal choices will coincide. However, even if property rights are clear, externalities may arise in several respects. Spatial externalities, e.g., because mono-cropping or short rotations can trigger accumulation of insects that may infect neighboring fields or, if a critical mass is exceeded, cause harmful infestation more broadly (Meisner and Rosenheim 2014). Effluents from more intensive use of fertilizer, pesticides, or manure than would be necessary with more diverse rotations may contribute to water pollution. Systematic overuse of chemicals can lead to pests developing resistance, as has been reported for GM crops (Qiao *et al.* 2017) and glyphosphate (Mitchell 2011). Moreover, if property rights are insecure, land users may maximize short-term benefits, e.g., 'mine' soil, without regard to long-term sustainability.

If supply is price elastic, modest price changes or subsidies may be sufficient to make farmers shift from monocropping to rotations. Hennessy (2006) shows that, depending on prices for maize and fertilizer, a maize-soybean rather than a maize-maize-soybean rotation will be chosen for subsidy levels between US$0 and US$40 per acre. Livingston *et al.* (2015) suggest that an incentive payment of US$4/acre could incentivize reductions in maize monoculture and fertilizer application. Study of rotational choices in other settings could help to better understand trade-offs to be considered in this context. Yet constraints on data availability have traditionally limited the scope for such investigation.

The main data source for assessing economic impacts has been from controlled trials run by agricultural experiment stations, often in the United States. Studies using such data suggest that rotating crops increases resilience (Krupinsky *et al.* 2006) and provides economic benefits by reducing input cost or increasing yields.[2] Maize and soybean grown in annual rotation yielded 13% and 10% more than when grown in continuation, with slightly higher effects on marginal soils (Porter *et al.* 1997). Hennessy (2006) and Livingston *et al.* (2015) find rotation effects of about 25% in similar settings. But apart from being geographically concentrated, agronomic trials are costly, associated with long lead times, and consider only a limited set of crops or management options.

To overcome these limitations, efforts to exploit alternative data sources proceeded in two directions. One way to improve data availability builds on collaboration with the private sector to develop samples, e.g., from suppliers of agro-processing plants (Koch *et al.* 2018), or on use of field-level records from crop insurance (Seifert *et al.* 2017). By highlighting spatial heterogeneity in effect size and a significant role of crop varietal traits, climate (temperature and soil moisture) and management (treatment of crop residue, tillage options, and planting dates), such studies suggest that real-world data can add significant insights to what is available from agronomic trials.

A second avenue is to rely on information derived from remotely sensed imagery. For the United States, the cropland data layer (CDL), based on 30 m pixels from Landsat imagery that are available for all 48 states since 2008 (Boryan *et al.* 2011) are a frequently used source. With proper methodology (Lark *et al.* 2017), the CDL can be used for field-level analysis. Wang and Ortiz-Bobea (2019) use 2005-14 CDL data for the US midwest to show that the 2005-09 biofuel boom triggered a significant increase in maize monocropping and GM maize adoption, especially in the vicinity of newly established biofuel plants.[3] Hendricks *et al.* (2014) use CDL data to argute that estimates of supply elasticities based on aggregate data may be biased and to show that use of field-level data that take into account crop rotation benefits results in estimates that are larger in the short than the long run, contrary to standard models.

[2] For example, comparing a two-year standard maize-soybeans rotation with more diverse a three- or four-year rotations that includes small grains and leguminous fodder crops (clover or alfalfa) led to equal or greater profit, reduced input use, and significantly lower toxicity (Davis *et al.* 2012).
[3] Economic factors may also lead producers to forgo the yield advantage of rotating crops for higher profitability from monoculture as has indeed been observed in the United States where major crops are increasingly grown in monoculture patterns (Plourde *et al.* 2013).

2.2 Why study of crop rotations is of relevance for Ukraine

Studying determinants of adoption and economic impact of different crop rotations in Ukraine is of relevance as the country experienced a significant shift in cropping patterns towards maize and sunflower that has raised concerns about soil mining and associated loss of soil fertility. Owing to its fertile and rich soils, the country was historically the former Soviet Union's breadbasket. With some 41.5 million hectares of agricultural land, more than Argentina (34 mn. ha) or the country's three western neighbors, Poland (11 mn. ha), Germany (12 mn. ha) and France (18 mn ha.) combined, Ukraine has the potential to be a key player in global agriculture.[4] Indeed, following drops in output in the early 1990s that can be attributed to the transition from collectivized farming, the agricultural sector more than doubled value added and now accounts for 10% of GDP and 42% of total exports.

In the late 1990s and early 2000s, close to 30 million hectares of land was privatized to some 7 mn. owners who each received an area of about 4 ha each. Of this, some 20 mn. ha is currently cultivated by legal entities: 6 mn. ha by large farms or agri-holdings and some 9 mn. ha by middle-size farms, both of which emerged mainly from collective farms and more than 4 mn ha is cultivated by so-called individual farmers (*fermerski gospodarstva*). A moratorium on sales of agricultural land introduced in 2001 implies they can only lease but not own their farmland.[5] Several factors, including the lack of a unified lease registry, a mix of digital and paper documents, corruptible registrars, courts and police forces, and legislation prohibiting registration of leases shorter than 7 years that tends to drive landowners into informality, further reduce users' ability to enforce leases, diverting resources from investment or reducing incentives for it.[6] While the resulting tenure insecurity is widely alleged to incentivize unsustainable land management, 'inspections' by government agencies to identify and prosecute violations of environmental rules are perceived as discretionary and ineffective.

Tenure insecurity also reduces Ukraine's potential for irrigation and agricultural value added. The amount of land irrigated, which stood at 2.2 mn. ha in Soviet times, still stands at its low of 0.3 mn. hectares due to insecure land rights, threatening to jeopardize climate resilience and labor-intensive investment to generate jobs. While commercial agriculture is technologically advanced and capital intensive, focus is on mobile capital and production of bulk commodities such as maize, wheat, and

[4] Brazil's total agricultural area is 64 mn. ha, though much of it is of significantly lower fertility than Ukraine's famous black soil. By comparison, total farmland in the United States in 2017 was 364 mn. ha (900 mn ac) with an average farm size of 178 ha (USDA 2019).
[5] While the median farm size is close to 2,000 ha (imposing considerable transaction cost by requiring management of about 500 lease contracts), some large farms are organized in agri-holdings of 100,000 ha or more that access foreign technology, capital and equity. About 12 mn. ha are cultivated by small and household farms that produce half of output value but have remained largely informal and some 9.2 mn. ha remain under state or communal ownership.
[6] The ability to register fraudulent contracts or to obtain connivance from local enforcement agencies triggered a wave of so-called 'raider attacks', often by individuals showing up to harvest crops they have not sown or through strategic cancellation of contracts to cause maximum disruption. Large enterprises hire their own militias to deal with this but medium sized operations who cannot afford this are thought to be particularly vulnerable.

sunflower to the detriment of dairy, horticulture, and agri-processing, all of which would require land-attached investment. To address tenure security constraints, the government recently passed legislation that would allow farmers to sell or buy agricultural land.[7] Subsidy programs to promote diversification of crops and a move towards higher value addition have also been put in place. Partly due to lack of data to underpin appropriate designs, these have been ill-targeted and ineffective.

3. Data, descriptive statistics, and econometric approach

While national statistical data suggest that the diversity of the country's agricultural output mix has decreased over time, they provide no information on mono-cropping that would allow to assess the incidence and impact of crop rotations. To generate such data, we combine a crop cover map generated from freely available satellite imagery via machine learning with statistical data for 2016, 2017, and 2018. With some 24% of area in maize, 17% in soybeans, 10% in sunflower, and 20% in cereals having been cropped in at least two consecutive seasons on the same field, detailed study of associated yield effects seems warranted and we describe the econometric approach for doing so.

3.1 Data sources and key outcome variables

Remotely sensed data provide objective information at low cost and the number, resolution and revisit frequency of optical or radar sensors to provide such data have expanded significantly. Imagery covering the entire world at 10 m spatial resolution every 5 days is available freely from the European Space Agency Sentinel constellation since 2013. Access to cloud computing platforms to process these at scale implies that location-specific training data are rapidly emerging as a binding constraint to generation of crop cover maps at field-scale. To generate such training data for Ukraine, *in-situ* data collection along main roads was undertaken every year during the 2016-18 period along different routes as displayed in figure 1. Experts covered these routes two times each year to capture images of winter as well as summer crops to serve as training data following standard guidelines. Appendix table 1 provides information on the distribution of fields by crop type.

To generate crop cover estimates for 2016, 2017, and 2018, all optical data from Sentinel-2 and SAR data from Sentinel-1 during the vegetation period were used to run a convoluted neural network on the Amazon Web Services cloud computing platform.[8] For classifier training, half of the data was randomly assigned to training and independent validation samples and accuracies obtained based on independent test sets for each of the crops received are in appendix table 1. Resulting maps for the four crops of interest are displayed in figure 2 with differences across years visible even at this scale.

[7] Recently passed legislation would partly address this constraint through a phased opening of agricultural land purchase markets starting on July 1, 2021 when agricultural land purchases will be allowed only by individuals with an ownership limit of 100 ha per person.
[8] See Kussul *et. al* (2018 and 2019) for a detailed description of methodology.

The above process generated raster data that provide information on crops grown in 2016, 2017 and 2018 for every 10*10 m pixel in Ukraine. Combining pixel-level data across years provides the basis for computing the area devoted to different types of crop rotations just by counting pixels. To keep the regressions manageable, we focus on maize, sunflower, cereals (wheat and barley) and soybean as the crops of interest.[9] For each crop cultivated in 2017 or 2018, we identify the immediate predecessor crop from this set plus an 'other' category and, for the 2018 data with mono-cropping (i.e., cultivation of the same crop in 2017) also whether the pixel was mono cropped in 2016 as well. Overlaying the resulting raster file with boundaries for some 7,600 rural village councils (VC) from the State Service of Geodesy and Cadaster (SGC) allows us to obtain the total area with different rotational practices at VC level which will be a key dependent variable in the regressions reported below.

We use statistical and remotely sensed information as our two outcome variables. Information on area cultivated and output by crop is collected annually by the State Statistics Service of Ukraine (SSSU). In contrast to more detailed information on input use and output prices that is only available for a subset of producers in a way that makes it impossible to infer the actual location where production took place,[10] these data are obtained from the universe of registered producers and reported by village councils where production took place. Aggregation and merging these data with rotational information from satellite imagery allows us to add information on village level yields for the crops of interest as key dependent variables in our regressions.

As statistical data may still be affected by biases or measurement error, we complement statistical yield data with field-level indices of vegetative health and soil water content for pixels with different crop rotations as alternative outcome variables. To do so, we consider the Normalized Difference Vegetation Index (NDVI) and Enhanced Vegetation Index (EVI) at 250m and the Leaf Area Index (LAI), Fraction of Absorbed Photosynthetically Active Radiation (FAPAR) and Land Surface Water Index (LSWI) at 500m resolution from MODIS data.[11] To avoid unequal number of observations due to clouds or shadows, a median filter was applied to produce 16-day composites for each of the indices that were then summed over relevant crops' vegetation periods. The crop classification map was then converted to the same spatial resolution as these indices and for each VC, crop, and index, the overlap

[9] As budget-related factors precluded collection of enough training data in all years to distinguish wheat and barley in the imagery, we aggregate the two and note that, based on statistical data reported below, an increasing share of cereal is devoted to wheat.

[10] The data used come from 'form F29'. Detailed information on inputs and outputs is supposed to be collected from a sample of those farming fewer than 200 ha (or a threshold of value added) via 'form F2' and for the universe of farms above 200 ha via 'form F50'. As these data are collected at the enterprise level and as even for those farming only one operational unit, fields are generally dispersed across several VCs, matching becomes much more difficult. See Deininger *et al.* (2018) for a more detailed description of the data.

[11] As fields are relatively large and to ease computation burden and capitalize on higher temporal frequency of MODIS as compared to Sentinel, these indicators were produced at the field level by aggregating daily MODIS data.

between crop mask and the accumulated index was then used as a second set of dependent variables. Two advantages of this procedure are that it covers areas cultivated by non-registered enterprises and that it can be applied even in settings where statistical yield data may be entirely unavailable.

3.2 Descriptive statistics

Table 1 combines data on area cropped from form 29 in 2018 (column 2) with the same information for 2005 (column 1). In 2018, of 27.7 mn. ha of cultivated land captured by the statistical system, 33% was under wheat and barley, 22% under sunflower, 17% under maize and 6% under soybean. Compared to 2005, area expanded by some 3.6 mn. ha and crop composition shifted with quadrupling of the area under soybean, tripling of maize area, and a 75% increase in sunflower. While wheat area remained constant, area under barley declined by 40% and that under root and fodder crops by 30%. These shifts were partly policy-induced as in the case of sunflower for which Ukraine introduced 23% export duty in 1999 that was gradually reduced to 10% in 2014. The associated expansion of sunflower crushing capacity helped Ukraine to develop into the leading global supplier of sunflower oil , creating demand for local sunflower seed and resulting in the area expansion visible in the data. Data on yield, input cost and imputed profit from form 50, an expanded questionnaire for the subset of farms greater than 200 ha, suggest this reflected a move towards relatively more profitable crops. Imputed profits from maize and sunflower are 2.4 and 2.2 times that of barley while soybean and wheat profits exceed those from barley by 56% and 22%.

To assess implications for crop rotation, table 2 combines remotely sensed with statistical information for the four main crops in 2017 and 2018. In 2017, some 20% of cereals, 10% of sunflower, 24% of maize, and 17% of soybean was mono cropped (i.e., grown on the same field where the same crop had also been cultivated in 2016) while figures for 2018 are 17%, 10%, 28%, and 15%, respectively. Availability of data on the last two crops grown in 2018 suggest that 12% and 4.6% of cereal area were cropped in two or three consecutive seasons with analogous figures of 7.4% and 2.5% for sunflower, 19% and 9% for maize, and 11% and 4% for soybean, respectively.

Based on the 2018 figures, cereals were most often preceded by sunflower (33%), others (26%), cereals (17%), maize (14%) or soybean (10%). Most of the area planted to sunflower was previously occupied by cereals (39%) or maize (29%), other crops (15%), sunflower (10%) or soybean (7%). For maize, mono-cropping (28%; 19% with maize grown consecutively for two years and 9% for three years) is the most frequent sequence, followed by rotations where maize was preceded by sunflower (24%), cereals (20%), others (19%) or soybean (9%). Soybean was most often preceded by maize (28%), cereals and others (21% each), sunflower (15%) or itself (15%). In most cases, cultivated area as determined by remote sensing exceeds that reported by state statistics data, with a mean difference of some 30%. This is plausible as these data are known to exclude farms not formally registered as well as household plots. With 40, 25, 73 and 24 dt/ha for cereals, sunflower, maize and soybean, respectively, yields in 2018 were uniformly above those in 2017, a year in which precipitation was well below the long-term average.

3.3 Econometric approach

Indexing crops by j and village councils by v, we characterize a field's crop rotation history by the crops that precede crop i to assess the impact of different predecessor crops on yield of this crop via an OLS regression of the form

$$\log(Y_{ivt}) = \alpha + \sum_{j \neq i} \beta_{j,t-1} S_{jv,t-1} + \gamma_{i,t-1} M_{iv,t-1} + \gamma_{i,t-2} M_{iv,t-2} + \delta_i \log(F_{ivt}) + \varepsilon_{ivt} \quad (1)$$

where Y_{ivt} is the yield of crop i in village council v at time t, $S_{jv,t-1}$ is the share of area planted with crop i in t that was under predecessor crop j in t-1; $M_{iv,t-1}$ and $M_{iv,t-2}$ are the share of area with crop i that had been cropped with the same crop consecutively for two or three years; F_{ivt} is the average farm size for crop i in village v to adjust for suitability and ε_{ivt} is a random error term. Coefficients of interest are $\beta_{j,t-1}$, the effect of having crop i preceded by j as well as $\gamma_{i,t-1}$ and $\gamma_{i,t-2}$, the effect of planting the same crop i for two and three consecutive years, respectively. Parameters are estimated separately for 2017 and 2018. As data on crop cover are available only for 2016-2018, the effect of monocropping is estimated for two consecutive years in 2017 and for two and three years together in 2018. Indices for vegetative cover (EVI, LAI, FAPAR) and soil water content (LSWI) for the areas covered by relevant crops constructed as described above from remotely sensed data are used as alternative left-hand side variables. For both yields and these indices, (1) is estimated separately for maize, soybean, cereals (wheat and barley) and sunflower.

4. Results

Results from yield regressions support the notion of continuous cultivation of the same crop having negative effects but, contrary to findings in the literature for (largely) US-based rotations, suggest that having sunflower as a predecessor crop leads to worse outcomes. For maize and sunflower, estimated marginal effects of three-year monocropping are significantly different from those of two-year monocropping. Regressions with indices of vegetative development or soil water content (based on remotely sensed imagery) as dependent variables yield similar results, allaying concerns about our results being driven by shortcomings in statistical data and suggesting that exploration of such indices' suitability for broader analysis may be warranted.

4.1 Yield effects of crop rotation

Regression results to assess the impact of different predecessor crops in two-year and three-year sequences in tables 4 and 5 highlight that in virtually all cases, crop rotation variables are highly significant and explain one-quarter to an eighth of variation in the data. Table 4 presents results for maize (col. 1-3) and soybean (col. 4-6) in the unusually dry season of 2016/17 (cols. 1 and 4), the 2017/18 season (cols. 2 and 5), and all three years (cols 3 and 6), with coefficients in the top and elasticities in the bottom panel. Having soybean precede itself once or twice is estimated to be no worse than having maize (the excluded category), possibly because of crop residue management. The most negative predecessor crop to soybean is estimated to be sunflower with a coefficient of -0.67 in 2016/17 and -0.44 overall, possibly because both crops are susceptible to stem rot (*sclerotina*), a fungal disease. A more pronounced negative effect in dry years (2016/17) may be due to sunflower's highly developed root system leaving little moisture for successor crops.[12] By comparison, having cereals precede soybean has an insignificant effect in 2016/17 and a slightly negative effect with a marginal effect of -0.11 overall. The inconsistency in the effect of other crops on soybean yield between 2016/17 (a statistically positive effect with a coefficient of 0.2) and insignificant negative effect overall could likely be due to the change in the composition other crops (i.e., consisting of crops other than maize, soybean, sunflower, wheat and barley) between the two periods.

With a coefficient of -0.38 in 2016/17 and -0.37 in 2017/18, there is evidence of significant self-incompatibility in maize. Moreover, with -0.51, the coefficient of having maize thrice in succession is statistically significantly different from that for a two-year maize sequence with estimated coefficient of -0.29, suggesting that in the case at hand, the absolute value for the impact of having maize planted three years in a row is significantly larger than that of maize in two-year succession. What is more surprising is that, compared to having soybean as predecessor crop (the excluded category), having maize preceded by other crops (-0.80 in 2016/17 and -0.58 in 2017/18), cereals (-0.52 in 2016/17 and -0.97 in 2017/18), or sunflower (-1.41 in 2016/17 and -1.16 in 2017/18) is estimated to have a more negative impact than a maize mono-crop, in line with the notion that, under conditions where water availability is a key constraint, sunflower's high levels of water extraction may undermine subsequent crops' possibilities for development.

Results for sunflower in table 5 suggest that, compared to soybean (the excluded category), maize is a preferable predecessor crop in normal years such as 2017/18 (a coefficient of 0.13) but not in dry

[12] Extension advice highlights that sunflower grows well under dry conditions because, with a root depth of more than 2 meters, well in excess of other agricultural crops, it can effectively utilize subsoil moisture. This high level of water extraction implies that the soil is depleted of water for the subsequent crop. It is thus recommended to have deep-rooted crops such as sunflower follow small grain and that in dry years there may need to be followed by fallow. Crop scientists suggest a minimum four-year rotation cycle that includes sunflower, e.g., a spring wheat-dry pea-barley-sunflower rotation where the broadleaf crops help boost yield of subsequent small grains, and alternating legumes and oilseed crops avoids build-up of plant diseases (National_Sunflower_Association 2013).

ones such as 2016/17 (-0.14) that are unfavorable to development of fungi and the associated spread of stem rot so that soybean's ability to fix nitrogen outweighs its role as potential disease vector. While other crops have a slightly larger negative effect, monocropping of sunflower is estimated to have the most negative effects with coefficients of -0.94 in 2016/17 when, in addition to disease permanence, water access is likely to have acted as a constraint, and -0.61 in 2017/18. With coefficients of -0.45 and -0.90 a three-year sunflower succession has a more negative impact than a two-year one.

As limited training data made it impossible to distinguish between wheat and barley as the main cereal crops in satellite imagery, cereals' high levels of self-incompatibility, with coefficients of -0.30 and -0.79 for two-year cereal successions in 2016/17 and 2017/18, respectively are only slightly above those for maize (-0.22 and -0.19 in the two periods) and other crops (-0.29 and -0.51), reflecting the positive contribution of soybeans (the excluded category) as a predecessor crop due to N fixation. With -0.83, the estimated coefficient on a three-year cereal monocrop is not significantly different from that on a two-year sequence (-0.76).

4.2 Using alternative outcome variables

Vegetation indices based on remote sensing data can complement statistical data two ways. First, rather than having one outcome variable (yield), they offer additional variables such as soil water content or leaf area at much higher temporal resolution. This allows to take corrective action in near real time[13] and to assess channels for observed effects *ex post*. Second, by providing an independent source of information, they can allay concerns about bias in statistical data that may result from mis-reporting to statistical agencies (e.g., to avoid taxes); partial coverage such as the omission of non-registered or household farms in the case of Ukraine; or other types of measurement error.[14]

To address these issues within the confines of this paper,[15] we complement yield analysis with an aggregate analysis where the means/sums of four key vegetation indices over relevant crops' entire vegetation period is used as outcome variables in regression equation (1). These indices, widely used in the literature and practical work are the enhanced vegetation index (EVI), the leaf area index (LAI), fraction of photosynthetically active radiation (FAPAR), and the Land Surface Water Index (LSWI).

The EVI is based on the normalized difference vegetation index (NDVI) that is widely used in the literature and agronomy but minimizes canopy-soil variations and improves sensitivity in dense vegetations. The

[13] While several firms provide such services to large farms in Ukraine, they generally use satellite imagery to identify areas at risk and complement it with drone imagery or manual scouting.
[14] In our case, inability to match all the village councils is a relevant concern in this respect.
[15] While detailed discussion of temporal variation in these indices together with detailed climatic data would be of great interest to provide insights on the mechanisms through which crop rotation and other factors affect yields, it is well beyond the scope of this paper.

LAI, defined as the green leaf area per ground area is closely related to FAPAR, the fraction of radiation in the 400−700 nm spectrum that is absorbed by green elements of the vegetation canopy and thus available for photosynthetic activity by plants. Finally, the LSWI uses the shortwave infrared (SWIR) and near infrared (NIR) bands that are strongly absorbed by liquid water to provide a measure of the total amount of liquid water in vegetation and its soil background that will be relevant to check if excessive extraction of soil water by deep-rooted sunflowers is a plausible be mechanism for sunflower's negative impact on successor crops' yield as observed above.

As illustrated in table 6, crop rotation characteristics for maize based on data from 3 seasons explain 30% to 40% of the variation in vegetation indices. Results are consistent with what was obtained for yield in that sunflower is by far the worst predecessor crop for maize with an outsized effect (compared to that of other crops) especially on water availability as measured by the LSWI. The relatively better performance of other crops as compared to cereals may be attributable to their positive impact (compared to soybean as omitted category with which they may also share N-fixing properties) on water balances both in dry and in average years. Insignificant differences in terms of water balances between maize and soybean (with a negative effect only for maize being cultivated three times in a row) may also partly explain that the negative impact of continued maize cultivation is relative limited.

Results for vegetation indices for areas covered with soybean in table 7 reinforce the notion of soybean being a better successor crop than maize for other crops (where coefficients are positive and significant for all four indices) and itself in dry seasons (all indices for 2016/17 are positive and significant with the exception of the positive and insignificant LSWI). Cereals as predecessor crop have a positive impact on LSWI as they extract limited amounts of water but a negative one on LAI and, in 2018, FAPAR and EVI. The estimated magnitude of this effect on FAPAR and EVI (and to a lesser extent LAI), though remains modest compared to the negative and significant effect of sunflower not only on vegetation indices, likely a result of high susceptibility of soybean and sunflower (but not maize) to stem rot and on water availability, a result of sunflowers' deep root system that reduces moisture availability for successor crops.

Table 8 illustrates that, consistent with yield results and negative impacts as a predecessor for other crops as documented above, sunflower is also a bad successor to itself . Coefficients on the share of area where the crop is grown twice or thrice in a row for all four indices considered are negative, significant at 1%, and in most cases much larger than those for cereals, as the predecessor crop with the second most negative effect. Estimated impacts of growing sunflower thrice are statistically significantly larger than those of growing it twice in a row with the difference particularly large for the LSWI where the size of the negative impact of a triple sunflower rotation is estimated to be more than double that of a double sunflower rotation. Interestingly, preceding sunflower with other crops is estimated to have a positive effect (compared to soybean as the omitted category) not only on LSWI but, in 2018 (i.e., a year with average levels of precipitation), also on FAPAR and EVI. Maize also exhibits only a marginally negative effect on LSWI in 2018 suggesting that, albeit inferior to other crops and soybean (or sunflower itself), it is preferable to cereals as a predecessor crop for sunflower.

Finally, table 9 suggests that the impact of growing cereals twice or thrice in succession on indices of vegetative development is slightly worse than that of maize while sunflower is the worst predecessor crop, consistent with results from yield regressions. Surprisingly, estimated impacts of double cropping cereals on soil water content in 2017 are positive and significant. Spatial and temporal disaggregation of vegetation indices could be used to determine whether this can be attributed to drought-induces loss of (spring) cereals having led to *de facto* fallowing of parcels at a large scale with subsequent seeding of winter cereals.

5. Conclusion and implications for research

Use of machine learning techniques on remotely sensed data and their combination with statistical information allows us to analyze crop rotation effects in a real world rather than experimental setting. The findings of statistically significant and economically meaningful effects that are consistent with agronomic factors such as different crops' ability to capture soil water and their susceptibility to certain diseases are encouraging. Given this potential, high-quality training data seem to be a key public good and approaches to encourage and systematize their collection and sharing will be a key element in advancing the global research agenda.

For the specific case of Ukraine, our results open opportunities for policy-relevant research in three directions: First, further disaggregation of effects by variation in climate and type of producer and of variables based on remote sensing over the agricultural season and the country's territory could yield valuable insight and help assess whether the effects described here suffer from downward bias due to mitigation (e.g., applying more fertilizer). Second, quantifying rotation effects on profitability could help understand inherent trade-offs that could be used to design the government's agricultural support programs in ways to facilitate adoption of more diverse rotations. Finally, it will be of great interest to explore policy effects on rotational practices and similar types of non-contractible investment by exploiting variation in policy over time and space.

Table 1: Changes in cropped area, Ukraine 2004-2018 as well as yields and output prices in 2018

	Cultivated land F29 (millions of ha)		Three-year average F50 (2016-2018)			
			Yield	Price	Cost	Profit
	2005	2018	dt/ha	UAH/Q	UAH/ha	UAH/ha
Wheat	6.36	6.61	43	403	12,044	4,095
Barley	4.20	2.49	37	390	10,302	3,346
Maize	1.45	4.58	75	379	18,459	8,180
Sunflower	3.51	6.11	24	896	13,128	7,426
Soybeans	0.42	1.72	22	966	14,784	5,207
Roots, veg. & forage	5.19	3.59				
Total agricultural crops	24.07	27.68				

Source: Own computation from form 29 and 50 of state statistics database

Table 2: Key rotation characteristics for sunflower, maize, and soybeans

Crop	**Cereals**[a]		**Sunflower**		**Maize**		**Soybean**	
	2017	**2018**	**2017**	**2018**	**2017**	**2018**	**2017**	**2018**
Panel A: Remote sensing								
Share following itself	0.199	0.166	0.107	0.100	0.240	0.278	0.171	0.147
monocrop 2017/2018		0.120		0.074		0.192		0.112
monocrop 2016/17/18		0.046		0.025		0.086		0.035
same crop 2016/2018		0.337		0.262		0.152		0.129
Sunflower	0.344	0.331			0.249	0.241	0.136	0.150
Maize	0.102	0.142	0.243	0.287			0.280	0.284
Soybeans	0.128	0.095	0.073	0.075	0.086	0.088		
Cereals			0.426	0.391	0.247	0.201	0.215	0.207
Others	0.227	0.265	0.151	0.148	0.178	0.191	0.197	0.212
Area (ha/VC)	1166	1156	1240	1081	844	715	356	425
Panel B: Statistics								
Yield (dt/ha)	34.99	40.28	21.58	24.91	50.47	72.83	17.15	23.83
Area per farm (ha)	187.91	189.85	319.58	321.97	301.84	312.15	202.74	190.83
Area statistics (F29)	867	894	832	838	604	629	382	345
Number of VCs	6,540	6,610	5,356	5,495	4,767	4,785	3,811	3,654

[a]Yield for cereals is calculated as value of wheat and barley per unit of land divided by average price.

Source: Own computation from remotely sensed data for 2016, 2017, and 2018, and form 29 of state statistics database for yield and mean farm size. See text for further description.

Table 3: Transition matrices by major crop types for 2017 and 2018

Panel A: 2016/17

	2017				
2016	Sunflower	Maize	Cereals	Soybean	Total
Sunflower	2.29	1.14	3.71	0.64	7.78
Maize	2.21	2.38	0.79	1.06	6.44
Cereals	6.01	1.74	2.56	0.65	10.95
Soybean	0.42	0.83	0.69	0.67	2.62
Total	10.93	6.09	7.76	3.02	27.79

Panel B: 2017/18

	2018				
2017	Sunflower	Maize	Cereals	Soybean	Total
Sunflower	1.78	1.80	5.05	0.77	9.41
Maize	2.35	2.36	1.00	1.34	7.05
Cereals	5.18	1.26	2.58	0.83	9.84
Soybean	0.48	0.74	0.63	0.65	2.50
Total	9.79	6.15	9.26	3.59	28.79

Source: Own computation based on remote sensing-based crop classification as described in text.

Table 4: Estimated yield effects of different predecessor crops for maize and soybeans

	Maize			Soybeans		
	2016/17	**2017/18**	**2016-18**	**2016/17**	**2017/18**	**2016-18**
% after sunflower	-1.405***	-1.158***	-1.161***	-0.665***	-0.440***	-0.441***
	(0.0713)	(0.0606)	(0.0606)	(0.0574)	(0.0492)	(0.0493)
% after cereals	-0.519***	-0.974***	-0.978***	-0.0436	-0.105**	-0.106**
	(0.0741)	(0.0634)	(0.0634)	(0.0493)	(0.0450)	(0.0450)
% after others	-0.796***	-0.583***	-0.585***	0.202***	-0.0155	-0.0165
	(0.0811)	(0.0677)	(0.0676)	(0.0572)	(0.0458)	(0.0459)
% self twice	-0.379***	-0.374***		-0.0441	0.0188	
	(0.0798)	(0.0634)		(0.0590)	(0.0525)	
% self twice only			-0.293***			0.0379
			(0.0699)			(0.0633)
% self thrice			-0.510***			-0.0350
			(0.0802)			(0.112)
No. of VCs	4,767	4,785	4,785	3,811	3,654	3,654
R^2	0.224	0.274	0.275	0.110	0.0875	0.0875
Mean dep. var.	3.737	4.144	4.144	2.677	3.071	3.071
SD dep. var.	0.663	0.601	0.601	0.635	0.485	0.485
Elasticity: twice	-0.0909	-0.104	-0.0563	-0.00756	0.00276	0.00424
Elasticity: thrice			-0.0439			-0.00123
El after sunflower	-0.350	-0.279	-0.279	-0.0908	-0.0662	-0.0663
F-test						
Self 2x=self 3x			7.61***			0.29

Note: Results are from regressions at village council level for two-year periods in columns 1 and 4 (2016-17) as well as 2 and 5 (2017-18) and for the 3-year period 2016-18 in columns 3 and 6. Dependent variable is log of physical yield (dt/ha). The soybean-maze and maize-soybean rotations are the omitted categories, respectively. A constant as well as mean farm size is included throughout but omitted. Standard errors in parentheses: * p<0.10, ** p<0.05, *** p<0.010.

Table 5: Estimated yield effects of different predecessor crops for cereals and sunflower

	Cereals			Sunflower		
	2016/17	**2017/18**	**2016-18**	**2016/17**	**2017/18**	**2016-18**
% after sunflower	-0.484***	-0.705***	-0.705***			
	(0.0257)	(0.0352)	(0.0352)			
% after maize	-0.221***	-0.199***	-0.197***	-0.141***	0.126**	0.130***
	(0.0439)	(0.0435)	(0.0436)	(0.0532)	(0.0491)	(0.0491)
% after cereals				-0.330***	-0.360***	-0.356***
				(0.0473)	(0.0451)	(0.0451)
% after others	-0.294***	-0.509***	-0.507***	-0.469***	-0.290***	-0.288***
	(0.0318)	(0.0376)	(0.0377)	(0.0597)	(0.0535)	(0.0535)
% self twice	-0.299***	-0.786***		-0.936***	-0.607***	
	(0.0319)	(0.0434)		(0.0608)	(0.0574)	
% self twice only			-0.762***			-0.445***
			(0.0538)			(0.0752)
% self thrice			-0.831***			-0.902***
			(0.0748)			(0.105)
No. of VCs	6,540	6,610	6,610	5,356	5,495	5,495
R^2	0.199	0.168	0.168	0.118	0.142	0.144
Mean dep. var.	3.557	3.486	3.486	2.962	3.104	3.104
SD dep. var.	0.412	0.454	0.454	0.501	0.463	0.463
Elasticity: twice	-0.0595	-0.131	-0.0917	-0.100	-0.0604	-0.0331
Elasticity: thrice			-0.0382			-0.0228
El after sunflower	-0.166	-0.234	-0.234			
F-test						
Self 2x=self 3x			0.54			11.11***

Note: Results are from regressions at village council level for two-year periods in columns 1 and 4 (2016-17) as well as 2 and 5 (2017-18) and for the 3-year period 2016-18 in columns 3 and 6. Dependent variable is log of the output value of cereals (wheat and barley) per ha in columns 1-3 and the log of physical yield (dt/ha) in columns 3-6. The soybean-sunflower rotation the omitted category. A constant as well as mean farm size is included throughout but omitted. Standard errors in parentheses: * p<0.10, ** p<0.05, *** p<0.010.

Table 6: Vegetation indices and maize crop rotation in Ukraine, 2016-18 period

	EVI		LAI		FAPAR		LSWI	
	2017	**2018**	**2017**	**2018**	**2017**	**2018**	**2017**	**2018**
After sunflower	-1.967***	-2.154***	-51.73***	-53.14***	-9.450***	-9.866***	-1.218***	-1.483***
	(0.0738)	(0.0673)	(1.867)	(1.805)	(0.361)	(0.326)	(0.0967)	(0.0817)
After cereals	-0.857***	-1.697***	-26.76***	-45.67***	-3.915***	-7.300***	-0.0822	-1.287***
	(0.0770)	(0.0705)	(1.950)	(1.889)	(0.376)	(0.341)	(0.101)	(0.0856)
After other	-0.758***	-0.0289	-33.19***	-14.19***	-2.866***	0.630*	0.569***	0.921***
	(0.0871)	(0.0770)	(2.208)	(2.063)	(0.426)	(0.373)	(0.114)	(0.0934)
Self twice	-0.476***		-12.09***		-1.821***		0.0424	
	(0.0829)		(2.097)		(0.405)		(0.109)	
Self twice only		-0.479***		-7.894***		-1.758***		0.0385
		(0.0782)		(2.097)		(0.379)		(0.0950)
Self thrice		-0.498***		-4.844**		-1.873***		-0.228**
		(0.0887)		(2.378)		(0.429)		(0.108)
No. of VCs	4553	4570	4547	4563	4547	4563	4548	4563
R^2	0.253	0.426	0.263	0.379	0.264	0.418	0.144	0.338
Mean dep.	4.207	4.438	65.43	69.86	22.81	23.52	-0.172	0.414
SD dep..	0.678	0.732	17.29	18.84	3.339	3.517	0.831	0.827
Elast. twice	-0.0275		-0.0450		-0.0194		0.0601	
Elast. twice		-0.0209		-0.0219		-0.0145		0.0180
Elast. thrice		-0.00987		-0.00611		-0.00701		-0.0486
El a sunflower	-0.119	-0.118	-0.201	-0.185	-0.105	-0.102	-1.806	-0.873
F-test								
Self 2x=self 3x		0.048		1.700		0.786		6.338**

Note: Results are from regressions at village council level and the soybean-maize rotation is the omitted category. A constant is included throughout but omitted. Standard errors in parentheses: * p<0.10, ** p<0.05, *** p<0.010.

Table 7: Vegetation indices and soybean crop rotation in Ukraine, 2016-18 period

	EVI		LAI		FAPAR		LSWI	
	2017	**2018**	**2017**	**2018**	**2017**	**2018**	**2017**	**2018**
After sunflower	-1.082***	-1.415***	-28.75***	-24.20***	-4.933***	-5.189***	-0.911***	-0.927***
	(0.0592)	(0.0600)	(1.421)	(1.500)	(0.276)	(0.271)	(0.0835)	(0.0743)
After cereals	0.0479	-0.568***	-6.243***	-15.06***	-0.0661	-1.981***	0.301***	-0.220***
	(0.0545)	(0.0546)	(1.312)	(1.366)	(0.254)	(0.247)	(0.0771)	(0.0676)
After other	1.224***	0.944***	5.303***	7.335***	4.769***	5.111***	2.107***	1.826***
	(0.0683)	(0.0560)	(1.640)	(1.397)	(0.318)	(0.252)	(0.0964)	(0.0692)
Self twice	0.212***		3.665**		1.262***		0.122	
	(0.0627)		(1.505)		(0.292)		(0.0884)	
Self twice only		-0.160**		-1.451		-0.387		-0.264***
		(0.0770)		(1.926)		(0.348)		(0.0954)
Self thrice		0.169		11.94***		0.873		-0.0813
		(0.136)		(3.388)		(0.612)		(0.168)
No. of VCs	3146	3394	3129	3380	3129	3380	3129	3381
R^2	0.223	0.322	0.148	0.152	0.196	0.309	0.201	0.314
Mean dep.	4.467	4.895	70.63	78.41	23.79	25.08	-0.0188	0.745
SD dep..	0.646	0.644	14.77	14.36	2.949	2.873	0.897	0.791
Elast. twice	0.00838		0.00917		0.00937		1.143	
Elast. twice		-0.00368		-0.00207		-0.00173		-0.0397
Elast. thrice		0.00124		0.00547		0.00125		-0.00393
El a sunflower	-0.0362	-0.0430	-0.0608	-0.0460	-0.0310	-0.0308	-7.227	-0.185
F-test								
Self 2x=self 3x		3.987**		10.56***		2.869*		0.802

Note: Results are from regressions at village council level and the maize-soybean rotation is the omitted category. A constant is included throughout but omitted. Standard errors in parentheses: * p<0.10, ** p<0.05, *** p<0.010.

Table 8: Vegetation indices and sunflower crop rotation in Ukraine, 2016-18 period

	EVI		LAI		FAPAR		LSWI	
	2017	**2018**	**2017**	**2018**	**2017**	**2018**	**2017**	**2018**
After maize	-0.652***	-0.402***	-16.42***	-9.266***	-3.250***	-2.309***	-0.434***	-0.159*
	(0.0661)	(0.0652)	(1.563)	(1.595)	(0.321)	(0.313)	(0.0919)	(0.0872)
After cereals	-1.119***	-1.454***	-29.86***	-38.77***	-5.344***	-6.658***	-0.605***	-1.247***
	(0.0588)	(0.0598)	(1.389)	(1.462)	(0.285)	(0.287)	(0.0817)	(0.0799)
After other	-0.187**	0.149**	-13.67***	-5.705***	-0.107	1.469***	0.789***	0.971***
	(0.0771)	(0.0749)	(1.818)	(1.833)	(0.373)	(0.360)	(0.107)	(0.100)
Self twice	-2.530***		-67.10***		-12.90***		-1.916***	
	(0.0740)		(1.745)		(0.358)		(0.103)	
Self twice only		-1.986***		-51.89***		-9.586***		-1.171***
		(0.102)		(2.496)		(0.490)		(0.136)
Self thrice		-2.651***		-81.42***		-15.16***		-2.552***
		(0.138)		(3.363)		(0.660)		(0.184)
No. of VCs	5060	5211	5049	5200	5049	5200	5049	5200
R^2	0.280	0.409	0.323	0.452	0.305	0.413	0.140	0.302
Mean dep.	3.979	4.190	58.61	61.41	21.31	21.89	-0.445	0.0819
SD dep.	0.640	0.695	15.50	17.57	3.138	3.333	0.808	0.852
Elast. twice	-0.0700		-0.126		-0.0667		-0.475	
Elast. twice		-0.0351		-0.0625		-0.0324		-1.058
Elast. thrice		-0.0164		-0.0345		-0.0180		-0.809
F-test								
Self 2x=self 3x		13.08***		43.26***		40.00***		31.60***

Note: Results are from regressions at village council level and the soybean-sunflower rotation is the omitted category. A constant is included throughout but omitted. Standard errors in parentheses: * p<0.10, ** p<0.05, *** p<0.010.

Table 9: Vegetation indices and cereals crop rotation in Ukraine, 2016-18 period

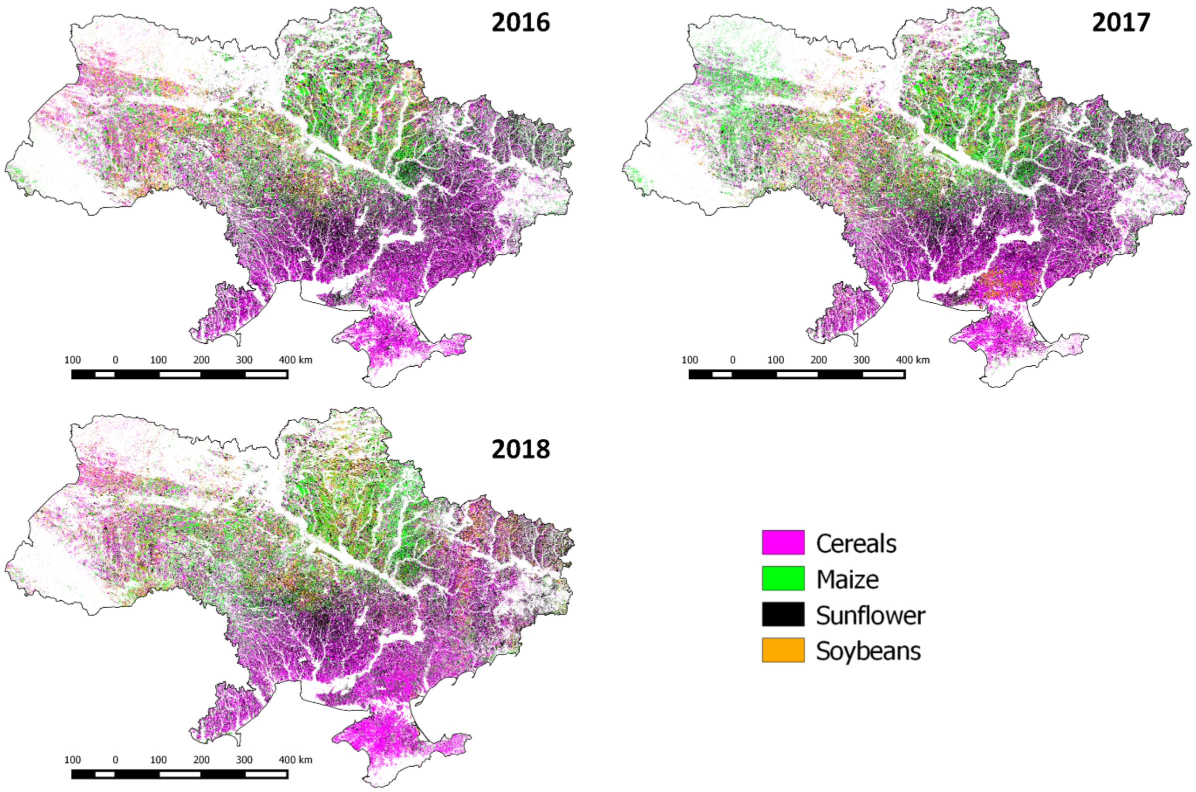
	EVI		LAI		FAPAR		LSWI	
	2017	**2018**	**2017**	**2018**	**2017**	**2018**	**2017**	**2018**
After sunflower	-1.099***	-1.264***	-32.30***	-36.54***	-4.714***	-5.828***	-0.520***	-1.194***
	(0.0312)	(0.0388)	(0.792)	(0.975)	(0.138)	(0.161)	(0.0413)	(0.0576)
After maize	-0.174***	-0.0497	-8.439***	-11.74***	-0.967***	-1.676***	-0.0535	-0.278***
	(0.0545)	(0.0480)	(1.379)	(1.208)	(0.242)	(0.199)	(0.0723)	(0.0714)
After other	0.0147	0.274***	-10.93***	-6.083***	-0.193	-0.440**	0.598***	0.348***
	(0.0394)	(0.0418)	(1.003)	(1.054)	(0.176)	(0.174)	(0.0527)	(0.0623)
Self twice	-0.231***		-12.15***		-1.088***		0.440***	
	(0.0392)		(0.992)		(0.174)		(0.0520)	
Self twice only		-1.027***		-31.18***		-4.647***		-1.283***
		(0.0597)		(1.503)		(0.248)		(0.0889)
Self thrice		-0.796***		-32.52***		-4.253***		-2.539***
		(0.0834)		(2.105)		(0.347)		(0.124)
No. of VCs	6436	6531	6409	6486	6381	6486	6381	6487
R^2	0.298	0.431	0.287	0.359	0.273	0.369	0.155	0.278
Mean dep.	3.504	3.482	53.92	52.63	18.49	17.30	0.319	1.401
SD dep..	0.535	0.599	13.45	14.17	2.322	2.352	0.644	0.789
Elast. twice	-0.0131		-0.0450		-0.0117		0.274	
Elast. Twice		-0.0354		-0.0712		-0.0323		-0.110
Elast. Thrice		-0.0105		-0.0285		-0.0113		-0.0834
Elast. a sunflower	-0.109	-0.121	-0.207	-0.232	-0.0890	-0.113	-0.569	-0.285
F-test								
Self 2x=self 3x		4.954**		0.261		0.838		65.83***

Note: Results are from regressions at village council level and the soybean-cereals (wheat and barley) rotation is the omitted category. A constant is included throughout but omitted. Standard errors in parentheses: * p<0.10, ** p<0.05, *** p<0.010.

Figures

Figure 1: Routes for in-situ data collection in 2016-2018

Figure 2: Crop cover maps 2016-2018



Appendix table 1: Distribution of in-situ data and accuracy of prediction by land cover type

Land cover type	Field number			Accuracy of prediction		
	2016	**2017**	**2018**	**2016**	**2017**	**2018**
Artificial	135	135	328	75	70.3	76.3
Bare land	114	108	161	60	60.8	76.9
Cereals	1819	1172	2292	60.4	96.9	92.9
Forest	676	683	1331	98.1	98.7	99.1
Grassland	798	544	1301	80.2	84.3	84.7
Maize	1027	688	972	93	92.1	60.6
Peas	32	49	76	70.9	91.8	91.8
Soybeans	545	381	649	82.5	81.6	31
Sugar beet	59	39	138	93.6	98.6	63.8
Sunflower	1645	1191	1751	94.3	95.6	78.6
Water	176	192	480	88.6	99.2	99.4
Wetland	48	70	132	77.3	62.5	78.7
Winter rape	188	213	391	83.6	97	95.2
Total	7,989	5,465	10,210	88.3	91	85

References:

Archer, D. W., M. A. Liebig, D. L. Tanaka and K. P. Pokharel. 2020. "Crop Diversity Effects on Productivity and Economics: A Northern Great Plains Case Study." *Renewable Agriculture and Food Systems*, ***35***(1), 69-76.
Arnade, C. and L. Hoffman. 2019. "Determination of the International Maize Price: What Is the Growing Role of Brazil and Ukraine?" *Agricultural Economics*, ***50***(6), 735-747.
Behnke, G. D., et al. 2018. "Long-Term Crop Rotation and Tillage Effects on Soil Greenhouse Gas Emissions and Crop Production in Illinois, USA." *Agriculture, Ecosystems & Environment*, ***261***, 62-70.
Boryan, C., Z. Yang, R. Mueller and M. Craig. 2011. "Monitoring Us Agriculture: The Us Department of Agriculture, National Agricultural Statistics Service, Cropland Data Layer Program." *Geocarto International*, ***26***(5), 341-358.
Bowles, T. M., et al. 2020. "Long-Term Evidence Shows That Crop-Rotation Diversification Increases Agricultural Resilience to Adverse Growing Conditions in North America." *One Earth*, ***2***(3), 284-293.
Boyabatli, O., J. Nasiry and Y. Zhou. 2019. "Crop Planning in Sustainable Agriculture: Dynamic Farmland Allocation in the Presence of Crop Rotation Benefits." *Management Science*, ***65***(5), 2060-2076.
Brunt, L. 1999. "An Arbitrage Model in Crop Rotation in 18th Century England." University of Oxford, Department of Economics, Economics Series Working Papers: 1999-W32.
Castellazzi, M. S., et al. 2008. "A Systematic Representation of Crop Rotations." *Agricultural Systems*, ***97***(1), 26-33.
Davis, A. S., et al. 2012. "Increasing Cropping System Diversity Balances Productivity, Profitability and Environmental Health." *PLOS ONE*, ***7***(10), e47149.
Deininger, K., D. Nizalov and S. K. Singh. 2018. "Determinants of Productivity and Structural Change in a Large Commercial Farm Environment: Evidence from Ukraine." *The World Bank Economic Review*, ***32***(2), 287-306.
Gentry, L., M. Ruffo and F. Below. 2013. "Identifying Factors Controlling the Continuous Corn Yield Penalty." *Agronomy Journal*, ***105***, 295.
Hendricks, N. P., A. Smith and D. A. Sumner. 2014. "Crop Supply Dynamics and the Illusion of Partial Adjustment." *American Journal of Agricultural Economics*, ***96***(5), 1469-1491.
Hennessy, D. A. 2006. "On Monoculture and the Structure of Crop Rotations." *American Journal of Agricultural Economics*, ***88***(4), 900-914.
Hilton, S., et al. 2013. "Impact of Shortened Crop Rotation of Oilseed Rape on Soil and Rhizosphere Microbial Diversity in Relation to Yield Decline." *PLOS ONE*, ***8***(4), e59859.
Koch, H.-J., K. Trimpler, A. Jacobs and N. Stockfisch. 2018. "Crop Rotational Effects on Yield Formation in Current Sugar Beet Production – Results from a Farm Survey and Field Trials." *Frontiers in Plant Science*, ***9***(231).
Krasnozhon, L. A. 2013. "Political Economy of Agricultural Market Reform in Ukraine: 'Good Bye Lenin'." *Journal of Private Enterprise*, ***29***(1), 119-140.
Krupinsky, J. M., et al. 2006. "Crop Sequence Effects of 10 Crops in the Northern Great Plains." *Agricultural Systems*, ***88***, 227-254.
Lark, T. J., R. M. Mueller, D. M. Johnson and H. K. Gibbs. 2017. "Measuring Land-Use and Land-Cover Change Using the U.S. Department of Agriculture's Cropland Data Layer: Cautions and Recommendations." *International Journal of Applied Earth Observation and Geoinformation*, ***62***, 224-235.
Livingston, M., M. J. Roberts and Y. Zhang. 2015. "Optimal Sequential Plantings of Corn and Soybeans under Price Uncertainty." *American Journal of Agricultural Economics*, ***97***(3), 855-878.
Meisner, M. H. and J. A. Rosenheim. 2014. "Ecoinformatics Reveals Effects of Crop Rotational Histories on Cotton Yield." *PLOS ONE*, ***9***(1), e85710-e85710.
Mitchell, P. D. 2011. "Economic Assessment of the Benefits of Chloro-S-Triazine Herbicides to U.S. Corn, Sorghum, and Sugarcane Producers." University of Wisconsin, Agricultural and Applied Economics, Staff Paper Series.
National_Sunflower_Association. 2013. "The Sunflower Production Guide." Mandan, ND.
Plourde, J. D., B. C. Pijanowski and B. K. Pekin. 2013. "Evidence for Increased Monoculture Cropping in the Central United States." *Agriculture, Ecosystems & Environment*, ***165***, 50-59.
Porter, P. M., et al. 1997. "Environment Affects the Corn and Soybean Rotation Effect." *Agronomy Journal*, ***89***(3), 442-448.
Qiao, F., J. Huang and X. Wang. 2017. "Fifteen Years of Bt Cotton in China: Results from Household Surveys." *World Development*, ***98***, 351-359.
Regis Mauri, G. 2019. "Improved Mathematical Model and Bounds for the Crop Rotation Scheduling Problem with Adjacency Constraints." *European Journal of Operational Research*, ***278***(1), 120-135.
Seifert, C. A., M. J. Roberts and D. B. Lobell. 2017. "Continuous Corn and Soybean Yield Penalties across Hundreds of Thousands of Fields." *Agronomy Journal*, ***109***(2), 541-548.
USDA. 2019. "2017 Census of Agriculture Highlights: Farms and Farmland." United States Department of Agriculture, National Agricultural Staistics Service, Washington DC.
Wang, H. and A. Ortiz-Bobea. 2019. "Market-Driven Corn Monocropping in the U.S. Midwest." *Agricultural and Resource Economics Review*, ***48***(2), 274-296.