Global Environmental Change 60 (2020) 101982

Contents lists available at ScienceDirect

Global Environmental Change

journal homepage: www.elsevier.com/locate/gloenvcha

Rural-urban connectivity and agricultural land management across the Global South

Fanny Boudet[a], Graham K. MacDonald[a,*], Brian E. Robinson[a], Leah H. Samberg[b,c]

[a] Department of Geography, McGill University, Montreal, Quebec, Canada
[b] Rainforest Alliance, New York, NY, USA
[c] Institute on the Environment, University of Minnesota, Minneapolis, MN, USA

ARTICLE INFO

Keywords:
Land use
Urbanization
Migration
Agriculture
Rural-urban linkages

ABSTRACT

Research on how urbanization affects rural agriculture has typically focused on loss of farmland due to urban expansion. However, more distal pathways that could link urbanization to rural agriculture, including enhanced connectivity through rural-urban migration and market access, remain poorly understood. Here, we assess whether greater rural-urban connectivity is associated with changes in agricultural land management across the Global South. Such associations are complex, and thus difficult to measure at this scale. We therefore take a two-step approach to investigate these relationships. First, using a multivariate clustering approach, we define a series of rural-urban connectivity typologies from existing spatial data on land use, demographics, rural market access, and rural population change (as a proxy for outmigration). We examine the variation in key agricultural outcome variables (mean cereal crop yields, % of attainable yields met, and cropping frequency) within the typologies, which shows that greater overall connectivity (market access and population change) is associated with higher cereal yields, yield attainment, and cropping frequency. Second, building on these clustering results, we develop hypotheses about the relationship between rural-urban connectivity and agricultural land use intensity. We then use propensity score matching to test these hypotheses by comparing locations with similar sociodemographic and land use characteristics. When controlling for gross domestic product (GDP) per capita, agricultural land, and population density, rural locations with relatively high market access, negative population change, and greater built-up area have significantly higher mean nitrogen application rates, irrigated areas, and cereal yields across the Global South. Results vary by region, but greater rural-urban connectivity and urban extents are generally associated with higher overall agricultural inputs and yields, particularly in Asia. However, we find little support for a relationship between connectivity and either % attainable yields met or field size. Our findings stress the need to better understand the mechanisms that link urbanization processes and agricultural management at different spatiotemporal scales.

1. Introduction

Urbanization, defined as changes in demographic composition and expansion of built-up areas, is one of the most important drivers of land-use change globally (Schneider et al., 2015). Urban expansion has contributed to the loss of substantial amounts of productive farmland in many regions (Bren d'Amour, 2016). However, urbanization is also linked to a myriad of physical, economic, or social processes occurring in rural areas (Jedwab et al., 2014). For example, a relationship between urbanization and diet trends toward more energy-dense processed foods and animal products could create feedbacks to rural agricultural production systems (Satterthwaite et al., 2010; Seto and Ramankuty, 2016). With increasing urbanization and food demand globally—especially in cities of the Global South—understanding the magnitude and variability of contemporary rural-urban interactions could provide insight on future land-use dynamics.

A growing body of research considers how urbanization impacts the agricultural systems that may supply food and other resources for cities. Rural and urban places coexist in a continuum, linked by multiple types of spatial and sectoral connections (Tacoli, 2004; Seto et al., 2012). Flows of people, goods, and information link urban populations to rural landscapes (e.g., Seto et al., 2012; Güneralp et al., 2013; Djurfeldt, 2015), which we refer to as rural-urban connectivity. Markets link rural activities to cities and larger regions via trade, enabling rural agricultural production in one location to benefit other distant places (Verburg et al., 2011). Market accessibility metrics can therefore be

---

[*] Corresponding author at: Department of Geography, McGill University, 805 Sherbrooke Street West, Montreal, Quebec, Canada H3A 0B9.
E-mail address: graham.macdonald@mcgill.ca (G.K. MacDonald).

https://doi.org/10.1016/j.gloenvcha.2019.101982
Received 27 November 2018; Received in revised form 4 September 2019; Accepted 9 September 2019

useful for examining spatial relationships between rural agricultural production systems and cities, as markets can affect decisions about agricultural investment and production (Sloat et al., 2018). Farmers close to urban markets can more easily purchase agricultural inputs, access services such as credit and insurance, and trade their products, which can lead to both increased agricultural productivity and agricultural specialization (Masters et al., 2013).

Outmigration from rural to urban areas can also link urbanization to agrarian change through the various connections between individuals' places of origin and destination (Tacoli et al., 2015). Urban centers often offer more diverse employment, educational, and cultural opportunities compared to rural areas, effectively extracting labor from agricultural sectors. Resulting changes in land-to-labor ratios can affect farm characteristics, including farm size, use of inputs, and commercialization of farms in general (ISPC, 2013; ISU, 2015). The remittances received by rural households from outmigration of family members may lead to greater investments in agricultural land in terms of labor and other inputs (Tacoli, 2004; Ospina et al., 2018), though some studies have found mixed effects of outmigration on agricultural outcomes (Gray and Bilsborrow, 2014; Ochieng et al., 2016). Conversely, remittances could lead to livelihood diversification, allowing households to move away from agricultural livelihoods altogether (Ellis, 1998).

Several local- and national-scale studies examine the mechanisms linking urbanization to agricultural land practices (e.g., Masters et al., 2013; Onwuchekwa and Ukandu, 2015; Jiang et al., 2013; Ospina et al., 2018). For example, Jiang et al. (2013) use a panel econometric model to assess linkages between urban expansion onto agricultural lands and changes in agricultural land use intensity in China. However, systematic analysis of rural-urban linkages is rarer in regional- and global-scale research (Seto et al., 2012). Global land systems classifications provide a conceptual starting point by highlighting underlying spatial patterns, such as regions where high rates of urbanization coincide with different forms of agriculture (van Asselen and Verburg, 2012; Václavík et al., 2013). Land system classifications increasingly incorporate different socioeconomic and agro-environmental factors, such as per capita incomes and land management characteristics (Ellis and Ramankutty, 2008; Letourneau et al., 2012). However, to our knowledge, connectivity variables, such as rural-urban migration and market accessibility, have not yet been integrated in land systems classifications at continental or global scales. Integrating rural-urban connectivity into land systems classification could help to better understand the drivers and feedbacks between urbanization and rural agricultural land use.

As a starting point for empirical analysis of potential relationships between urbanization and agricultural land use and management, we devised an analytical framework that combines the concept of rural-urban connectivity with key themes from the literature on global land systems (Fig. 1). Our framework depicts potential social and economic interactions between urban and rural areas embedded within a larger relationship between rural land system change and agricultural land management. We apply this framework to examine the extent to which urbanization processes may be associated with different agricultural inputs and outcomes across the Global South. Specifically, we focus on the role of rural-urban connectivity (i.e., rural market accessibility and rural outmigration) as a potential mechanism that could influence rural agriculture, distinguishing between the more proximate effects of urban land extent and the more distal effects of market access and migration. Agricultural land use intensity can be measured in multiple ways, including agricultural outputs per unit land per unit time (e.g., crop yield), cropping frequency, or inputs per unit area (e.g., fertilizer use) (Turner and Doolittle, 1978; Jiang et al., 2013; Erb et al., 2013). Accordingly, we adopt several indicators of agricultural inputs and outcomes, which are relevant to characterizing agricultural management and land use intensity.

Our aim in this study is to explore whether rural-urban connectivity, in terms of market access and human migration, is related to agricultural inputs (e.g., fertilizer application rates) and outcomes (e.g., crop yields) across the Global South. We use a two-step approach to test this. First, we conduct an exploratory analysis with multivariate clustering to examine how rural-urban connectivity relates spatially to several different agricultural and urbanization variables. Multivariate clustering facilitates exploration of complex relationships among multiple variables at the same time. By examining the spatial patterns in the resulting clusters and how these compare to key agricultural outcome variables (crop yields, % yield attainment, and cropping frequency), we draw three specific, directional hypotheses about the relationship between rural-urban connectivity and agricultural land use intensity. In a second step, we use an inferential statistical approach with propensity score matching techniques to test these hypotheses. Matching facilitates comparison among locations that differ in the connectivity measures by controlling for potential confounding factors, such as GDP per capita (*sensu* Jiang et al., 2013). In the matching analysis, rural-urban connectivity is compared across binary control/treatment groups (e.g., high vs. low market access) and each agricultural outcome/response variable (e.g., crop yields, % yield attainment, and cropping frequency) is considered in a separate model. Our findings from both approaches provide some evidence that urbanization processes may be associated with certain types of changes in nearby agricultural landscapes, stressing the need to consider what implications projected future urban growth could have for agricultural land use and management.

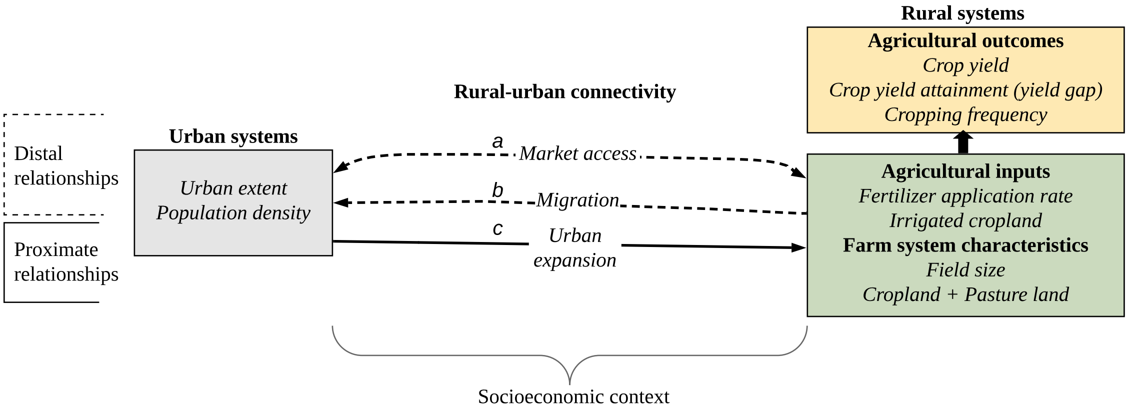

Fig. 1. Guiding analytical framework for assessing if changes in rural–urban connectivity (due to market access and migration) and urbanization (i.e., the expansion of urban extent) relate to changes in rural agricultural land use and management. The dotted arrows represent distal relationships and the solid arrow represents a proximate relationship. The key variables used in the study are italicized. Letters *a–c* refer to the three hypotheses tested in this study (described in Section 3.3). In this study, we account for the potential moderating influence of socioeconomic context on these relationships by considering gross domestic product (GDP) per capita as well as carrying out separate regional analyses.

2. Methods

2.1. Analytical framework

In this study, we use two variables to represent urban systems (urban extent and population density) and compare them to a series of variables that represent agricultural land management in rural systems—differentiating between agricultural inputs (e.g., fertilizer application rates), outcomes (e.g., crop yields), and farming system characteristics (e.g., field size) (Fig. 1). Although we distinguish between 'urban' and 'rural', our intent is to examine the role of rural-urban connectivity as a mechanism by which urban processes could influence rural areas. While urban expansion onto agricultural lands is a proximate link between urbanization and rural agriculture, market accessibility is a mechanism for potential distal links; for example, farmers who are closer to markets may have greater access to agricultural machinery, inputs, information, and buyers. Another distal link is rural outmigration, which may lead areas experiencing a loss of agricultural labor or receiving higher levels of remittances to see changes in farmland prices resulting in consolidation of land onto larger farms.

2.2. Study area

Our study area encompasses Latin America, Sub-Saharan Africa, and South and East Asia (individual countries are listed in Table S1) using two hexagon grids equal to ~13,000 km^2 (sub-regional scale) and ~130 km^2 (local scale), respectively (detailed in Section 2.4). These three Global South regions were chosen because of distinct and more recent processes of urbanization compared to countries in the Global North (Nagendra et al., 2018; Jedwab et al., 2014). In addition to the full study area, separate regional analyses helped to account for potential differences in the relationship between urbanization and rural agriculture across continents.

We divide agricultural lands into three overarching types of land use systems; *cropland, grazing*, and *mixed cropland-grazing* land systems using globally consistent datasets (Ramankutty et al., 2008). Agricultural lands were grouped into these categories based on either cropland or pasture being dominant (Fig. 2). We excluded *grazing*-dominated areas (where > 75% of the agricultural land is pasture) from further analysis given our emphasis on cropland systems and because less global data is available to assess land use intensity in pastures. Hexagons identified as *cropland* are concentrated mainly in South and Southeast Asia, as well as along the western coast of Africa (e.g., Nigeria and Cameroon) and East Asia; a smaller number are present in Mexico, Peru, Brazil, Suriname, and Argentina. The majority of agricultural land considered in the rest of Africa and Central and South America, as well as in some of East Asia, is classified as *mixed land* systems (Fig. S1).

2.3. Data sources and preparation

We collated existing global spatial datasets relating to key aspects of land use/land cover, demographic context, rural-urban connectivity, agricultural inputs, and farming system characteristics (Table 1). Dataset selection was based on coverage for the full study area in the period 2000 to 2005.

We used a global map of urban built-up areas (locations dominated by the built environment) from Schneider et al. (2009) to reflect circa 2001–2002 urban extent based on Moderate Resolution Imaging Spectroradiometer (MODIS) imagery. The MODIS dataset was used to distinguish between urban and non-urban areas throughout our analysis. As a second urban variable, we used circa-2000 population density maps from the Gridded Population of the World, Version 4 (GPWv4) (CIESIN, 2016).

We focus on two variables describing potential mechanisms for rural-urban connectivity: market access and population change. Verburg et al. (2011) computed a global market access index (ranging from 0–1) by using travel times to domestic and international markets and to cities with more than 50,000 inhabitants. Accounting for rural outmigration is particularly challenging because global spatial data on rural-urban internal migration are currently unavailable. To provide a globally consistent proxy for potential rural outmigration intensity across the Global South, we used the spatial dataset of Global Estimated Net Migration by Decade from the Center for International Earth Science Information Network (CIESIN) (de Sherbinin et al., 2012, 2015). This indirect estimation of net migration represents the number of people migrating into an area minus the number of people migrating out, after accounting for rates of natural increase from births and deaths. We use these data as a proxy to distinguish between rural areas with net inmigration and net outmigration.[1] To better reflect rural areas, we masked the MODIS urban extents from both the market access and the population change data in their respective base resolutions prior to summarizing in the hexagons (i.e., grid cells overlapping with built-up areas were removed).

We define agricultural management in *cropland* and *mixed* land systems based on key inputs (e.g., fertilizer and irrigation), farming system characteristics (e.g., field size), and farm outputs (e.g., crop yield, yield attainment, and cropping frequency). As a proxy for fertilizer nutrient use, weighted average synthetic nitrogen (N) fertilizer application rates for 140 crops (listed in Table S2) were calculated using harvested areas of each crop as weights (Mueller et al., 2012). As a proxy for irrigation water use, we use a global gridded dataset of monthly irrigated areas for 26 crop classes (listed in Table S2) to calculate the percentage of irrigated harvested areas out of the total harvested area (Portman et al., 2010). As a proxy for farm size, we used the global cropland field size index from Fritz et al. (2015), which ranges from 10 (very small field size of < 0.5 ha) to 40 (very large field size of > 100 ha). Total agricultural extent (cropland + pasture) was also summarized for each hexagon.

Agricultural outcome variables considered include yield gap, major cereal yields, and cropping frequency (crops are listed in Table S2). We used yield gaps from Mueller et al. (2012) that are based on observed crop yields, crop yield potentials, and climate data for 16 major crops. We calculated a weighted average yield gap by weighting the crop-specific yield gaps by the harvested areas of the 16 crops and then standardized this as a percentage of attainable yields met (*% attainable yields met = 100 − yield gap %*) for consistency with other variables. Mean cereal yields for 7 major cereals were summarized by hexagon (Mueller et al., 2012). Lastly, cropping frequency was calculated from the ratio of maximum monthly growing area to harvested area for 26 crops by using the approach of Siebert et al. (2010) to calculate the metric they call "cropping intensity", excluding fallow areas.

2.4. Hexagon grid

Spatial datasets were retained in their original resolutions and then aggregated to a coarser scale to help reduce uncertainties arising from different resolutions and input data quality (especially for the rural outmigration proxy). A tessellation of 5,172 hexagons, each with an area approximately equal to 13,000 km^2 at the equator was created in ArcGIS 10.5.1 (ESRI, 2017). Hexagons were chosen because this is the closest shape to a circle that can be tessellated, reducing distortion compared to a square grid in terms of grid cell area and thus also reducing inconsistency in terms of how much agricultural land was being compared in each hexagon. The hexagon area of 13,000 km^2 was chosen primarily due to the coarse resolution of the historical population grids used as an input to estimate subnational net-migration (de Sherbinin et al., 2015). However, to account for potential scale-dependence and to take advantage of the relatively higher spatial

---

[1] We caution that some areas identified as having rural population growth may actually reflect finer-scale urban population change due to the coarse resolution of the historical population grids (~10,000 km^2) used in calculating population change by de Sherbinin et al. (2015).

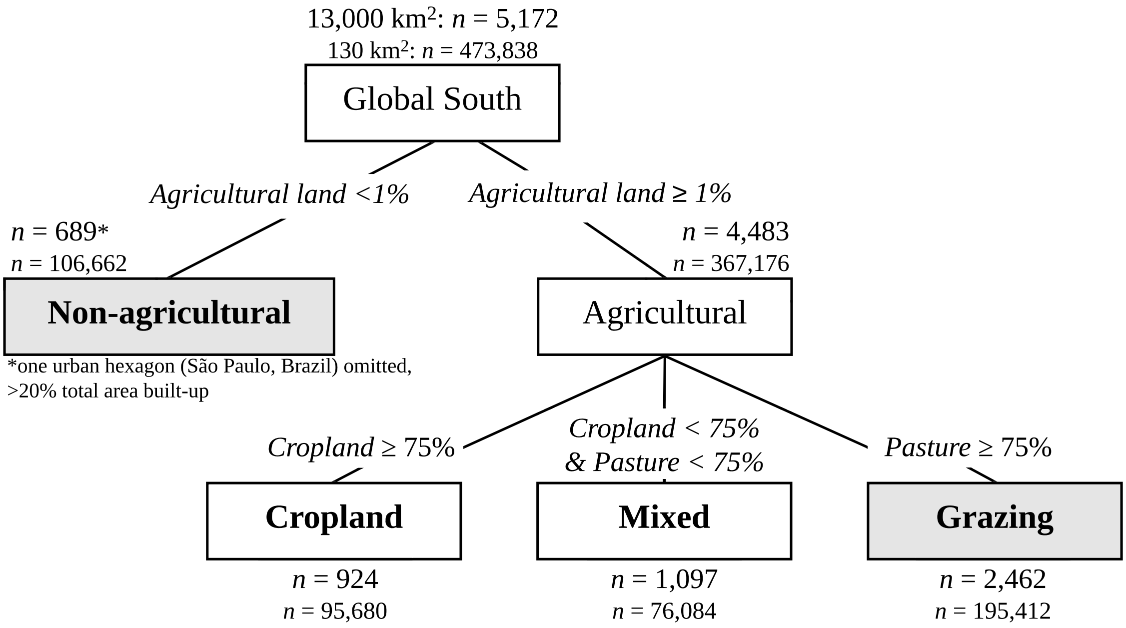

Fig. 2. Classification of land systems across the Global South. Sample size numbers indicate the count of hexagons for the 13,000 km^2 and 130 km^2 resolution analyses, respectively. Hexagons in the shaded groups were excluded from the analysis. See Fig. S1 for a map of the classification.

resolutions of most of the datasets (Table 1), we repeated our hypothesis testing analysis at a finer-resolution of approximately 130 km^2 (explained in Section 2.6). For this, we created an additional tessellation of 473,838 hexagons with an area of ~130 km^2. To avoid bias from coastal areas, hexagons with centroids that fell outside a country boundary from the Database of Global Administrative Boundaries were omitted (Hijmans, 2009), and the 130 km^2 hexagons were further clipped to the spatial extent of the cropland dataset (Ramankutty et al., 2008). Spatial variables were then summarized in each hexagon using the aggregation procedure shown in Table 1.

2.5. Step 1: multivariate approach to characterize 'urban' and 'rural' systems

As an exploratory analysis used for hypothesis development, we cluster the Global South into two typologies that address rural-urban connectivity from different perspectives that we call *urban-connectivity* and *rural-connectivity* typologies. These typologies give alternate vantage points on connectivity—one urban-focused and one rural-focused. Eight land cover, demographic, agricultural, and rural-urban connectivity variables are included (Table 1). Clustering is a multidimensional statistical approach that partitions objects with the goal of identifying the optimal number of natural groupings (Legendre and Legendre, 2012). We use *k*-means, an unsupervised clustering method that iteratively assigns each hexagon to the closest cluster centroid (Tan et al., 2006), given its efficiency with partitioning large datasets (Jain, 2010). Details on the clustering methodology are given in Appendix A. We calculated mean values of all variables (and standard deviations), from which we described and named each cluster (Table S3 and Table S4). All analyses were performed using the R Statistical Programming Language version 3.5.3 (R Core Team, 2018).

We summarize three key agricultural outcome variables (mean cereal yields, % yield attainment, and cropping frequency) within these typologies to assess the degree to which a gradient of rural-urban connectivity across the clusters coincides with different agricultural outcomes. We then test for significant differences in the outcome variables across the clusters by using Wilcoxon rank-sum tests (also known as Mann-Whitney tests). Wilcoxon tests can be used to compare differences between independent groups with a variable that is not normally distributed, which is the case for several of our agricultural variables. Controlling for inflation of type I error was done with Bonferroni correction, equal to $\alpha = 0.05$ / *number of comparisons*. The spatial patterns in the multivariate cluster analysis combined with the inferences from the Wilcoxon rank-sum tests led us to generate a set of hypotheses for formal testing (in Step 2).

2.6. Step 2: testing hypotheses about connectivity variables with matching

Building from the multivariate analysis, we further analyze the links between rural connectivity and agricultural land management by preprocessing the hexagons to reduce selection bias. In the absence of randomization, we adjust for potential selection bias through propensity score matching. Matching techniques facilitate an "apples-to-apples" comparison of outcomes in treatment versus control groups (Stuart, 2010) by placing more weight on comparisons between locations that have more similar socioeconomic and land use characteristics. Matching helps justify the assumption that both treatment and control groups come from the same population, allowing inference from comparing outcomes in the 'treatment' group to a 'control' that acts as a baseline for counterfactual comparison. We assess the relationship between the three treatments and the six diverse agricultural outcome/response variables: nitrogen application rates, irrigated cropland, % attainable yields, cereal yields, field size, and cropping frequency (Table 2). The matching models were developed for the entire study area as well as for each region independently (Africa, Asia, and Latin America) in order to explore potential differences within and across regions for the hypotheses in Fig. 1.

Three covariates were used to develop the propensity score upon which hexagons were matched, using a caliper width of 0.1 for all models (detailed in Appendix A): gridded GDP per capita data at 5 arc-minute resolution for the year 2000 from Kummu et al. (2018), as well as population density (CIESIN et al., 2016) and total agricultural land (Ramankutty et al., 2008). GDP per capita was included as a covariate that can be associated with both of the treatment variables and with the agricultural outcome/response variables. For example, a fairly strong correlation exists between per capita GDP and urbanization (Cohen and

Table 1
Datasets used for the multivariate and matching analyses.

Variable	Dataset	Base unit	Base spatial resolution	Temporal coverage	Aggregation method used in this study	Source
Rural connectivity						
Market access	Market access index	Index (0–1)	5 arc-min	~2000	Median market access after urban extent is masked (5 arc-min)	Verburg et al., 2011
Rural outmigration proxy	Global estimated net-migration grid	Population number	30 arc-s	1990–2000	Sum of population change after urban extent is masked (30 arc-s)	de Sherbinin et al., 2015
Urban connectivity						
Urban extent	MODIS	Binary	15 arc-s	~2001–2002	% of total hexagon area built-up	Schneider et al., 2009
Population density	Gridded Population of the World, v4	Population number	30 arc-s	2000	Median in hexagon	CIESIN et al., 2016
Farming systems and agricultural inputs						
Total agricultural land	EarthStat	km^2	5 arc-min	~2000	Sum of cropland & pasture in hexagon	Ramankutty et al., 2008
Nitrogen (N) application rate	EarthStat	$kg\ N\ ha^{-1}\ yr^{-1}$	5 arc-min	~2000	Grid cell weighted mean summarized as simple mean in hexagon	Mueller et al., 2012
Irrigated cropland area	MIRCA2000	ha	5 arc-min	~2000	% of cropland irrigated in hexagon	Portmann et al., 2010
Field size	Field size index	Index (10–40)	30 arc-s	~2005	Majority field size class in hexagon	Fritz et al., 2015
Agricultural outcomes						
Mean cereal yields	EarthStat	$t\ ha^{-1}\ yr^{-1}$	5 arc-min	~2000	Grid-cell level mean for 7 crops summarized as simple mean in hexagon	Monfreda et al., 2008
Yield gap	EarthStat	$t\ ha^{-1}$	5 arc-min	~2000	Grid cell weighted mean summarized as simple mean in hexagon	Mueller et al., 2012
Cropping frequency	MIRCA2000	# of crops harvested yr^{-1}	5 arc-min	~2000	Mean in hexagon	Portmann et al., 2010

Simet, 2018). Matching therefore helps to control for differences in socioeconomic, population, and land use characteristics (the covariates) across hexagons. For the 'treatment', we test three different types of models that estimate the association between our connectivity and urbanization variables (hypothesized links *a, b*, and *c* in Fig. 1) on agricultural outcomes among matched locations:

- Market access (*treatment* = high market access; *control* = low market access),
- Rural outmigration proxy (*treatment* = negative population change; *control* = positive population change), and
- Urban extent (*treatment* = high urban extent; *control* = low urban extent).

While these groupings are not 'treatments' in the classical sense, matching facilitates 'like' comparisons across our connectivity variables to evaluate their hypothesized associations with the outcome/response variables (Table 2). In the first model type, high market access (the 'treatment' group) hexagons are paired with more remote hexagons (the 'control' group) to estimate the impact of market access on agricultural activities by controlling for similar GDP per capita, population, and land use characteristics. While market access is continuous (Verburg et al., 2011), we transformed it into a binary variable in which areas were deemed "high access" at values ≥ 0.29 (the 'treatment' group, > 75th percentile of market access across the 13,000 km^2 hexagons) or "low access" with values < 0.29 (the 'control' group). A second model type is based on a binary variable that pairs hexagons with negative rural population change (the 'treatment' group) to comparable areas with positive population change (the 'control' group) to estimate the association between negative population change (a proxy for rural outmigration) and agricultural activities. Finally, a third model type pairs hexagons with high urban extent ($\geq 0.82\%$ built up area, the 'treatment' group, representing the 75th percentile of built-up area across the 13,000 km^2 hexagons) with comparable hexagons with low urban extent ($< 0.82\%$ built-up area, the 'control' group) to estimate the relationship between built-up environment and agricultural activities.

Overall, we test 72 models based on 6 different agricultural variables over the 3 'treatment' model types for each of 4 scales (the Global South, Africa, Americas, and Asia). To assess whether relationships are robust at a finer spatial resolution, we repeat the 72 models using data at the 130 km^2 resolution, with the same treatment/control group thresholds as in the 13,000 km^2 analysis. The matched hexagons for the study area were more highly represented in Asia (from 49% up to 58% depending on the hexagon resolution and matching model sample), so we emphasize this region in Section 3 but present separate results for each regional subset. For each agricultural variable, significant differences between the two groups were tested using the non-parametric Kruskal-Wallis (H) test at the $P = 0.05$ level, which is equivalent to a one-way analysis of variance (ANOVA) but can be used for variables with non-normal distributions.

3. Results

We found that the spatial patterns in our multivariate urban- and rural-connectivity clustering typologies are generally structured by the rural-urban connectivity variables (market access and population change). This is based on the general spatial overlap in the clusters for both of the typologies, where the connectivity variables are the two common variables (compare the urban-focused map in Fig. 3A with the rural-focused map in Fig. 4A). When summarizing the agricultural outcome variables according to the clusters in each of the two typologies (Fig. 3B–D and Fig. 4B–D), a general association emerges between clusters with greater connectivity and the three agricultural outcome variables (cereal yields, % attainable yields, and cropping frequency). These patterns are broadly similar for both the urban and rural typologies.

The propensity score matching models developed based on the

Table 2
Datasets used in the propensity score matching analyses.

Covariates for matching – used across all models (3)		Treatment variables (3)	Agricultural outcome/response variables (6)
		18 models estimated at 4 scales (72 models total)	
(1) GDP per capita	→	(1) Market access	(1) Nitrogen application rate
(2) Population density		(2) Rural outmigration proxy	(2) Percent irrigated cropland
(3) Total agricultural land		(3) Urban extent	(3) Field size
			(4) Major cereal yields
			(5) % of attainable yields met
			(6) Cropping frequency

spatial patterns evident in the multivariate clustering indicate that matched locations with greater rural-urban connectivity and urbanization have significantly higher values for roughly four of the six agricultural response variables at the Global South scale (Table 3). However, the strongest positive relationships between high-connectivity and high agricultural intensity locations are typically in Asia (Figs. 5 and 6), contributing disproportionately to the observed Global South scale effects.

3.1. Patterns of rural-urban connectivity across the typologies

The *k*-means clustering indicated five clusters for each of the urban- and rural-connectivity variables (Figs. 3A and 4A); in both, mean market access follows a general gradient from high to low across the clusters, with the urban typologies characterized by higher absolute market access (Tables S3 and S4). The relatively high-connectivity clusters in both the urban and rural typologies tend to overlap, although the urban typology clusters (Fig. 3) are more heterogeneous due to the effect of population density. The spatial patterns of rural-urban connectivity differ across the regions in both typologies, with the greatest connectivity in Asia. In the urban typologies, the greatest rural-urban connectivity is concentrated around large cities in eastern China and northern India, where urban extent and population density are relatively higher than other regions, as shown in the '*very high market access with rural inmigration cluster*' (Fig. 3A). In the rural-connectivity typologies, areas with the highest levels of agricultural inputs (i.e., nitrogen application rates and irrigated areas) are also found in eastern China and northern India, such as in the '*very high market access croplands*' cluster (Fig. 4A).

Clear patterns of low connectivity are also found in the clustering results for both typologies. The urban and rural clusters characterized by low market access emphasize more remote areas with lower population densities and lower cropland extents. For the urban typologies, the '*very low market access with rural outmigration*' cluster includes slightly built-up areas with low population densities dispersed across large parts of Southeast Asia, sub-Saharan Africa, and South America. Many of these areas are also encompassed by the rural typologies of '*very low market access mixed grazing-cropland with outmigration*' cluster, which includes more remote pasture and cropland characterized by moderate field sizes (influenced particularly by large fields in southern

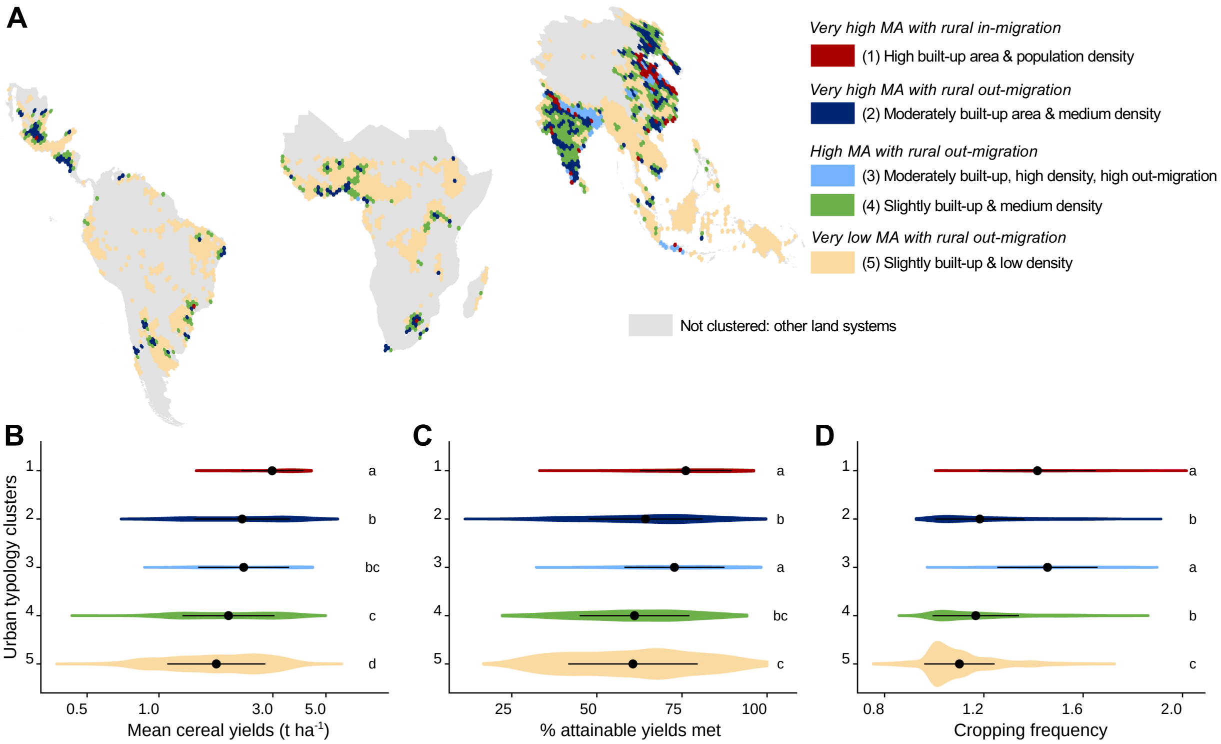

Fig. 3. Distribution of the multivariate urban-connectivity typologies across the Global South (A), with clusters grouped by comparable intervals of market access (MA) within each typology for the *cropland* and *mixed* land systems. Clusters are based on four variables: MA, rural population change (a proxy for rural outmigration), urban extent, and population density. Panels B, C and D show summaries of the agricultural outcome variables (mean cereal yields, % attainable yields met, and cropping frequency) according to the urban-connectivity typologies in A. Note the logarithmic X-axis scale in panel B. Means (black dots) and standard deviations (black lines) are shown; violins indicate data distributions. Letters next to each violin indicate results from the Wilcoxon rank-sum tests with Bonferroni-correction; within each panel, any rows not containing the same letter indicates a significant difference between clusters.

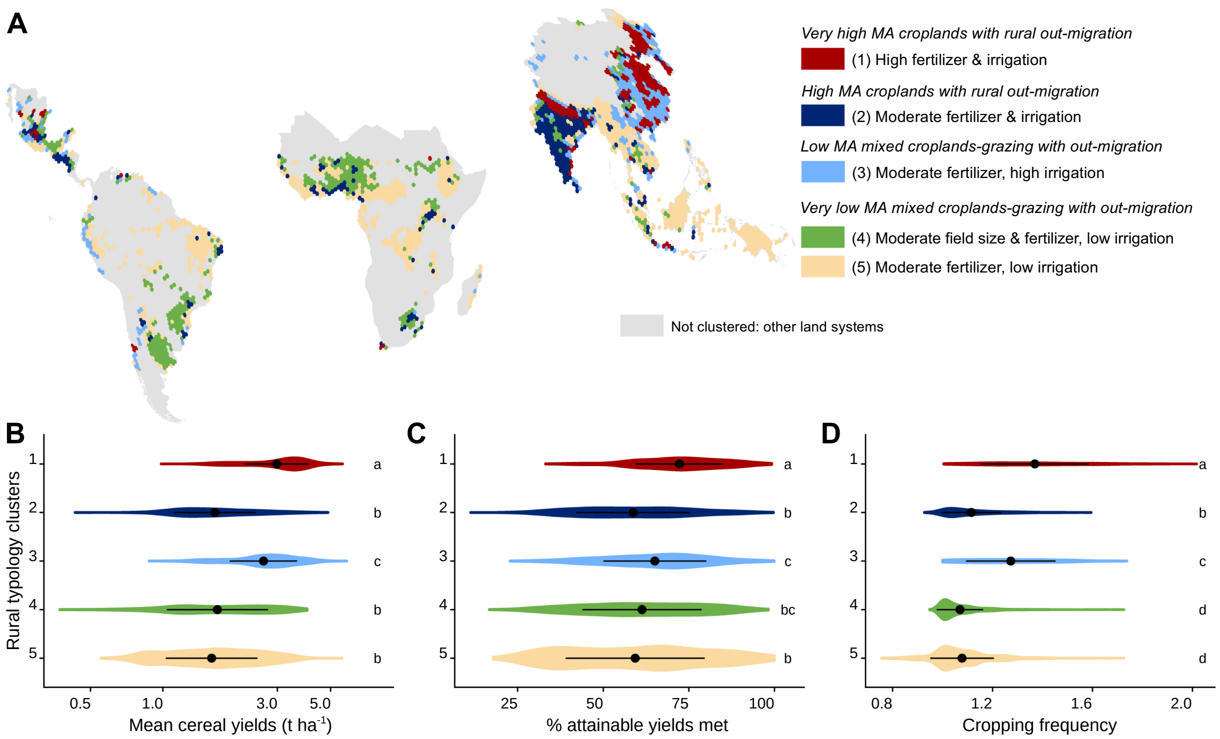

Fig. 4. Distribution of the multivariate rural-connectivity typologies across the Global South (A), with clusters grouped by comparable intervals of market access (MA) within each typology for the *cropland* and *mixed* land systems. Clusters are based on seven variables: MA, rural population change (a proxy for rural out-migration), nitrogen application rate, irrigated cropland, cropland area, pasture area, and field size. Panels B, C and D show summaries of the agricultural outcome variables (mean cereal yields, % attainable yields met, and cropping frequency) according to the rural-connectivity typologies in A. Note the logarithmic X-axis scale in panel B. Means (black dots) and standard deviations (black lines) are shown; violins indicate data distributions. Letters next to each violin indicate results from the Wilcoxon rank-sum tests with Bonferroni-correction; within each panel, any rows not containing the same letter indicates a significant difference between clusters.

South America), moderate nitrogen application rates, and low irrigation. While market access follows a clear gradient in both typologies, rural population change is more variable both within and across clusters in both typologies—however, higher market access clusters tend to have either inmigration (in the urban typologies; Table S3) or are characterised by higher absolute outmigration, on average (in the rural typologies; Table S4).

3.2. Agricultural outcome variables across the urban- and rural-connectivity clusters

The distributions of the agricultural outcomes are similar across the urban- and rural-connectivity typologies (e.g., Figs. 3B and 4B). The three agricultural outcomes (cereal yields, % attainable yields, and cropping frequency) are positively associated with the higher market access clusters. Mean cereal yields differ significantly across most urban clusters and some of the rural clusters, with significantly higher cereal yields in the cluster with the highest market access. For example, the urban and rural clusters #1 (high market access) have mean cereal yields (both 3.1 t ha^{-1} yr^{-1}) that are significantly greater than that of the urban and rural clusters #5 (1.9 and 1.8 t ha^{-1} yr^{-1}, respectively) with low rural market access and lower rural population change. There are also moderate trends for the more remote *cropland* and *mixed land* systems being characterized by lower % attainable yields (i.e., higher yield gaps) and cropping frequency (i.e., fewer cropping cycles per year) than higher market access clusters, which are relatively closer to their yield potentials.

3.3. Hypotheses testing results

Based on the general associations between higher connectivity clusters and the three outcome variables in our multivariate analysis, we devised three specific hypotheses (Table 3) to test the potential relationships between connectivity and urbanization variables for the six agricultural response variables in statistically matched hexagons. These hypothesized relationships reflect multiple dimensions of agricultural land management, including input, outcome, and farming system variables, which we generalize as 'land use intensity' (where higher values for any variable indicate greater intensity).

Hypothesis a. Locations with greater rural market access have higher agricultural land use intensity than low market access areas (link a in Fig. 1)

Matched high market access areas across the Global South have significantly higher mean nitrogen application rates (+27–40%), irrigated areas (+41–62%), cereal yields (+13–15%), and cropping frequency (+2–8%) compared to low market access areas (ranges reflect variation across the two analysis resolutions; Table 4). In the local (130 km^2 resolution) analysis, % yield attainments are also typically greater in high market access areas (Table 5). High market access areas have significantly higher irrigated areas than lower market access areas across all regions and at both resolutions, as well as higher cereal yields in all regions for the local resolution analysis (Table 5). However, there were few significant differences for matched high and low market access areas in Africa and no models were significant for field size.

Hypothesis b. Locations with negative rural population change (outmigration) have higher agricultural land use intensity than locations with positive rural population change (inmigration) (link b in Fig. 1)

Matched negative rural population change areas across the Global South have significantly higher nitrogen application rates (+13–31%), irrigated areas (+12–33%), and cereal yields (+11–12%) than areas with positive population change at both resolutions (Table S7). However, counter to our hypothesis, negative rural population areas have lower (−3 to −4%) yield attainment across the Global South at both

Table 3

Hypothesis development from multivariate clustering and overview of subsequent hypothesis testing results from the matching analysis.

Insight from multivariate clustering approach leading to hypothesis		Hypotheses	Support from matching – 13,000 km^2 hexagons	Support from matching – 130 km^2 hexagons	Key findings
Clear gradient in mean rural market access (high to low) across both urban and rural clusters (Tables S3 and S4). High rural market access clusters associated with relatively higher mean cereal yields, yield attainment, and cropping frequency.	→	*a*: Rural market access positively related to land use intensity	12 of 24 market access models +, significant ($P < 0.05$ or lower)	17 of 24 market access models +, significant ($P < 0.05$ or lower)	In Asia and at the Global South scale, matched locations with high market access have significantly higher: • mean N application rate • % irrigated cropland • cereal yields • cropping frequency at both 13,000 km^2 and 130 km^2 resolutions. Irrigated cropland has consistent positive relationship with market access in all regions.
Moderate association between relatively high population change clusters, higher mean cereal yields, yield attainment, and cropping frequency (Tables S3 and S4). Population change direction and magnitude varies across clusters.	→	*b*: Rural population change (proxy for outmigration) positively related to land use intensity	8 of 24 rural outmigration models +, significant ($P < 0.05$ or lower)	15 of 24 rural outmigration models +, significant ($P < 0.05$ or lower)	At the Global South scale, matched locations with rural outmigration have significantly higher: • mean N application rates • % irrigated cropland • cereal yields at both 13,000 km^2 and 130 km^2 resolutions. Results for individual regions are variable, but mean N application rates and cropping frequency are consistently greater in outmigration areas (130 km^2 resolution).
Built-up area has a clear influence on cluster spatial patterns in urban typologies, spatially coinciding with high-input rural clusters. Urban typology cluster #1 (highest built-up area) associated with clearly higher mean cereal yields, yield attainment, and cropping frequency relative to low built-up area clusters.	→	*c*: Urban extent positively related to land use intensity	14 of 24 urban extent models +, significant ($P < 0.05$ or lower)	17 of 24 urban extent models +, significant ($P < 0.05$ or lower)	In Asia and at the Global South scale, matched locations with high urban extent have significantly higher: • mean N application rate • % irrigated cropland • cereal yields • yield attainment • cropping frequency at both 13,000 km^2 and 130 km^2 resolutions. Irrigated cropland has consistent positive relationship with urban extent in all regions.

resolutions. At the regional scale, Asia has the most coincidence between negative rural population change and greater overall agricultural intensity, with significantly higher mean nitrogen application rates and cereal yields, as well as cropping frequency (at the 130 km^2 resolution), relative to areas with positive population change (Table 5). Results for Latin America and Africa are more variable at both resolutions.

Hypothesis c. Locations with high urban extent have higher agricultural land use intensity than low urban extent areas (link c in Fig. 1)

Matched high built-up areas in the Global South have significantly higher nitrogen application rates (+32–36%), irrigated areas (+48–60%), cereal yields (+18–19%), yield attainments (+10%), and cropping frequency (+5–8%) at both resolutions (Table S8). These results are consistent in Asia but more variable for the other regions in both of the analysis resolutions. For example, no significant differences were found between high built-up areas and cereal yields or cropping frequency in Africa at either scale. Significant differences in some regional models for nitrogen application rates and irrigated areas (at the 13,000 km^2 resolution) were sensitive to the specific subset of control/treatment hexagons chosen by the matching procedure (Table 5).

3.4. Evidence for and against hypotheses

Although we find mixed evidence in support of our three specific hypotheses across all 6 outcome/response variables (Table 3), findings in support of the hypotheses are generally robust at both sub-regional (13,000 km^2) and local (130 km^2) resolutions. Positive effects supporting hypotheses *a, b*, and *c* are clearest in Asia, where higher rural market access, negative rural population change, and higher built-up areas coincide with greater nitrogen application rates, irrigated areas, cereal yields, and cropping frequency in the local resolution models (Table 5 and Fig. 6). However, we find little or no support for a relationship between connectivity and the % yield attainment and field size variables at either resolution. There is also less support overall across the tested models for our rural outmigration proxy (negative population change) at the sub-regional scale.

At the Global South scale, when controlling for GDP per capita, population density, and total agricultural land, higher connectivity and more built-up areas have significantly higher means across at least four of the six agricultural response variables in both the local and sub-regional models. However, these trends are quite variable for the regional models. In some cases, results differ across the regions and the two resolutions (e.g., a positive effect observed at the local scale versus a negative effect at the sub-regional scale) due to different matching sample subsets. There are typically weaker trends in agricultural outcome variables in Latin America and especially Africa compared to Asia across both resolutions. For example, Africa has low overall fractions of irrigated cropland and matched areas with negative rural population change have significantly lower irrigated cropland (20–40% lower) than positive population change areas. Therefore, support for our hypotheses is not generalizable across regions.

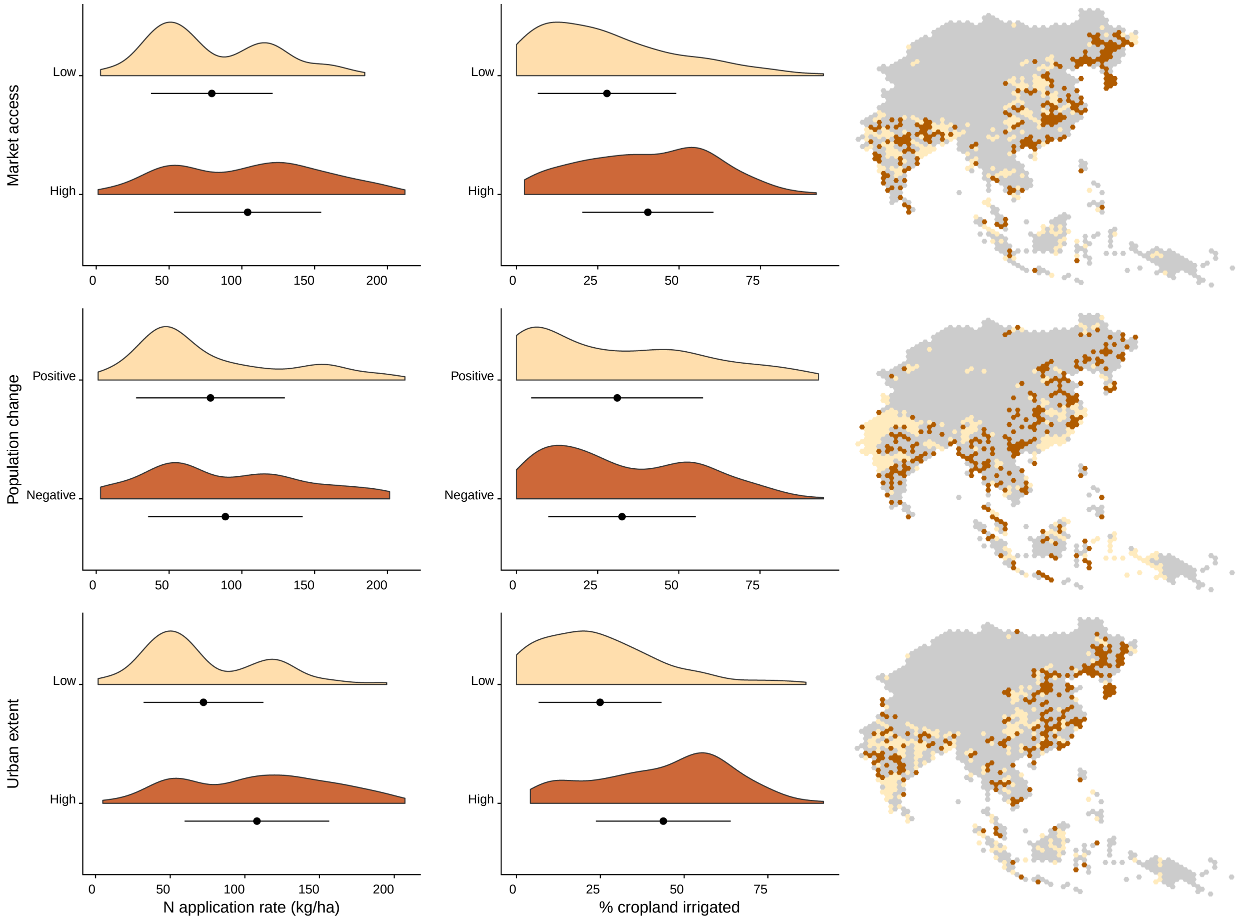

Fig. 5. Examples of matching comparisons for key agricultural input variables among the control and treatment subsets for Asia at the 13,000 km^2 hexagon resolution: average nitrogen application rates (first column) and % irrigated croplands (second column) for matched areas in Asia (matched subsets for each explanatory variable are shown in the third column). The graphs compare matched hexagons with high/low market access (top row), negative/positive population change (middle row, a proxy for rural outmigration and inmigration), and high/low urban extent analyses (bottom row). The top part of each plot shows the distribution shape (as a partial violin plot) and the bottom part shows the means and standard deviations as thick black dots and lines. Note that five of the six models are significant at the $P < 0.05$ level or lower (control/treatment means are not significantly different for % cropland irrigated in the population change model).

4. Discussion

Our results empirically show some evidence that greater rural-urban connectivity and urban extents are associated with differences in agricultural land management, particularly in Asia. We found some support for our hypothesis that areas closer to markets that experience rural outmigration tend to have greater agricultural land use intensity compared to relatively remote areas or those that experience inmigration. This is especially the case for irrigation, where locations with high connectivity and urban extents have significantly more irrigated cropland compared to areas with lower connectivity and urban extents across the Global South for both the local and sub-regional scales (significant positive relationships were found for 21 of the 24 tested models for irrigated areas overall in Table 5). This finding is in line with the analysis by Thebo et al. (2014), which showed that 60% of the global irrigated cropland area is located within 20 km of urban areas. Although one of the key strengths of statistical matching is that pre-processing of the data can better facilitate causal inference (Stuart, 2010), we caution against interpreting a causal link between rural-urban connectivity and agricultural change at this scale, particularly given that consistent panel data was not available.

In contrast, we found little to no support for a consistent relationship between urbanization processes and field size at either the sub-regional or local scales. Our rationale for including field size (as a proxy for farm size) was based on historic trends toward consolidation of crop production onto relatively larger farms in parts of countries such as the United States (MacDonald et al., 2013) and Brazil (Ferreira Filho and Vian, 2016) as production became more specialized over time. However, croplands on peripheries of urban areas may experience different farm size dynamics due to land constraints and higher population densities (Samberg et al., 2016). For example, urban expansion in the Hang-Jia-Hu region of China between 1994–2003 led to fragmentation and isolation of agricultural patches (Su et al., 2011). Rather than no relationship, there may be a directional relationship between urbanization and field size that works differently within and across regions (Djurfeldt, 2015; Jayne et al., 2016). Farm size may also interact with other agricultural inputs and outcome variables. In rural areas experiencing outmigration, reduced household labor availability could lead to reduced agricultural investments or land abandonment (Gray, 2009; Castelhano et al., 2016). In contrast, Wu et al. (2018) found that a 1% increase in farm size in China was associated with a 0.3% and 0.5% decrease in fertilizer and pesticide use per hectare, respectively.

Regional disparities in the effects observed here (Table 5) indicate that any mechanistic relationship between rural-urban connectivity and rural agricultural change likely varies within and across countries. For example, analysis of links between urbanization and agricultural land use intensity in China showed a negative relationship between urban expansion and cropping frequency (nationally; Jiang et al., 2013) but a positive relationship for fertilizer use (in Henan province; Jiang and Li, 2016). Declining cropland area per capita and resulting land scarcity

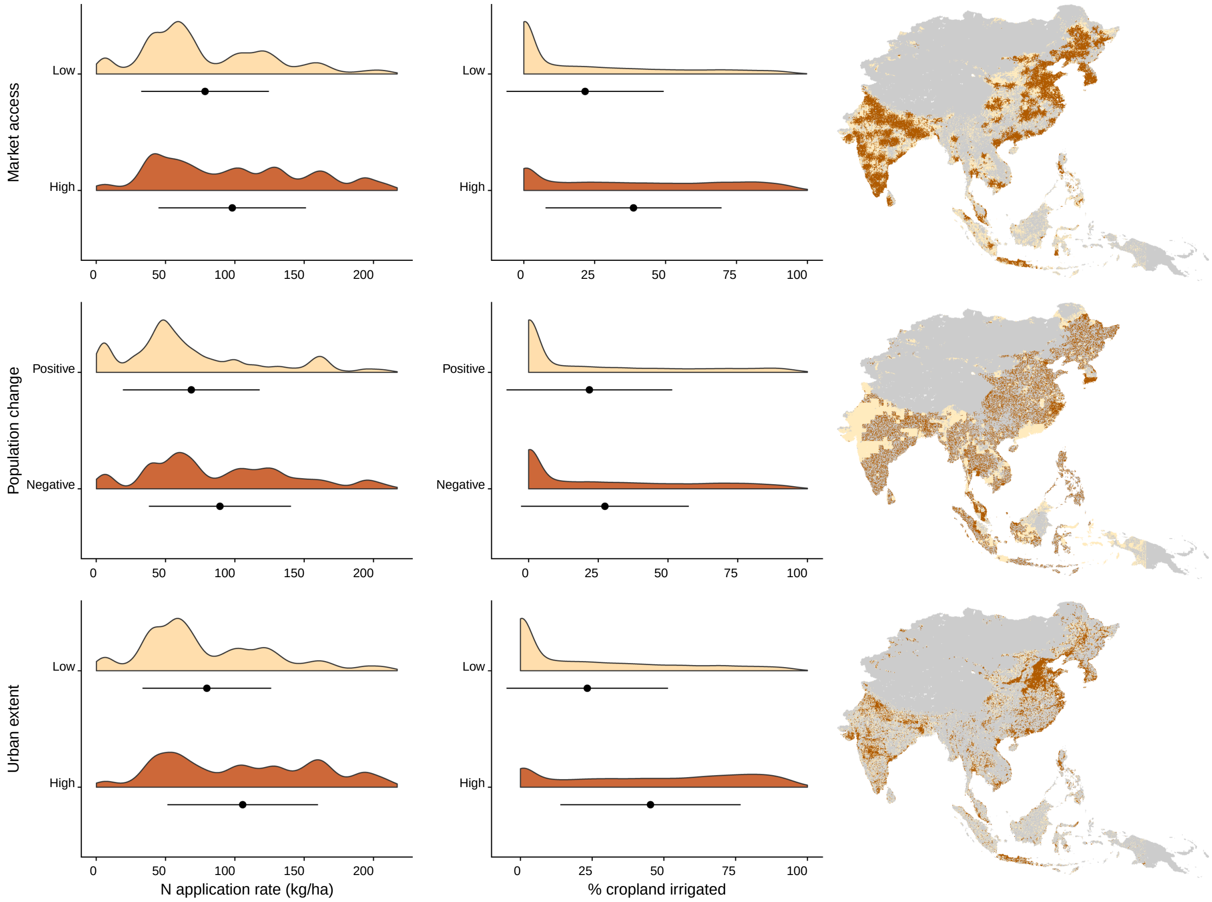

Fig. 6. Examples of matching comparisons for key agricultural input variables among the control and treatment subsets for Asia at the 130 km^2 hexagon resolution: average nitrogen application rates (first column) and % irrigated croplands (second column) for matched areas in Asia (matched subsets for each explanatory variable are shown in third column). The graphs compare matched hexagons with high/low market access (top row), negative/positive population change (middle row, a proxy for rural outmigration and inmigration), and high/low urban extent analyses (bottom row). The top part of each plot shows the distribution shape (as a partial violin plot) and the bottom part shows the means and standard deviations as thick black dots and lines. Note that all models are significant at the $P < 0.05$ level or lower.

as well as increasing urban wages were potential explanatory factors (Jiang and Li, 2016). Similarly, a study exploring the impacts of urbanization on agricultural land use in Pennsylvania found that the pressures put on farmers included moving to more land-intensive systems, diversifying their income from off-farm work, selling their land and reducing agricultural activities (Larson et al., 2001). Accounting for country-specific responses to urbanization and sub-regional land-use policies is therefore an important next step for empirical analysis of the relationship between urbanization characteristics and agricultural land management characteristics at different scales.

Our findings are dominated by trends that are clearer in Asia, a region which represents from 49% to 58% of the hexagons in our matching analysis (the range varies based on the matching models and the two resolutions). Differences in urbanization histories and agricultural development among Asia, Africa, and Latin America may therefore help to explain regional differences in our results. For example, while Asia has a much

Table 4

Examples of results for selected market access (MA) models at the Global South scale showing the percent differences in means of four response/outcome variables with significant, positive relationships (at least $P < 0.05$ or lower) at both the local and sub-regional hexagon resolutions. For the Global South scale, 58% of hexagons in the 13,000 km^2 and 57% of hexagons in the 130 km^2 market access matching models are located in Asia, respectively. Tables S7 and S8 provide similar results for the rural outmigration proxy and urban extent models, respectively.

Response/outcome variable	Hexagon resolution	Means from matching subsets [Treatment / Control]	% difference in mean for high MA areas relative to low MA areas	Standard deviation [Treatment / Control]	Sample size [Treatment + Control]
Nitrogen application rate *(kg N ha^{-1} yr^{-1})*	13,000 km^2	70 / 51	+27%	52 / 47	*n* = 692
	130 km^2	78 / 47	+40%	46 / 56	*n* = 82,064
Irrigated cropland *(%)*	13,000 km^2	31 / 19	+41%	24 / 21	*n* = 692
	130 km^2	31 / 12	+62%	23 / 32	*n* = 82,064
Mean cereal yields *(tonnes ha^{-1} yr^{-1})*	13,000 km^2	2.4 / 2.1	+13%	1.0 / 0.9	*n* = 692
	130 km^2	2.4 / 2.0	+15%	0.9 / 1.1	*n* = 82,064
Cropping frequency *(# of crops harvested yr^{-1})*	13,000 km^2	1.18 / 1.16	+2%	0.19 / 0.17	*n* = 692
	130 km^2	1.22 / 1.13	+8%	0.19 / 0.25	*n* = 82,064

Table 5

Results of the 72 matching models tested at the 13,000 km^2 (sub-regional) scale and the 72 matching models tested at the 130 km^2 (local) scale, respectively. Shaded cells indicate where there is a significant, positive relationship between the rural-urban connectivity treatment variables (high market access and rural outmigration), or the high urban extent treatment variable, and one of the six agricultural outcome/response variables. Each model was tested with two subsets based on a different random number seed, and some models were only significant in one of the two subsets (shown with a *). Cells that are not shaded had either a non-significant difference between the treatment and control means or a significant negative relationship. Numerical details for the matching models (chi-square and exact *P* values) are given in Tables S5 and S6.

Sub-regional scale (13,000 km^2)

	Global South	Africa	Latin America	Asia	Global South	Africa	Latin America	Asia
High market access	■	░ *		■	■	░ *	■	■
Rural outmigration	░			▓	■		■	
High urban extent	■	░ *		■	■	░ *	░ *	■
	Nitrogen application rate				Irrigated cropland			
High market access	■			■				■
Rural outmigration	■		▓	■				
High urban extent	■			■	■			■
	Mean cereal yields				% Attainable yields			
High market access	░			░				
Rural outmigration							░ *	
High urban extent	■		░ *	■				
	Cropping frequency				Field size			

Local scale (130 km^2)

	Global South	Africa	Latin America	Asia	Global South	Africa	Latin America	Asia
High market access	■		■	■	■	■	■	■
Rural outmigration	■	■	■	■	■		■	■
High urban extent	■			■	■	■	■	■
	Nitrogen application rate				Irrigated cropland			
High market access	■	■	■	■	■	■		■
Rural outmigration	■	■		■		■		
High urban extent	■		■	■	■	▓	■	■
	Mean cereal yields				% Attainable yields			
High market access	■		■	■				
Rural outmigration	■	■	■	■				
High urban extent	■		■	■			■	
	Cropping frequency				Field size			

■ P<0.001 ▓ P<0.01 ░ P<0.05 * Significant in only 1 subset

higher absolute urban population and a larger number of cities with populations > 1 million, the fraction of population living in urban areas was already much higher in Latin America (> 70%) than either Asia or Africa (both < 40%) by the 1990s (Montgomery, 2008). While all regions covered in our study had declining average farm size from 1960 to 1990, Latin America as a whole had a rise in farm size by 2000 (Lowder et al., 2016). Conversely, smallholder farming and declining farm sizes have persisted in Africa and Asia, at least partly related to increasing rural population densities, although some authors discuss possible recent shifts toward farm consolidation with rural demographic change in Asia (Masters et al., 2013; Jayne et al., 2016; Lowder et al., 2016).

Overall, the many non-significant relationships between rural-urban connectivity and different agricultural variables in our study may be insightful in that they show disparities in how rural agriculture can respond differently to urbanization trajectories across regions. However, although the cereal crops considered in our study represent 37–53% of the total per capita caloric food supplies for our study regions circa 2000 (see Table S9), the importance of these crops to diets is lower on average in Latin America and Sub-Saharan Africa, which may in part be reflected in their lower yields and lack of significant relationships to rural-urban connectivity in our study. Additionally, our method for aggregating cereal yields is based on tonnes per hectare, which may bias the yield values among regions growing different cereal crops (e.g., maize versus rice) since some cereals have very different yields on a mass-basis but similar yields on a caloric basis (Cassidy et al., 2013).

4.1. Priorities for follow-up research

Rural-urban linkages affecting agriculture that happen at fine peri-urban scales may be important and pronounced for some cities (e.g., Karg et al., 2016). In our study, we found that some associations between rural-urban connectivity and agricultural inputs and outcomes are robust at both local (130 km^2) and sub-regional (13,000 km^2) scales. More systematic analysis of these rural-urban dynamics and driving factors between urban expansion and agricultural land management over time could provide further important insights on agricultural sustainability. As many urban areas are located near some of the world's most productive croplands (Thebo et al., 2014), the direct competition between urban expansion and food production is growing, particularly in Africa and Asia (Bren d'Amour et al., 2016). Between

2000 and 2040, urban growth could displace around 65 million tonnes of crop production (van Vliet et al., 2017). In rapidly urbanizing regions, such as China, displacement effects impact national resource-use efficiency by shifting crop production to regions with relatively lower crop yields and lower crop-water productivity (Zuo et al., 2018).

Our study focuses on rural-urban connectivity as a distal variable, which has some parallels with the urban teleconnections framework for examining links between cities and their hinterlands (Seto et al., 2012). For example, Güneralp et al. (2013) give examples of 'short-distance' teleconnections between urban areas and surrounding peri-urban or regional systems, which they contrast with 'long-distance' teleconnections linking cities to inter-regional and international systems. While we find some support that market access and migration could be related to agricultural outcomes at the regional scale, our study does not capture the specific factors that might be involved in this (such as migrant remittances; Ospina et al., 2018). Our study also omits long-distance interactions across regions. Countries in the Global North have often shifted environmental pressures from their own lands to the Global South by importing food and other commodities (e.g., DeFries et al., 2010; Davis and Caldeira, 2010; Lambin and Meyfroidt, 2011). This displacement effect may also occur within Global South countries, such as in the case of soy exports from Latin America to China (Lathuillière et al., 2014). Macro-scale factors and policies that shape these relationships are further important considerations for understanding changes in rural agriculture across the Global South.

Our study does not incorporate land use policies that may explain some of the regional variation in our findings. Institutional, policy, and economic contexts often determine how efficient markets are and regulate how agricultural land and other natural resources are used (Lambin et al., 2001). Many countries have regulatory policies intended to protect agricultural lands around urban areas, through different levels of land use restrictions. As mentioned above, the Chinese government encouraged agricultural intensification through historic subsidies for fertilizers and other inputs (Jiang et al., 2013). An analysis of the role of socioeconomic and policy factors in agricultural land use intensity in China found that GDP per capita and agricultural investments were positively associated with cropping frequency, but that urban expansion was negatively associated with cropping frequency at the provincial level (Jiang et al., 2013). Price and trade policies may also be designed and implemented to protect rural producers. Investments in agricultural research and technology may also have significant impacts on driving changes in agricultural land use and management (Cassman et al., 2005). Thus, national policies and regulations can impact farmers' decisions related to changes in rural-urban connectivity.

4.2. *Data limitations*

The variable with greatest uncertainty in our study is rural population change, which we use as a proxy for rural outmigration since the underlying dataset makes no distinction between internal or international migration (de Sherbinin et al., 2015). Studying migration flows is inherently difficult due to the complexity of the process (Kahanec and Zimmermann, 2008). The majority of internal migration flows involve movement from rural to urban areas, yet movement of populations also takes place between rural places, or even from urban to rural areas (Tacoli, 2004). For example, migration between rural places has historically been a predominant type of migration in India (Bhattacharya, 2000). We chose the population change dataset from de Sherbinin et al. (2015) as a proxy for rural outmigration primarily because the gridded spatial dataset offered greater comparability with the other variables, enabled masking of urban extent, and was consistent across the Global South. Other datasets on internal migration flows exist but are further limited by coarse spatial resolutions, non-standardized approaches to data collection, or more limited spatial coverage. For example, the WorldPop database provides data on internal migration by estimating internal human migration flows between subnational administrative units for malaria endemic countries between 2005 and 2010 (Sorichetta et al., 2016). The International Public Use Microdata Series-International (IPUMS) database also includes data on internal migration based on harmonized census microdata, but this was limited to 19 countries as of 2018. Overall, the effect of rural outmigration on agricultural activities is often mixed and context-dependent (Gray and Bilsborrow, 2014; Ochieng et al., 2016).

We use global data from the 2000–2005 period, which capture the magnitude of urbanization that occurred in the 1990s but overlook the more rapid rates of urbanization that have occurred since 2005—particularly in Asia (Schneider et al., 2015). Although our study period is static, we capture a gradient in urban extents (i.e., low to high degrees of urban extent) by focusing on intra- and inter-regional variation in urban extent. This urban gradient mimics a 'space-for-time' substitution approach, which may provide some insights on urban expansion more generally. Other indicators of the full array of potential impacts that urbanization may have on agricultural land systems include effects on land tenure, farm mechanization, agricultural subsidies, and trade of agricultural products, but consistent spatial data across the Global South is currently very limited. For example, we used field size as a proxy for farm size, since the currently best available farm size data is reported at often relatively large administrative unit scales, such as provinces in China (Samberg et al., 2016). Still, our analysis is a useful step by integrating rural-urban connectivity as a distal mechanism of rural land system change. Finally, we created binary classifications for the market access, population change and urban extent variables to reproduce a "treatment" effect, which is required for hypothesis testing using propensity score matching. However, different thresholds in the classifications would likely yield distinct results. Such issues are representative of the modifiable areal unit problem (MAUP) for geographic data more generally (e.g., Jelinski and Wu, 1996).

5. Conclusions

Our study explores links between rural-urban connectivity and agricultural land management within and across the Global South by using two complementary approaches. Our findings demonstrate some evidence for links between connectivity and some agricultural inputs and outcome variables that are robust across regions and scales (~ 130 and 13,000 km^2 hexagons, respectively). At the Global South level, locations of greater overall connectivity (market access and population change) tend to be associated with higher mean nitrogen fertilizer application rates, irrigated areas, and cereal yields at both resolutions. Significant positive relationships are most common in Asia, where matched locations with high rural market access, negative rural population change, and high built-up areas have significantly greater fertilizer application rates and mean cereal yields at both resolutions. Findings from our study highlight the need for further empirical research on the linkages between rural and urban areas at different scales—including at longer distances internationally via trade (DeFries et al., 2010; Seto et al., 2012)—and the role of urbanization in the sustainable use of global land resources.

Our approach to hypothesis testing enables comparison of locations with similar socioeconomic and land use characteristics using propensity score matching, which is to our knowledge a novel application for research on the links between urbanization and land system change at this spatial scale. Matching techniques could be used for other large-scale land change research to examine the effects of urbanization on land use and land management.

Research on rural-urban connectivity and agriculture is key to understanding how future urbanization could impact food production systems both more locally and globally. Enhanced rural-urban connectivity could theoretically influence agricultural sustainability by increasing farm productivity in urbanizing regions, but this needs to be accompanied with appropriate policies to limit displacement of croplands lost due to urbanization and measures to reduce the negative externalities of intensification. Thus, identifying generic patterns of agricultural land use intensity and how it varies spatially relative to rural-urban connectivity could help to formulate policy responses to rapid urbanization. With forecasts of future urbanization, particularly in Sub-Saharan Africa (Seto et al., 2012b), changes in urban extent, market access, and migration could have unexpected effects on regional agriculture.

Declaration of Competing Interest

None.

Acknowledgments

Comments from Alain Boudet and Andréane Bellon de Chassy, and discussions with Camille Ouellet-Dallaire and Alessandro Sorichetta were important contributions to this research. We appreciate helpful comments from two anonymous reviewers on the manuscript. This work was supported by the McGill Sustainability Systems Initiative (MSSI) and the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grants Program (grant number RGPIN-2016-04920 to GKM). BER was supported by the Social Science and Humanities Research Council (SSHRC) of Canada (grant number 430-2014-00861). LS was supported by the Institute on the Environment at the University of Minnesota.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.gloenvcha.2019.101982.

References

van Asselen, S., Verburg, P.H., 2012. A Land System representation for global assessments and land-use modeling. Glob. Change Biol. 18 (10), 3125–3148. https://doi.org/10.1111/j.1365-2486.2012.02759.x.

Bhattacharya, P.C., 2000. An analysis of rural-to-rural migration in India. J. Int. Dev. 12 (5), 655–667.

Bren d'Amour, C., Reitsma, F., Baiocchi, G., Barthel, S., Güneralp, B., Erb, K.-H., Haberl, H., ... Seto, K.C., 2016. Future urban land expansion and implications for global croplands. Proc. Natl. Acad. Sci. U. S. A. 114 (34), 8939–8944. https://doi.org/10.1073/pnas.1606036114.

Cassidy, E.S., West, P.C., Gerber, J.S., Foley, J.A., 2013. Redefining agricultural yields: from tonnes to people nourished per hectare. Environ. Res. Lett. 8 (3), 034015.

Cassman, K.G., Wood, S., Choo, P.S., Cooper, H.D., Devendra, C., Dixon, J., et al., 2005. Cultivated systems. In: Hassan, R. (Ed.), Ecosystems and Human Well-Being: Current state and trends. Vol. 1. Conditions and trends. Millennium Ecosystem Assessment. Island Press, Washington, DC, pp. 745–794 Chapter 26.

Castelhano, M., Lawell, C.Y.C.L., Sumner, D.A., Taylor, J.E., 2016. The effects of migration and remittances on productive investment in rural Mexico. In: Working Paper. University of California, Davis Retrieved from. http://www.des.ucdavis.edu/faculty/Lin/Mexico_migration_remittances_investment_paper.pdf.

Center for International Earth Science Information Network (CIESIN), Columbia University, International Food Policy Research Institute (IFPRI), The World Bank, & Centro Internacional de Agricultura Tropical (CIAT), 2016. Global Rural-Urban Mapping Project, Version 1 (GRUMPv1): Population Density Grid. NASA Socioeconomic Data and Applications Center (SEDAC), Palisades, NY. https://doi.org/10.7927/H4VT1Q1H.

Cohen, M., Simet, L., et al., 2018. Macroeconomy and urban productivity. In: Elmqvist, T., Bai, X., Frantzeskaki, N., Griffith, C., Maddox, D., McPhearson, T. (Eds.), Urban Planet: Knowledge towards Sustainable Cities. Cambridge University Press, Cambridge, pp. 130–146.

Davis, S., Caldeira, K., 2010. Consumption-based accounting of CO_2 emissions. PNAS 107 (12), 5687–5692. https://doi.org/10.1073/pnas.0906974107.

DeFries, R.S., Rudel, T., Uriarte, M., Hansen, M., 2010. Deforestation driven by urban population growth and agricultural trade in the twenty-first century. Nat. Geosci. 3 (3), 178–181. https://doi.org/10.1038/ngeo756.

Djurfeldt, A.A., 2015. Urbanization and linkages to smallholder farming in sub-Saharan Africa: implications for food security. Glob. Food Sec. 4, 1–7. https://doi.org/10.1016/j.gfs.2014.08.002.

Ellis, F., 1998. Household strategies and rural livelihood diversification. J. Dev. Stud. 35 (1), 1–38.

Ellis, E.C., Ramankutty, N., 2008. Putting people in the map: anthropogenic biomes of the world. In: Frontiers in Ecology and the Environment. 6. pp. 439–447. https://doi.org/10.1890/070062.

Erb, K.H., Haberl, H., Jepsen, M.R., Kuemmerle, T., Lindner, M., Muller, D., Verburg, P.H., Reenberg, A., 2013. A conceptual framework for analysing and measuring land-use intensity. Curr. Opin. Environ. Sustain. 5 (5), 464–470. https://doi.org/10.1016/j.cosust.2013.07.010.

ESRI, 2017. ArcGIS Desktop. Release 10.5.1. Environmental Systems Research Institute, Redlands, CA.

Ferreira Filho, J.B.D.S., Vian, C.E.D.F., 2016. The evolving role of large and medium farms on Brazilian agriculture. Agric. Econ. 47 (S1), 215–225.

Fritz, S., See, L., McCallum, I., You, L., Bun, A., Moltchanova, E., Duerauer, M., ... Obersteiner, M., 2015. Mapping global cropland and field size. Glob. Chang. Biol. 21 (5), 1980–1992. https://doi.org/10.1111/gcb.12838.

Gray, C.L., 2009. Rural out-migration and smallholder agriculture in the southern Ecuadorian Andes. Popul. Environ. 30, 193–217. https://doi.org/10.1007/s11111-009-0081-5.

Gray, C.L., Bilsborrow, R.E., 2014. Consequences of out-migration for land use in rural Ecuador. Land use Policy 36, 182–191. https://doi.org/10.1016/j.landusepol.2013.07.006.

Güneralp, B., Seto, K.C., Ramachandran, M., 2013. Evidence of urban land teleconnections and impacts on hinterlands. Curr. Opin. Environ. Sustain. 5 (5), 445–451. https://doi.org/10.1016/j.cosust.2013.08.003.

Hijmans, R. (2009). Global administrative areas boundaries without limits, 2009. Retrieved from:http://www.gadm.org/.

ISCP (Independent Science and Partnership Council). (2013). Urbanization and farm size in developing countries: iImplications for agricultural research. Retrieved from: http://ispc.cgiar.org/sites/default/files/ISPC_StrategyTrends_FarmSize_Synthesis.pd.

ISU (International Sustainability Unit). (2015). Food in an urbanised world: The role of city region food systems in resilience and sustainable development. Retrieved from:http://www.fao.org/fileadmin/templates/FCIT/documents/Food_in_an_Urbanised_World_Report_DRAFT_February_2015.pdf.

Jain, A.K., 2010. Data clustering: 50 years beyond K-means. Pattern Recognit. Lett. 31 (8), 651–666. https://doi.org/10.1016/j.patrec.2009.09.011.

Jayne, T.S., Chamberlin, J., Traub, L., Sitko, N., Muyanga, M., Yeboah, F.K., Anseeuw, W., Chapoto, A., Wineman, A., Nkonde, C., Kachule, R., 2016. Africa's changing farm size distribution patterns: the rise of medium-scale farms. Agric. Econ. 47 (S1), 197–214.

Jedwab, R., Christiaensen, L., Gindelsky, M., 2014. Demography, urbanization and development: rural push, urban pull and… urban push. J. Urban Econ. 98, 6–16. https://doi.org/10.1016/j.jue.2015.09.002.

Jelinski, D.E., Wu, J., 1996. The modifiable areal unit problem and implications for landscape ecology. Landsc. Ecol. 11 (3), 129–140.

Jiang, L., Deng, X., Seto, K.C., 2013. The impact of urban expansion on agricultural land use intensity in China. Land Use Policy 35, 33–39. https://doi.org/10.1016/j.landusepol.2013.04.011.

Jiang, L., Li, Z., 2016. Urbanization and the change of fertilizer use intensity for agricultural production in Henan Province. Sustainability 8 (2), 186. https://doi.org/10.3390/su8020186.

Kahanec, M., Zimmermann, K.F., 2008. Migration and globalization: challenges and perspectives for the research infrastructure. In: Discussion paper No. 3890. The Institute for the Study of Labor (IZA).

Karg, H., Drechsel, P., Akoto-Danso, E.K., Glaser, R., Nyarko, G., Buerkert, A., 2016. Foodsheds and city region food systems in two West African cities. Sustainability 8 (12), 1175. https://doi.org/10.3390/su8121175.

Kummu, M., Taka, M., Guillaume, J.H.A., 2018. Gridded global datasets for gross domestic product and human development index over 1990–2015. Sci. Data 5, 180004. https://doi.org/10.5061/dryad.dk1j0.

Lambin, E.F., Turner, B.L., Geist, H.J., Agbola, S.B., Angelsen, A., Bruce, J.W., Coomes, O.T., ... Xu, J., 2001. The causes of land-use and land-cover change: moving beyond the myths. Glob. Environ. Change 11 (4), 261–269. https://doi.org/10.1016/S0959-3780(01)00007-3.

Lambin, E.F., Meyfroidt, P., 2011. Global land use change, economic globalization, and the looming land scarcity. PNAS 108 (9), 3465–3472. https://doi.org/10.1073/pnas.1100480108.

Lathuillière, M.J., Johnson, M.S., Galford, G.L., Couto, E.G., 2014. Environmental footprints show China and Europe's evolving resource appropriation for soybean production in Mato Grosso, Brazil. Environ. Res. Lett. 9 (7), 074001. https://doi.org/10.1088/1748-9326/9/7/074001.

Larson, J.M., Findeis, J.L., Smith, S.M., 2001. Agricultural adaptation to urbanization in Southeastern Pennsylvania. Agric. Resou. Econ. Rev. 30 (1), 32–43. https://doi.org/10.1017/S1068280500000526.

Legendre, P., Legendre, L., 2012. Numerical Ecology. Elsevier, Amsterdam Chapter 8: Cluster Analysis.

Letourneau, A., Verburg, P.H., Stehfest, E., 2012. A land-use systems approach to represent land-use dynamics at continental and global scales. Environ. Model. Softw. 33, 61–79. https://doi.org/10.1016/j.envsoft.2012.01.007.

Lowder, S.K., Skoet, J., Raney, T., 2016. The number, size, and distribution of farms, smallholder farms, and family farms worldwide. World Dev. 87, 16–29.

MacDonald, J.M., Korb, P., Hoppe, R.A., 2013. Farm Size and the Organization of U.S. Crop Farming, ERR-152. U.S. Department of Agriculture, Economic Research Service, Washington D.C.

Masters, W.A., Djurfeldt, A.A., De Haan, C., Hazell, P., Jayne, T., Jirström, M., Reardon, T., 2013. Urbanization and farm size in Asia and Africa: implications for food security and agricultural research. Glob. Food Sec. 2 (3), 156–165. https://doi.org/10.1016/j.gfs.2013.07.002.

Monfreda, C., Ramankutty, N., Foley, J.A., 2008. Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. Glob. Biogeochem. Cycles 22 (1). https://doi.org/10.1029/2007GB002947.

Montgomery, M.R., 2008. The urban transformation of the developing world. Science 319 (5864), 761–764. https://doi.org/10.1126/science.1153012.

Mueller, N.D., Gerber, J.S., Johnston, M., Ray, D.K., Ramankutty, N., Foley, J.A., 2012. Closing yield gaps through nutrient and water management. Nature 490 (7419), 254–257. https://doi.org/10.1038/nature11420.

Nagendra, H., Bai, X., Brondizio, E.S., Lwasa, S., 2018. The urban south and the predicament of global sustainability. Nat. Sustain. 1 (7), 341–349.

Ochieng, J., Knerr, B., Owuor, G., Ouma, E., 2016. Migration and agricultural intensification at origin: evidence from farm households in Central Africa. Migr. Dev. 6, 161–176. https://doi.org/10.1080/21632324.2015.1123837.

Onwuchekwa, I., Ukandu, I., 2015. Effect of urbanization on agricultural production in Abia State. Int. J. Agric. Sci. Res. Technol. Ext. Educ. Syst. 5 (2), 83–89. Retrieved from: http://ijasrt.iau-shoushtar.ac.ir/article_521607_df37537c006d2c542ffcdf755d3fa019.pdf.

Ospina, D., Peterson, G.D., Crépin, A.-S., 2018. Migrant remittances can reduce the potential of local forest transitions - a social-ecological regime shift analysis. Environ. Res. Lett in press. https://doi.org/10.1088/1748-9326/aaf0e.

Portmann, F.T., Siebert, S., Döll, P., 2010. MIRCA2000—global monthly irrigated and rainfed crop areas around the year 2000: a new high-resolution data set for agricultural and hydrological modeling. Glob. Biogeochem. Cycles 24 (1). https://doi.org/10.1029/2008GB003435.

Ramankutty, N., Evan, A.T., Monfreda, C., Foley, J.A., 2008. Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000. Glob. Biogeochem. Cycles 22 (1). https://doi.org/10.1029/2007GB002952.

R Core Team, 2018. R: A Language Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria Retrieved from: http://www.R-project.org/.

Samberg, L.H., Gerber, J.S., Ramankutty, N., Herrero, M., West, P.C., 2016. Subnational distribution of average farm size and smallholder contributions to global food production. Environ. Res. Lett. 11 (12), 124010. https://doi.org/10.1088/1748-9326/11/12/124010.

Satterthwaite, D., McGranahan, G., Tacoli, C., 2010. Urbanization and its implications for food and farming. Philos. Trans. R. Soc. Lond. B Biol. Sci. 365 (1554), 2809–2820. https://doi.org/10.1098/rstb.2010.0136.

Schneider, A., Friedl, M.A., Potere, D., 2009. A new map of global urban extent from MODIS satellite data. Environ. Res. Lett. 4 (4). https://doi.org/10.1088/1748-9326/4/4/044003.

Schneider, A., Mertes, C.M., Tatem, A.J., Tan, B., Sulla-Menashe, D., Graves, S.J., Patel, N.N., ... Dastur, A., 2015. A new urban landscape in East–Southeast Asia, 2000–2010. Environ. Res. Lett. 10 (3), 34002. https://doi.org/10.1088/1748-9326/10/3/034002.

de Sherbinin, A., Levy, M., Adamo, S., MacManus, K., Yetman, G., Mara, V., Razafindrazay, L., Goodrich, B., Srebotnjak, T., Aichele, C., Pistolesi, L., 2012. Migration and risk: net migration in marginal ecosystems and hazardous areas. Environ. Res. Lett. 7 (4), 045602. https://doi.org/10.1073/pnas.0906974107.

de Sherbinin, A., Levy, M., Adamo, S., MacManus, K., Yetman, G., Mara, V., Razafindrazay, L., Goodrich, B., Srebotnjak, T., Aichele, C., Pistolesi, L., 2015. Global Estimated Net Migration Grids By Decade: 1970–2000. NASA Socioeconomic Data and Applications Center (SEDAC), Palisades, NY. https://doi.org/10.7927/H4319SVC.

Seto, K.C., Reenberg, A., Boone, C.G., Fragkias, M., Haase, D., Langanke, T., Marcotullio, P., ... Simon, D., 2012. Urban land teleconnections and sustainability. Proc. Natl. Acad. Sci. U. S. A. 109 (20), 7687–7692. https://doi.org/10.1073/pnas.1117622109.

Seto, K.C., Ramankutty, N., 2016. Hidden linkages between urbanization and food systems. Science 352 (6288), 943–945. https://doi.org/10.1126/science.aaf7439.

Seto, K.C., Güneralp, B., Hutyra, L.R., 2012b. Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools. Proc. Natl. Acad. Sci. U. S. A. 109 (40), 16083–16088. https://doi.org/10.1073/pnas.1211658109f.

Siebert, S., Portmann, F.T., Döll, P., 2010. Global patterns of cropland use intensity. Remote Sens. 2 (7), 1625–1643. https://doi.org/10.3390/rs2071625.

Sloat, L.L., Gerber, J.S., Samberg, L.H., Smith, W.K., Herrero, M., Ferreira, L.G., Godde, C.M., ... West, P.C., 2018. Increasing importance of precipitation variability on global livestock grazing lands. Nat. Clim. Chang. 8 (3), 214–218. https://doi.org/10.1038/s41558-018-0081-5.

Sorichetta, A., Bird, T.J., Ruktanonchai, N.W., zu Erbach-Schoenberg, E., Pezzulo, C., Tejedor, N., Waldock, I.C., Sadler, J.D., Garcia, A.J., Sedda, L., Tatem, A.J., 2016. Mapping internal connectivity through human migration in malaria endemic countries. Sci. Data 3, 160066. https://doi.org/10.1038/sdata.2016.66.

Stuart, E.A., 2010. Matching methods for causal inference: a review and a look forward. Statist. Sci. 25 (1), 1–21. https://doi.org/10.1214/09-STS313.

Su, S., Jiang, Z., Zhang, Q., Zhang, Y., 2011. Transformation of agricultural landscapes under rapid urbanization: a threat to sustainability in Hang-Jia-Hu region, China. Appl. Geogr. 31 (2), 439–449.

Tacoli, C., McGranahan, G., Satterthwaite, D., 2015. Urbanisation, rural-urban migration and urban poverty. In: IIED *Working Paper*. London. Retrieved from: http://pubs.iied.org/pdfs/10725IIED.pdf.

Tacoli, C., 2004. Rural-urban linkages and pro-poor agricultural growth: an overview. In: Agriculture and Pro-Poor Growth Task Team Helsinki Workshop, 17–18 June 2004, Retrieved from: http://www.oecd.org/dataoecd/25/8/36562896.pdf.

Tan, P.N., Steinbach, M., Kumar, V., 2006. Introduction to data mining. In: Chapter 8: Cluster Analysis: Basic Concepts and Algorithms, Retrieved from: https://www-users.cs.umn.edu/~kumar001/dmbook/ch8.pdf.

Thebo, A.L., Drechsel, P., Lambin, E.F., 2014. Global assessment of urban and peri-urban agriculture: irrigated and rainfed croplands. Environ. Res. Lett. 9 (11), 114002. https://doi.org/10.1088/1748-9326/9/11/114002.

Turner, B., Doolittle, W.E., 1978. The concept and measure of agricultural intensity. Prof. Geog. 30 (3), 297–301. https://doi.org/10.1111/j.0033-0124.1978.00297.x.

van Vliet, J., Eitelberg, D.A., Verburg, P.H., 2017. A global analysis of land take in cropland areas and production displacement from urbanization. Glob. Environ. Change 43, 107–115.

Václavík, T., Lautenbach, S., Kuemmerle, T., Seppelt, R., 2013. Mapping global land system archetypes. Glob. Environ. Change 23 (6), 1637–1647. https://doi.org/10.1016/j.gloenvcha.2013.09.004.

Verburg, P.H., Ellis, E.C., Letourneau, A., 2011. A global assessment of market accessibility and market influence for global environmental change studies. Environ. Res. Lett. 6 (3), 34019. https://doi.org/10.1088/1748-9326/6/3/034019.

Wu, Y., Xi, X., Tang, X., Luo, D., Gu, B., Lam, S., ... Chen, D., 2018. Policy distortions, farm size, and the overuse of agricultural chemicals in China. Proc. Natl. Acad. Sci. U. S. A. 115 (27), 7010–7015. https://doi.org/10.1073/pnas.1806645115.

Zuo, L., Zhang, Z., Carlson, K.M., MacDonald, G.K., Brauman, K.A., Liu, Y., Zhang, W., Zhang, H., Wu, W., Zhao, X., Wang, X., Liu, B., Yi, L., Wen, Q., Liu, F., Xu, J., Hu, S., Sun, F., Gerber, J.S., West, P.C., 2018. Progress towards sustainable intensification in China challenged by land-use change. Nat. Sustain. 1, 304–313. https://doi.org/10.1038/s41893-018-0076-.