# Developing country-wide farming system typologies

## An analysis of Ethiopian smallholders' income and food security

by

**Esther Boere**
**Aline Mosnier**
**Géraldine Bocquého**
**Alessandra Garbero**
**Tamás Krisztin**
**Petr Havlík**
**Thomas Elhaut**

**32** *IFAD RESEARCH SERIES*

The IFAD Research Series has been initiated by the Strategy and Knowledge Department in order to bring together cutting-edge thinking and research on smallholder agriculture, rural development and related themes. As a global organization with an exclusive mandate to promote rural smallholder development, IFAD seeks to present diverse viewpoints from across the development arena in order to stimulate knowledge exchange, innovation, and commitment to investing in rural people.

**Authors:**
Esther Boere, Aline Mosnier, Géraldine Bocquého, Alessandra Garbero, Tamás Krisztin,
Petr Havlík, Thomas Elhaut

# IFAD

Investing in rural people

# Developing country-wide farming system typologies

An analysis of Ethiopian smallholders' income and food security

by

**Esther Boere**
**Aline Mosnier**
**Géraldine Bocquého**
**Alessandra Garbero**
**Tamás Krisztin**
**Petr Havlík**
**Thomas Elhaut**

# Acknowledgements

## About the authors

**Esther Boere** is an agricultural economist in the International Institute for Applied Systems Analysis (IIASA) Ecosystem Services and Management (ESM) program. Her research focuses on analysing the driving factors of land use change through short- and long-term responses to agricultural shocks and regional- and country-level impacts on food security, deforestation and restoration policies. Prior to joining IIASA, Dr Boere received her PhD from Wageningen University on the effect of the European Union's Common Agricultural Policy reform on land use changes.

**Aline Mosnier** is a research scholar at IIASA in the ESM program. She holds a PhD in agricultural economics from the University of Life Sciences of Vienna (Austria) and a Master's degree in development economics from the Centre d'Etudes et de Recherches sur le Développement International (CERDI, France). She has contributed to the development of GLOBIOM, a global partial equilibrium model on land-based sectors, since 2008. Her main research topics are tropical deforestation and climate change impacts, mitigation and adaptation in agriculture.

**Géraldine Bocquého** is a researcher at the French National Institute for Agronomic Research (INRA). She holds a PhD in economics from AgroParisTech (France) and two master's degrees from the same institute, in rural development and in environmental economics. At IIASA, she worked on alternative development pathways in Congo Basin countries, Brazil and Ethiopia. More generally, her research interests lie in land use policies, services provided by forest and agricultural ecosystems, and decision makers' behaviour with respect to risk and time.

**Alessandra Garbero** is the senior econometrician in the Research and Impact Assessment (RIA) Division of the International Fund for Agricultural Development (IFAD). Her work focuses on impact assessment methodologies and applied econometrics. She supports RIA strategic management related to the impact assessment production cycle and is in charge of the estimation and projection of corporate impacts. Her first degree is in economics, with specializations in demography and statistics. She also holds a PhD in epidemiology and population health from the London School of Hygiene and Tropical Medicine.

**Tamás Krisztin** is a research scholar at IIASA. He studied spatial econometrics and completed his PhD at the WU Vienna University of Economics and Business. His current research focuses on econometric downscaling techniques and time-series analysis of agricultural commodities.

**Petr Havlík** is Deputy Program Director of the IIASA ESM program and coordinates research on the use and development of forest and agricultural sector optimization models at both the European and global levels. He is leading the development of the GLOBIOM model, applied to academic and policy questions on the future development of the land and food sector, climate change impact and mitigation, and ecosystems services. He holds a PhD in agricultural economics from the University of Montpellier and Mendel University.

**Thomas Elhaut** is currently public policy economics counsel, affiliated with WeCitizens (Belgium). At the time of the research reported here, he was director of the Statistics and Studies for Development Division at IFAD, on sabbatical leave, as senior research scholar, at IIASA and, as visiting senior fellow, at the Brooks World Poverty Institute, University of Manchester.

# Table of contents

# Abstract

Smallholder farms are often the focus of strategies to reduce poverty, inequality and hunger. They are also a very diverse group, leading to calls for more context-specific strategies to support smallholder farms. Until recently, a lack of both household and macro-level data prevented policymaking tailored to the context of smallholder farms, i.e. what types of farms at what locations may benefit most from, for example, better irrigation infrastructures, intensification or improved access to markets? Increased effort to collect more information at different scales and scopes enables us to systematically establish the context in which smallholder farms operate. The objective of this study is to combine household survey, agricultural census and land cover data to analyse food security and poverty at the micro (farm) and macro (regional and country) levels. To do so, the study developed a novel methodology that combines different data sources to establish country-wide farm typologies that enable an analysis of poverty and food security that is both farm-system specific and spatially explicit. To test our methodology, we analysed the poverty and food security situation of Ethiopian smallholder farms. Our results show that the combination of activities and agroecological zone largely determined the food security situation of the farm. In terms of poverty, all farms were below the poverty line; in particular, pure livestock farms operating in poor biophysical conditions exhibited low market interactions and high incidence of poverty. The farming-system- and location-specific poverty and food security indicators developed in this study provide guidance for better targeting of policy strategies.

# **1** Introduction

More than 80 per cent of the 570 million farms in the world operate land areas smaller than 2 hectares. These farms are often the focus of strategies to reduce poverty, inequality and hunger (Lowder, Skoet and Raney, 2016). A wide array of goals and targets has emerged for environmental sustainability, food security and agricultural productivity, most notably those listed in the Sustainable Development Goals (SDGs) (Sachs, 2012). However, smallholder farms are also a very diverse group, which calls for context-specific strategies to support them (Graeub et al., 2015; Rapsomanikis, 2015). It is therefore of critical importance to define spatially targeted and context-specific interventions. Spatially targeted interventions can increase lagging areas' connection to the market, mobilize untapped resources for development, and take the regional importance of sectors into account without leading to inefficient and costly investments (Herrero et al., 2014).

Research to date has categorized small and family farms, focusing on sustainable development paths for the agricultural sector and the role of smallholder farms (Graeub et al., 2015; Wiggins et al., 2010; Lowder, Skoet and Raney, 2016). These studies all used agricultural census data to categorize smallholder farms and analyse their complexities and differences. Using census data enabled them to provide a country- or even world-wide picture; however, they were unable to account for potential heterogeneity inside administrative units. The scope of these analyses therefore remained limited. In an attempt to obtain more detailed information on such heterogeneity (e.g. farm management, food consumption and sales), there has been an increased effort to collect more information at the farm household level (World Bank, 2016). Subsequently, these data were used to analyse the determinants of food security and food availability (Frelat et al., 2016). However, household surveys are generally carried out at limited scales and therefore provide limited representativeness.[1] Remote sensing data are able to capture spatial and temporal context at different levels of analysis and have therefore been used to analyse economic growth and poverty-environment relationships (Keola, Andersson and Hall, 2015; Mertens et al., 2000; Watmough et al., 2016). However, unless combined with statistics, identifying the causes of these relationships is more difficult.

In this paper, we develop a novel methodology that combines different data sources to establish country-wide farm typologies that enable analysis of poverty and food security that is both farm-system specific and spatially explicit. The methodology reconciles different data types to better capture spatial heterogeneity among smallholder farms at both the micro and macro levels. We used household survey data to compute average land use practices, productivity, consumption patterns and market indicators at the farm level for each farming system and extrapolated this information to the areas which are not covered by the household survey using a combination of multinomial regression model and a harmonization procedure with agricultural statistics and land cover map. This resulted in a comprehensive

---

1. In case of large-scale household survey data, representativeness may not always be a problem. However, most farm household survey data are unfit to be aggregated to subnational or national levels, which is necessary for an ex ante policy impact evaluation.

mapping and characterization of smallholders across the whole territory which can be used for policy design, linked to (for example) governmental aims or the SDGs. This methodology was tested and applied to the Ethiopian context to develop a farming systems typology and analyse smallholders' food security and income.

Ethiopia is used as a case study to analyse smallholders' food security and income. Ethiopia is one of the fastest growing economies in the world, with an average annual growth in gross domestic product (GDP) of 10 per cent (Paul et al., 2016). The largest share of the GDP of the country (46.9 per cent) comes from agriculture (Diao, Hazell and Thurlow, 2010). Smallholder farms represent the majority of the rural population in Ethiopia and, despite an average farm area of only 1 hectare and low productivity, they are responsible for about 90 per cent of the total agricultural output (Dorosh and Rashid, 2012; Hanjra, Ferede and Gutta, 2009). A rapidly increasing population (expected to double by 2050), slow productivity growth and climate-related disasters such as droughts increase food insecurity. Consequently, competition for available land, water, energy and other inputs is increasing, putting pressure on rural livelihoods, income and food security (Bryan et al., 2009; Garnett et al., 2013).

The remainder of the paper is organized as follows. In the next section, the development of the novel methodology is described, including how the various data were combined to analyse the scale and scope of the poverty and food security situation of smallholder farms in Ethiopia. The analysis was done on the third administrative level, the *woredas*. Section 3 analyses the current state of poverty and food security. Our conclusions, discussion and policy recommendations follow.

# **2** Methods and data

To evaluate each farming system in each spatial unit in terms of the favourability of biophysical constraints, activities and management, smallholder farms were categorized according to their agroecological zone (AEZ), dominant activities and management.

The AEZ captures the integration of smallholder farms into the broader agro-economic context in which they operate, which defines their constraints and opportunities and drives their livelihood and production choices. Smallholders often operate mixed crop-livestock farming. The biophysical conditions such as climate, topography and soil type, as well as historical preferences for certain products, defines their portfolio of activities. Access to markets and the orientation of the farm defines the management system.

A combination of household survey, agricultural census and land cover data was used to classify different types of smallholder farming systems and their socio-economic characteristics throughout the country. We describe the different steps and datasets used to obtain a country-wide representation of smallholders in Ethiopia and their food security and poverty status.

## 2.1 The household survey and the farming systems typology

We used the Ethiopian Rural Socioeconomic Survey (ERSS) dataset from the Living Standards Measurement Study – Integrated Surveys on Agriculture (LSMS-ISA) project led by the World Bank and funded by the Bill & Melinda Gates Foundation (World Bank, 2016) to define smallholders' main farming systems and their characteristics. The first wave of the survey consisted of three rounds of questionnaires administered between September 2011 and March 2012. The ERSS dataset covers information on crop and livestock technology (e.g. variable and fixed inputs used, yields obtained), market access (e.g. degree of inputs bought and outputs sold to the market) and socio-economic household information (e.g. farm size). In total, 3,969 households are included in the ERSS, of which 3,216 had enough household, crop and livestock information to be used for this analysis.

Using expert knowledge, statistical clustering and literature review (Dixon, Gulliver and Gibbon, 2001), we grouped smallholder farms from the household survey into five types: smallholders who focus solely on keeping livestock ("pure livestock") and mixed crop-livestock farms, dominated by four types of main crop combinations: millet-sorghum, maize, wheat-barley-teff and perennial crops. Mixed crop-livestock farms were allocated to a certain group based on the most important crop cultivated in terms of area. If this was not millet, sorghum, maize, wheat, barley, teff or a perennial, the second or third most important crop was considered. The remaining farms that could not be classified (e.g. those focusing mostly on horticultural crops) were dropped from the analysis (179 farms).[2]

---

2. These were 5.5 per cent of the sample and, especially with increasing urbanization and population density in areas with good market access, horticultural farms might be on the increase in Ethiopia. However, because horticultural farms are typically not land-based and the focus of this paper is on the spatial allocation of smallholder farms, especially in rural areas, these types of farms largely fall outside the scope of the paper.

**Figure 1** Number of farms by *woreda* (third administrative level) that are included in the ERSS survey



Number of farms by *woreda*

- 1 - 5
- 6 - 14
- 15 - 22
- 23 - 34
- 35 - 103
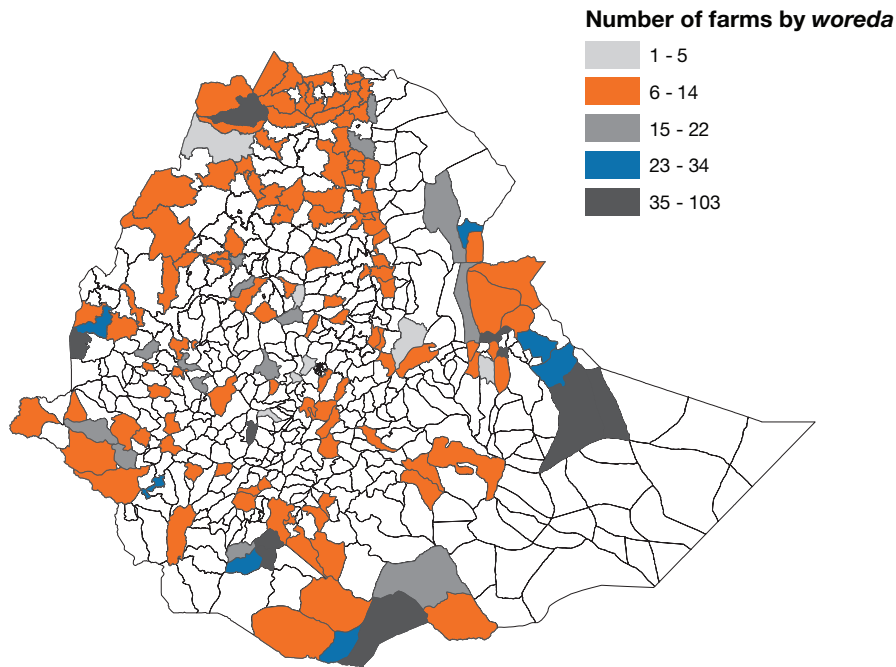
For the farms classified, we computed number, production, use, revenues and costs for animal activities and average land use per activity, input use, input costs, productivity, and the share of the production which is used for home consumption, sale or other use for crop production, as well as the income obtained from off-farm activities. However, the household survey data do not encompass the whole country,[3] nor are they representative at the *woreda* level (figure 1).

## 2.2 Prediction of the distribution of farming systems by *woreda* using a multinomial regression model

To overcome the problem of limited spatial coverage of the household survey, we predicted the distribution of the five classified farming systems by type in all *woredas* using a multinomial regression model.

We used three main criteria to define farmers' decisions on what to produce and the intensity and use of their production: the agroecological growing areas, the orientation of the farm and the decisions of neighbouring *woredas*. The AEZ limits the scope of potential farm activities. The orientation of the farm represents socio-economic characteristics at the *woreda* level that determine the choice of the most important activity in the main season as well as the wider portfolio of on-farm activities (Demissie and Legesse, 2013). And smallholders' production decisions are likely to be spatially dependent; collective decisions to sell products may lead to new markets, which in turn influence production decisions of nearby farms. Spatial dependence was incorporated in the model by using the spatially lagged characteristics of

---

3. Addis Ababa, three zones (second administrative level) in the Afar region, six zones in the Somali region and many *woredas* (third administrative level) were not surveyed. Moreover, in most of the surveyed *woredas*, only between 6 and 14 farms were found to have reliable information.

neighbouring *woredas* as explanatory variables. This is in effect a spatial lag on the explanatory variables.[4] To establish which *woredas* are neighbours to each other, we used a *k*-nearest neighbour spatial structure, based on the shortest possible distance of the *woreda* centroids, with *k* = 10.[5] The *woreda*-specific explanatory variables are reported in table 1. They are obtained from Tadese et al. (2006), except for the AEZ, which has been obtained from the Ministry of Agriculture, and size of the *woreda*, which was obtained from the GAUL dataset (IES-JRC, 2000). Descriptive statistics of the covariates can be found in the annex (table A1).

**Table 1** Variables used for extrapolation

| Variable | Description |
| --- | --- |
| **Dependent variable** | |
| Activity set | Pure livestock, millet-sorghum, maize, wheat-barley-teff, perennial farms |
| **Ability to grow certain crops and/or raise livestock** | |
| Average elevation | Average *woreda* elevation in metres above mean sea level |
| Average slope | Mean slope as percentage |
| Average rainfall | Mean monthly rainfall in millimetres |
| Dominant crop | Dominant crop in the *woreda*: dummy for barley, maize, millet, pastoral, sorghum, teff or wheat |
| Secondary crop | Secondary crop in the *woreda*: dummy for barley, maize, millet, pastoral, sorghum, teff or wheat |
| AEZ | Dummy for drought-prone area and arid pastoral area (moisture reliable treated as base category) |
| Area | Size of the *woreda* in square kilometres |
| **Socio-economic context of the farm** | |
| Population density | Population density in persons per square kilometre |
| Road density | All weather road density (m/km2) |
| Cooperative and type in the *woreda* | Split into coffee marketing, multi-purpose, no cooperative, saving and credit, and vegetable marketing |
| Number of cooperatives | Number of cooperatives in *woreda* |
| Total members involved | Total members active in a cooperative in a *woreda* |
| Spatially lagged variables | For population, elevation, slope, rainfall, road density and area |

4. Note that, while it would technically be possible to also estimate a model with spatial lags in the error terms or the dependent variables, these models could not be used for extrapolation of the data. This is due to the fact that these spatial autoregressive specifications assume a closed spatial system in the lag and thus cannot be used to extrapolate over missing data (for details, see LeSage and Pace, 2009).

5. Note that we experimented with using alternative neighbourhood specifications, such as contiguity or alternative specifications of *k*. We selected *k* = 10, based on the highest Bayesian Information Criteria value. The comparison results are available upon request.

In each *woreda*, farmers face a discrete choice problem in selecting one of the five dominant activity sets, based on their own, as well as the neighbouring *woredas'* attributes listed in table 1. The multinomial logit model is the classic choice from a microeconometric point of view to model such a problem (see e.g. Cameron and Trivedi, 2005). To run the regression, we assumed that information in the *woredas* is missing-at-random and not structurally missing. Based on this assumption, we used multiple imputations to predict farmers' choices in all *woredas* based on the 143 *woredas* for which smallholder decisions are available from the household survey.

More formally, let us denote the observed choice of farm portfolio $j$ – out of $J$ non-intersecting choice combinations of activity portfolios – in *woreda* $i$ ($i = 1,…, N$), as $\gamma_{ij}$. Recall that the possible choices of activity portfolios consist of: pure livestock, millet-sorghum, maize, wheat-barley-teff and perennial farms. Our goal is to extrapolate from the $N$ observed *woredas* to $M$ missing *woredas*, where the total number of *woredas* in Ethiopia is denoted by $N^*$ and $N^* = M + N$.

We observe a multinomial sample $\gamma_i = \{\gamma_{ij}\}^J_{j=1}$ that records the number of farmers choosing activity portfolio $j = 1, …, J$[6] and the total number of responses $n_i$ in *woreda i*. According to the logistic link, the probability $p_{ij}$ of randomly drawing a single response from the $j$th category in *woreda i* is given as:

$$p_{ij} = \frac{\exp \mu_{ij}}{\sum^J_{i=1} \exp \mu_{ik}} \tag{1}$$

where the log odds $\mu_{ij}$ is modelled in terms of $P$ explanatory variables and their spatially lagged counterparts:

$$\mu_{ij} = \alpha_j + x_i \beta_j + \sum^{N^*}_{q=1} w^*_{iq} x^*_i \eta_j \tag{2}$$

where $\alpha_j$ is the intercept, $x_i$ is a $1{\times}P$ vector of *woreda*-specific explanatory variables, with an associated $P{\times}1$ parameter vector $\beta_j$, and $w^*_{iq}$ is a typical element of the $N^*{\times} N^*$ exogenously given spatial weight matrix $W$. If *woreda i* and *woreda q* are neighbours, $w^*_{iq} > 0$, otherwise $w^*_{iq} = 0$. Moreover, the spatial weight matrix is defined in a fashion that no *woreda* is a neighbour to itself, therefore, $w^*_{iq} = 0 \; \forall \; i=q$. $W$ is assumed to be row-standardized, that is $\sum^{N^*}_{q=1} w^*_{iq} = 1 \; \forall$ $i = 1,…, N$. Note that $W$ and the associated $1{\times}P$ vector of explanatory variables $x^*_i$ extends across all *woredas*, even the $M$ missing ones. This is due to the fact that even though we do not observe smallholder farmers' choices in the *woredas*, we have a full set of explanatory variables $x^*_i$. $\eta_j$ is the $P{\times}1$ parameter vector of spatially lagged coefficients. Note that, for the purposes of identification, we set $\alpha_J$, $\beta_J$ and $\eta_J$ equal to zero (see e.g. Holmes and Held, 2006).

---

6. Where the 1, …, J activities consist of pure livestock, millet-sorghum, maize, wheat-barley-teff and perennial farms introduced in earlier section 2.1.

We favour a Bayesian approach for the purpose of data imputation, since it not only allows us to naturally obtain uncertainty bounds over the extrapolated data (Frühwirt-Schnatter and Frühwirt, 2012; Holmes and Held, 2006), but also allows us to easily include variable selection in our model. We use an efficient Gibbs sampling algorithm to estimate the multinomial logit model in equation (1), in the spirit of Polson, Scott and Windle (2013). The main difference from their approach is that we use a Bayesian variable selection prior, as pioneered by George and McCulloch (1993) and Kuo and Mallick (1998).

Since we use a Bayesian estimation procedure, we have to set up a set of priors for the regression coefficients. For easier notation, let us denote $z_i = [\iota_N \; x_i \; \sum_{q=1}^{N} w_{iq} \, x_q]$, with ($q = 1,\dots, N$) and $\theta_j = [\alpha_j \, \beta_j' \, \eta_j']'$.

To implement Bayesian variable selection, we use spike-and-slab priors for on $\theta_{pj}$ (p=1,...,P):

$$p(\theta_{pj} \mid \delta_{kj}) = \delta_{pj} \, N(0, \tau_1^2) + (1 - \delta_{pj}) N(0, \tau_0^2) \tag{3}$$

where $\delta_{pj}$ is a binary random variable and the prior variances are chosen so that $\tau_0^2 \gg \tau_1^2$.[7] We set the priors for $\tau_0^2$ and $\tau_1^2$ using the automated procedure described by George and McCulloch (1993). We used 20,000 cycles for our Gibbs sampling algorithm, where the first 2,000 draws were discarded as burn-in. Posterior means and standard deviations, the residual deviance, as well as the model's Bayesian information criterion (BIC) are reported in the annex (table A2). Based on the obtained posterior density $p(\theta_j \mid \cdot)$, we can extrapolate to the $M$ missing *woredas*. The resulting uncertainties associated with the extrapolation are summarized per *woreda* in the annex (figure A1).

To ensure that our spatial multinomial logit model is robust, we explored multiple alternative models. As a first step, we considered using different spatial weight matrix specifications. For this, we considered spatial weight matrices based on first-order contiguity and based on three to 15 $k$-nearest neighbours. We compared the resulting model estimates based on the BIC, and chose $k = 10$ with a BIC of 4,274.84. Moreover, we compared the model to a non-spatial specification (where $\eta_j = 0, j = 1,..., P$). We evaluated the performance of these two models based on the deviances of the value of the BIC. Furthermore, to assess the predictive performance of the model, we used 20 randomly selected *woredas* as a holdout sample and thus evaluated the predictive performance of the spatial and non-spatial specifications, as measured on their root mean squared errors (RMSE) and their log-predictive scores. All of these criteria favoured the spatial model. The results of the tests are summarized in table 2.

---

7. This so-called spike and slab prior is a useful tool for drawing posterior inference on the relative importance of a variable to the specific choice class. In the case of $\delta_{pj} = 0$, $\theta_{pj}$ is given a relatively flat and agnostic prior (the so-called slab), whereas in the case of $\delta_{pj} = 1$, the Gaussian prior on $\theta_{pj}$ has a large part of its centre mass on zero and therefore the parameter is aggressively shrunk towards zero. Otherwise, we use an auxiliary mixture sampling procedure, with an associated Pólya-Gamma prior, as suggested by Polson, Scott and Windle (2013). The conditional posteriors take on standard closed form, therefore we refrain from reporting them.

**Table 2** Specification tests for the spatial and non-spatial multinomial logit models

| | Specification | 1fs | 2fs | 3fs | 4fs | 5fs |
|---|---|---|---|---|---|---|
| RMSE | Non-spatial | 4.70 | 5.92 | 4.84 | 6.13 | 1.69 |
| | Spatial | 4.59 | 5.90 | 4.86 | 6.19 | 1.65 |
| Log-predictive score | Non-spatial | -4.06 | -2.60 | -0.62 | -7.46 | -4.00 |
| | Spatial | -4.58 | -5.50 | -1.73 | -9.82 | -4.31 |
| BIC | Non-spatial | 3 561.97 | | | | |
| | Spatial | 3 596.04 | | | | |
| Deviance | Non-spatial | 9 821.70 | | | | |
| | Spatial | 10 072.32 | | | | |

BIC, Bayesian information criterion; RMSE, root mean squared error.

## 2.3 Adjustment of farming systems' parameters to match agricultural statistics per *woreda*

The sum of the total area by crop and animal numbers by type for each activity set resulting from the household survey resulting from the calculations in the previous section does not equal the sum of the total area by crop and animal numbers by type from the agricultural census data. To ensure representativeness at the *woreda* level, we minimize the variance (e.g. the sum of the squared deviations between the area of teff over all farming systems in a *woreda* and the area of teff from the agricultural census data in the same *woreda*). At the end of this procedure, we obtained the distribution of farms in each farming system for each *woreda* – e.g. in a *woreda* not covered by the survey, 30 per cent of the farms are allocated to the farming system "wheat-barley-teff", 10 per cent to "pure livestock" and 50 per cent to "maize". The results are described in section 3. We refer to this as the "harmonization process". The crops used in the harmonization process are all important staple crops (teff, wheat, barley, maize, sorghum, millet, rice, potato, sweet potato), the most important pulses (beans, chickpea), some oilseeds (sesame, groundnut, rapeseed, sunflower, soybean) and perennials (coffee). Cows, sheep, goats and poultry represent the animals used in the harmonization process. The harmonization procedure results in data in terms of crop area, livestock numbers, inputs and yields by activity set and *woreda* that, when aggregated over the activity sets, is fully consistent with *woreda*-level data for the base year.

The Ethiopian Agricultural Sample Enumeration (EASE) census data of 2011-2012 was used as the agricultural census data for harmonization (CSA, 2011). The EASE census data provides data in six domains: land use, area, production, livestock, farm practice and socio-economic situation at the *woreda* level. Like the ERSS, the EASE does not cover three zones in the Afar region and six zones in the Somali region. For these regions, we used the GLCshare land cover map for the cropland area and FAO-ILRI livestock numbers (IES-JRC, 2000); with 8.6 million hectares of cropland, the total area of cropland from the land cover data is close to the 8.9 million hectares reported in the EASE statistics.

In order to analyse the food security and poverty status across smallholders, we need to have information on consumption and production patterns by farm. We estimated the total number of farms for each *woreda* using the rural population census data by *woreda* and the average farm household size from the ERSS. Once we know the number of farms in each farming system in each *woreda*, we can obtain the total cultivated area for each crop, the total production for each crop and the total livestock numbers for each *woreda* using the average production parameters per farm computed per farming system at the AEZ level based on the ERSS data.

## 2.4 Poverty and food security

Three indicators were established to analyse the poverty and food security status of smallholder farms: income, food security and food diversity. Income was computed as the net revenues from all crop and livestock products plus the revenues derived from off-farm work in United States dollars per person per day. The net revenues from all crop and livestock products were computed as the on-farm production multiplied by price minus variable costs. The income per farm[8] is then compared with the poverty line of US$1.25 per capita per year (2008 global line set by the World Bank), using the average 2011 exchange rate of 17.5 Ethiopian birr per United States dollar.

Food security was measured by the combination of on-farm food availability and the cost of food that needs to be bought to achieve the average regional diet. Food availability was calculated by the total on-farm livestock and crop production in calories per capita per day. It was compared with the minimum daily intake for Eastern Africa, which is 2,245 kcal (van Wesenbeeck, Keyzer and Nubé, 2009). Similarly to the income-based food security indicator developed by Antle, Adhikari and Price (2015), we calculated the kilocalories of each food type that need to be bought to fulfill the average regional diet and convert them to United States dollars. To respect local dietary preferences and availability of different food types, we used the Ethiopian national food consumption survey to calculate the share by food type consumed by the first administative zone in Ethiopia (EPHI, 2013). The cost of achieving the average regional food diet was compared with the net revenues achieved through sales to market and off-farm income.

---

8. We estimated the total number of farms by *woreda* using the rural population census data by *woreda* and the average farm household size from the ERSS. Once we know the number of farms in each farming system in each *woreda*, we can obtain the total cultivated area by crop, the total production per crop and the total livestock numbers by *woreda* using the average production parameters per farm computed per farming system at the AEZ level based on the ERSS data.
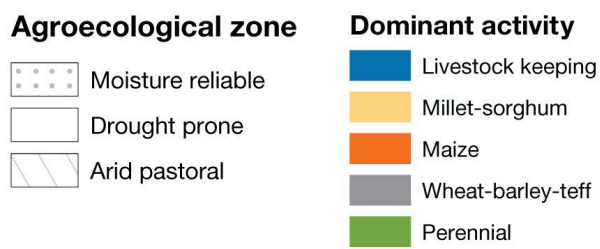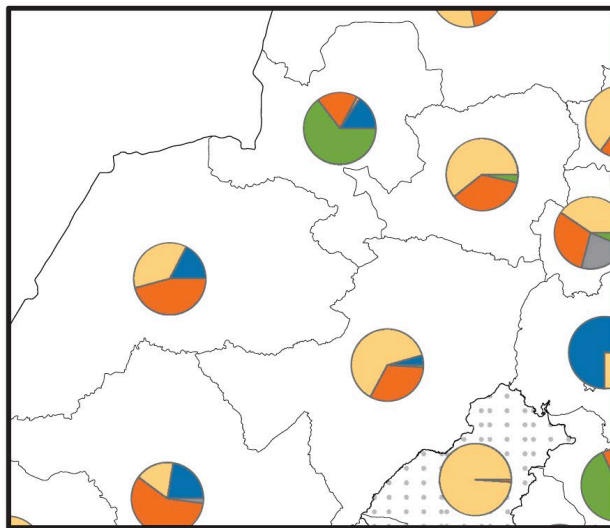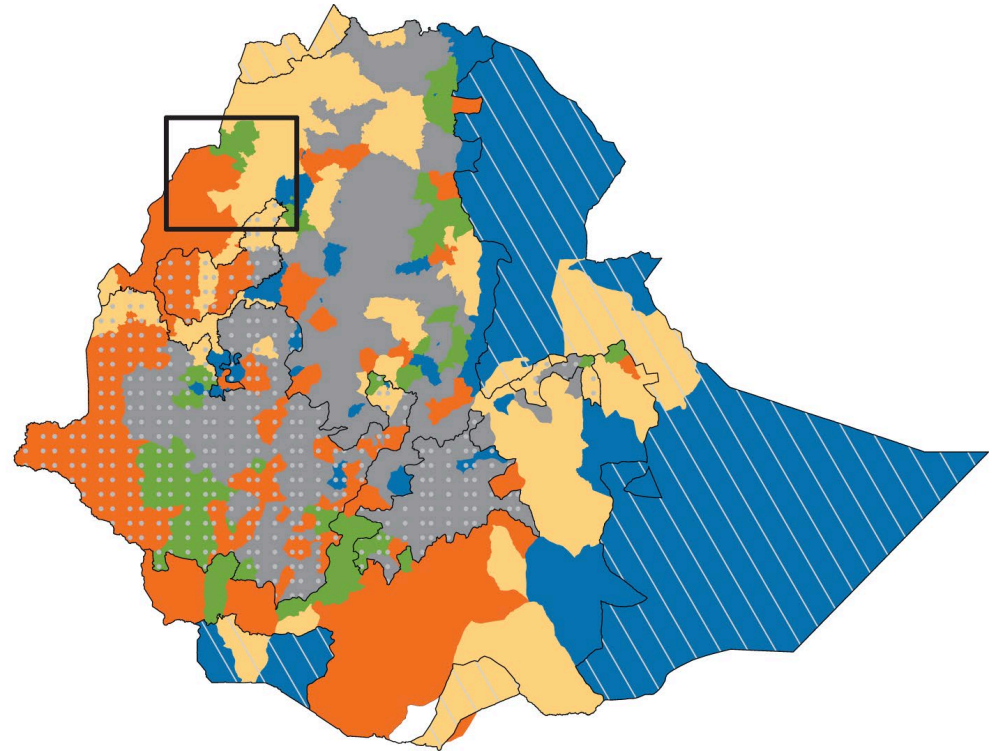
# **3** Results

In this section, we discuss the distribution of different types of smallholder farm systems, their characteristics in terms of farm size and farm management, and the poverty and food security status at the farm system and national levels.

## 3.1 Distribution of farming systems

The dominant combination of activity set and AEZ by *woreda* and the distribution of farming systems for the subset in the rectangle is shown in figure 2. Farmers' production activities are strongly dependent on the biophysical conditions in which they operate, leading to the exclusion of the wheat-barley-teff and perennials activity sets from the arid-pastoral zone. The pure livestock and millet-sorghum activity sets dominate in almost all of the arid AEZ along the border with Somalia, which is less suited for crop production. The millet-sorghum activity set is also widely spread in the north of Ethiopia. The area allocated to maize spreads out over the whole of the western half of Ethiopia but is dominant in the west and south of the country. The wheat-barley-teff activity set is dominant in the highlands, which run north-south through the middle of Ethiopia. The perennial activity set is dominant in a few *woredas*, mostly in the zones of Southern Nations, Nationalities, and Peoples' (SNNPR) and Oromia Regions with generally sufficient rainfall.

The combination of dominant activity set and AEZ by *woreda* can be related to the existing literature on farming systems in sub-Saharan Africa. The area dominated by pure livestock farms fits quite well with the pastoral and arid-pastoral systems of Dixon, Gulliver and Gibbon (2001) and the pastoral and agropastoral zone of Otte and Chilonda (2002). Especially in the north-eastern part, this correlates with a high cattle population density. In general, the farming systems in this study show a higher tendency towards crop-based farming systems compared with those of Dixon, Gulliver and Gibbon (2001). This may be due to the rapid expansion of cropland over the past decade: total arable land increased by 54.8 per cent between 2001, the year of the study of Dixon, Gulliver and Gibbon (2001), and 2012, the year of the ERSS data (FAO, 2016).

**Figure 2** Combination of dominant activity set and AEZ by *woreda* (top) and the distribution of farming systems for the subset in the rectangle (bottom)



**Agroecological zone**

- [dotted pattern] Moisture reliable
- [blank] Drought prone
- [hatched pattern] Arid pastoral

**Dominant activity**

- [blue] Livestock keeping
- [yellow] Millet-sorghum
- [orange] Maize
- [grey] Wheat-barley-teff
- [green] Perennial

## 3.2 Poverty and food security by farming system

Figure 3 graphically shows the poverty, food availability and food diversity measures. The calculations comprise the average extrapolated and harmonized data on yields, crop area, inputs, animals and animal products, and use of products by smallholder farming system. For the poverty indicator, additional information on prices that are not available in the national statistics were necessary. Prices of variable inputs and crop outputs were therefore taken as averages from the ERSS data.

The food security indicator was measured by the combination of food availability and food diversity. Food availability was calculated by the total on-farm livestock and crop production in calories per capita per day and the fraction that gets consumed, sold or used for another purpose (middle graph of figure 3) and compared with the minimum daily intake for Eastern Africa (dotted line). Food diversity was measured following a similar method as the income-based food security indicator developed by Antle, Adhikari and Price (2015). To respect local dietary preferences and availability of different food types, we used the Ethiopian national food consumption survey to calculate the share by food type consumed for each of the main zones in Ethiopia (EPHI, 2013). Based on these shares, we calculated the kilocalories of each food type that needed to be bought to fulfill dietary wishes, after self-consumption has been accounted for. The necessary kilocalories to complement the diet were converted to United States dollars. The cost of food diversity was compared with the profit achieved through market interaction (dotted line in bottom graph in figure 3).

When total net revenues of all on-farm production activities were added to the revenues obtained from off-farm work, none of the farming systems meet the US$1.25 per person per day poverty line. On average, a farmer's net revenues are around US$0.5, of which only 6 per cent is obtained from off-farm work, and other revenues are split about evenly between arable and livestock activities (45 and 48 per cent, respectively). The drought-prone zone contains both the highest and lowest revenues per person per day (US$0.92 per person per day for the perennial activity set and US$0.26 for the pure livestock activity set). The perennial activity set had a small land size, indicating that, unlike what is suggested in the literature, not only the size of the farm, but also the value of the crop produced is important to determine poverty status (Hanjra, Ferede and Gutta, 2009; Jayne et al., 2003). The highest commercialization of agricultural activities was observed for millet-sorghum in the arid-pastoral zone (20 per cent), and maize (14.5 per cent) and wheat-barley-teff (15 per cent) in the moisture-reliable zone. The highest shares of off-farm work were found for pure livestock (15.6 per cent) and maize (16.6 per cent) in the moisture-reliable zone. Both findings are in line with Chamberlin, Pender and Bingxin (2006), who found higher commercialization of cereals and more engagement in off-farm work in the moisture-reliable zone compared with the drought-prone zone. On average, 11 per cent of smallholder farmers' total revenues is obtained from off-farm work, which is close to the 12 per cent found by Frelat et al. (2016) for households without enough food available and 14 per cent found by Rapsomanikis (2015).

# Figure 3 Measures of poverty, food availability and food diversity

## Poverty



## Food availability



## Food diversity
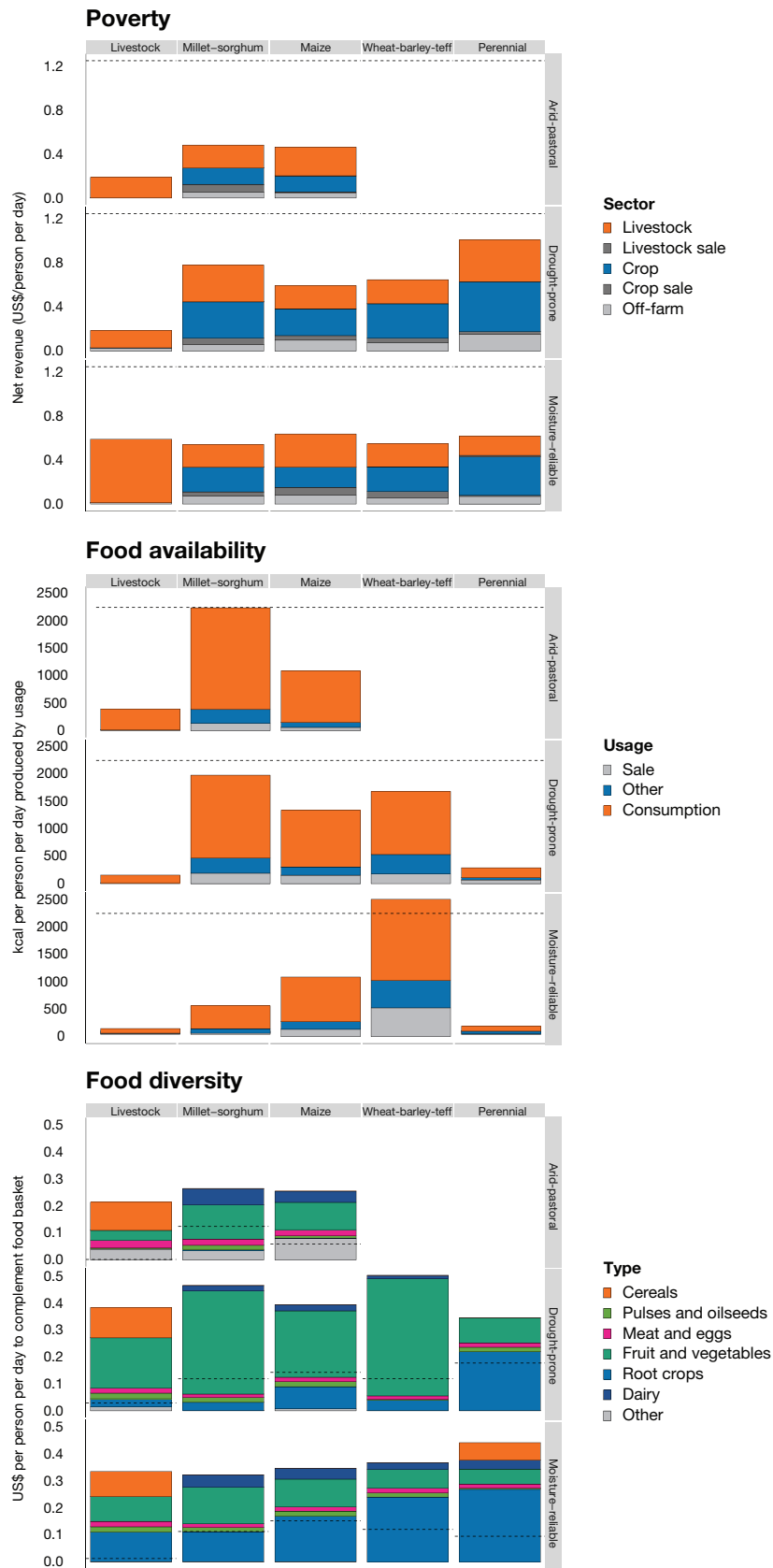
The importance of the biophysical context to enhance productivity is highlighted by the kilocalories obtained per farming system (middle graph of figure 3). With respectively 2,245, 1,976 and 1,680, and 2,403 kcal per capita, only millet-sorghum in the arid-pastoral zone, millet-sorghum and wheat-barley-teff in the drought-prone zone, and wheat-barley-teff in the moisture-reliable zone produce at or around the minimum daily intake of 2,245 kcal per capita (van Wesenbeeck, Keyzer and Nubé, 2009). This finding – that consumption of self-produced crops does not cover the food need for most farming systems – is in line with the results of Frelat et al. (2016) and leads us to conclude that per capita availability of calories is closer to the 1,810 estimated by FAO (2017) than to the 2,046 estimated by van Wesenbeeck, Keyzer and Nubé (2009). Crop, and especially cereal, production accounts for the largest source of energy, which is reflected by the very low amount of energy obtained from own production for pure livestock and perennial farms across all AEZs. This is in line with the results of Frelat et al. (2016), who found that 60 per cent of food availability comes from crop production.

When we focus on food diversity the picture is different (bottom graph of figure 3). Because of the focus on high-value nutrient-rich livestock products, smallholder farms in the arid-pastoral zone have lower costs to complement their diet compared with the drought-prone and moisture-reliable zones (US$0.26 compared with US$0.35 and US$0.36 per person per day). Only for the pure livestock dominant activity and the perennial activity in the moisture-reliable zone do diets need to be complemented with cereals. Diets need to be complemented mostly by fruits and vegetables in the drought-prone zone, root crops in the moisture-reliable zone and a mix of different items in the arid-pastoral zone.

## 3.3 Poverty and food availability by region

High population density and corresponding difficulties with access to land suitable for agricultural activities are prevalent throughout Ethiopia and are likely to persist as rural population densities continue to increase (Josephson, Ricker-Gilbert and Florax, 2014). To assess where Ethiopia is most prone to poverty and food insecurity, we calculated the food energy and gross revenues obtained from production (bottom maps of figure 4). The incidence of food insecurity and poverty is especially high in the arid-pastoral zone. The areas that are dominated by the perennial activity perform well in terms of net revenue. This confirms the results of figure 3. The depiction that per capita calorie intake is lowest in the pastoralist areas and food insecurity is higher in the Tigray Region and SNNPR is in line with Dorosh and Rashid (2012).

**Figure 4** Average farm size (top left), share of land intensified (top right), kcal/person per day produced (bottom left) and gross revenues from production (bottom right) per person per day by *woreda*
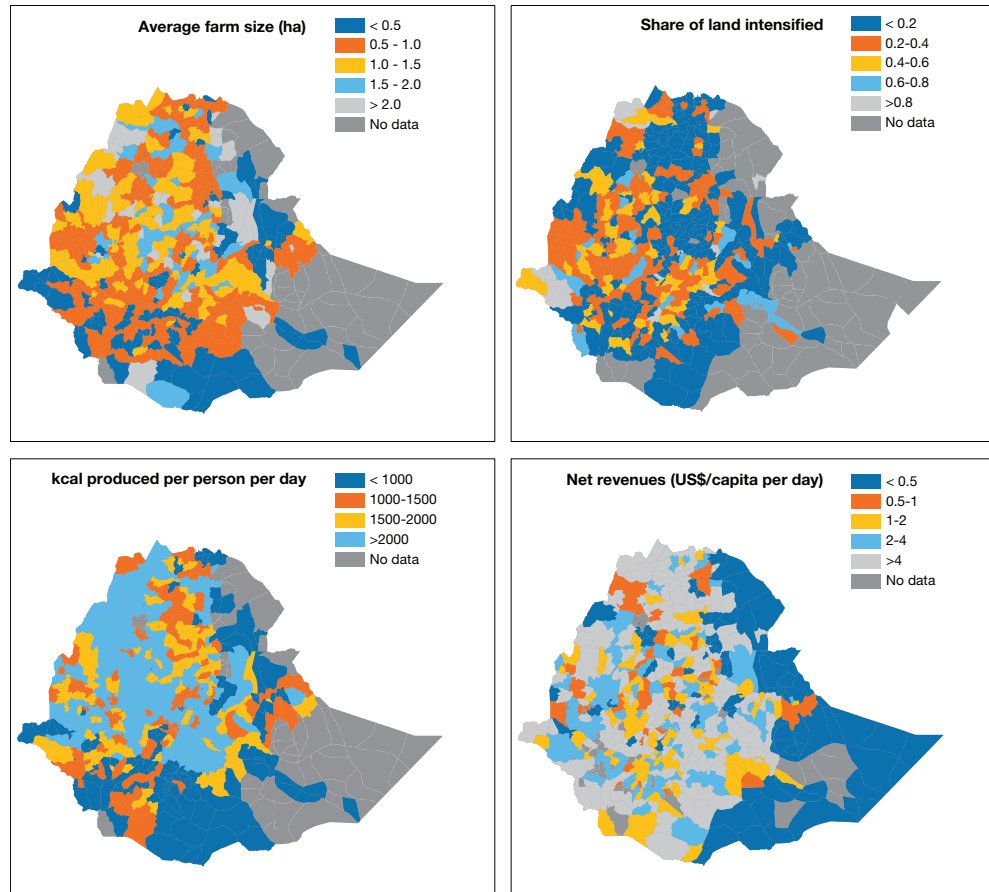
# **4** Discussion and conclusion

The objective of this study was to combine household survey, agricultural census and land cover data to analyse food security and poverty both at the micro (farm) and macro (regional and country) levels. To do this, we developed a novel methodology that combines different data sources to establish country-wide farm typologies that enable an analysis of poverty and food security that is both farm-system specific and spatially explicit. To test our methodology, we analysed the poverty and food security situation of Ethiopian smallholder farms.

## 4.1 Discussion

In the existing literature, farming systems and poverty and food security analyses often use either regional or national statistics in combination with expert knowledge or household-level data and statistical clustering techniques. Expert knowledge and macro-level statistics map farm structures and analyse poverty and food security across large regions and multiple countries (Cassman et al., 2005; Dixon, Gulliver and Gibbon, 2001; Seré and Steinfeld, 1996). However, this does not record heterogeneity within the coarse administrative units and therefore cannot be parameterized with data. A typology that is constructed using farm household data enables the construction and quantification of farm-type-specific levels of poverty and food security (Cecchi et al., 2010; Jayne et al., 2003; Lowder, Skoet and Raney, 2016; Otte and Chilonda, 2002; Oumer, Hjortsø and de Neergaard, 2013). However, this has limited coverage and representativeness. In this study, we combined the two approaches, using several datasets and statistical methods to capture spatial heterogeneity in land cover and land use at both the small- and large-scale levels, as well as to capture the heterogeneity of the actors operating the land.

The resulting typology reflected the importance of geographical factors influencing farmers' livelihood options. The arid-pastoral zone had mainly pure livestock and millet-sorghum farms, whereas the drought-prone and moisture-reliable zones had different mixed farming systems, but cereal-based farming systems dominated there. However, characterizing farming systems goes beyond the AEZ. Factors such as access to markets, population density, main production activities, and degree of off-farm work have been used by others to further understand heterogeneity between smallholder farms (Chamberlin, Pender and Bingxin, 2006; Oumer, Hjortø and de Neergaard, 2013; Tittonell, 2014). However, none of these studies has been able to distinguish these factors among farming systems. By spatially differentiating within the defined farming systems, this study moves beyond an entry point for development interventions targeting reduced poverty and increased food security towards a way to quantify the impact of different types of farming systems in different areas.

Regarding poverty, our results show that, even when farmers would sell all their produce, total revenues would be below the US$1.25 per day line for all farming systems. Moreover, the percentage of products sold and the share of off-farm work in the total value of production was very low. The highest degree of market interaction (sale of agricultural products and off-farm work combined) was found in the moisture-reliable zone, especially for the maize-based farming system. The pure livestock farming system and the arid-pastoral zone were characterized by especially low market interaction, which is reflected by the low average revenues in Afar Region and Somali.

Regarding food security, only millet-sorghum in the arid-pastoral zone, wheat-barley-teff in the moisture-reliable zone and both in the drought-prone zone produced at or around the minimum daily calorie intake. All farming systems needed to buy additional products from the market to secure a healthy diet and, currently, revenues obtained from market interaction are insufficient. For the pure livestock dominant activity and the perennial activity in the moisture-reliable zone, the food basket needs to be complemented with cereals. The relatively better food security status of the millet-sorghum and wheat-barley-teff activity sets is reflected by the higher per capita calorie availability in Tigray, Amhara and northern Oromia.

The farm-system- and location-specific poverty and food security indicators provide a first direction on whether the transformation of smallholder agriculture should be achieved via increased market orientation or increased productivity. While some studies indicate that large parts of the yield gap can be closed without the need for new technologies (Henderson et al., 2016), others highlight the importance of better access to markets (Frelat et al., 2016; Hanjra, Ferede and Gutta, 2009; Jayne et al., 2003; Sibhatu, Krishna and Qaim, 2015). Our results show that better market access may especially benefit the millet-sorghum, wheat-barley-teff and perennial activities and the moisture-reliable zone, where yields and production diversity are relatively high. Improved productivity and increased diversity may especially benefit the maize and pure-livestock systems and the arid-pastoral zone, which are characterized by low food availability and high costs to complement the food basket. Well-targeted policies that account for the diversity among smallholder farms and between regions is therefore key to improving poverty and food security.

## 4.2 Conclusion

We developed a novel methodology integrating household survey data with land cover data and subnational statistics to establish country-wide farm typologies that enable farm-system-specific analysis. To target policies aimed to counter poverty and improve food security, it is imperative to properly reflect what defines farmers' production choices at what locations. The typology in this paper focused on the AEZ, farm size and main activity set, enabling us to account for the diversity of smallholder farming systems while also taking large-scale commercial farms into account. To quantify and spatially differentiate the farming systems, farm household survey data were extrapolated using a multinomial regression and harmonized with land cover data and subnational statistics. Compared with other typologies that commonly only focus on the spatial distribution of farming systems, this methodology fills the resulting typology with data, enabling analysis of income and food security of different farming systems.

This method results in farming systems parameterized in a way that allows for transitions from one type to another. Therefore, such country-wide farming systems typology can be extremely useful for ex ante impact analyses as well as to better target and tailor development interventions to specific farming systems, focusing governmental efforts towards poverty reduction or, more broadly, towards meeting the SDGs. Additional and more dynamic elements to characterize farmers' poverty and food security, such as market access and the degree of intensification, can be incorporated into the existing methodological framework. This could provide a framework to analyse trade-offs and complementarities between targeting poverty and food security through improving market access and closing yield gaps. As such, the methodology developed can be used in farm household and large-scale market equilibrium models, which often require data on all household activities with a wide spatial coverage. Therefore, the method developed in this study is replicable for other countries and settings, as long as sufficient household survey data and subnational statistics are available.

# References

Antle, J., Adhikari, R. and Price, S. 2015. An income-based food security indicator for agricultural technology impact assessment. *Frontiers of Economics and Globalization* 15: 81-96. https://doi.org/10.1108/S1574-871520150000015006

Bryan, E., Deressa, T.T., Gbetibouo, G.A. and Ringler, C. 2009. Adaptation to climate change in Ethiopia and South Africa: Options and constraints. *Environmental Science and Policy* 12(4): 413-426. https://doi.org/10.1016/j.envsci.2008.11.002

Cameron, A.C. and Trivedi, P.K. 2005. *Microeconometrics – Methods and applications*. Cambridge: Cambridge University Press. https://doi.org/10.1007/s13398-014-0173-7.2

Cassman, K.G. and Wood, S. et al. 2005. Cultivated systems. In *Ecosystems and human well-being: Current state and trends*. Millennium Ecosystem Assessment, World Resources Institute and Island Press. www.millenniumassessment.org/documents/document.295.aspx.pdf

Cecchi, G., Wint, W., Shaw, A., Marletta, A., Mattioli, R. and Robinson, T. 2010. Geographic distribution and environmental characterization of livestock production systems in Eastern Africa. *Agriculture, Ecosystems and Environment* 135(1-2): 98-110. https://doi.org/10.1016/j.agee.2009.08.011

Chamberlin, J., Pender, J. and Bingxin, Y. 2006. Development domains for Ethiopia: Capturing the geographical context of smallholder development options. Development Strategy and Governance Division Discussion Paper No. 43. Washington, D.C.: International Food Policy Research Institute.

CSA. 2012. Ethiopian Agricultural Sample Enumeration 2011/2012. Central Agricultural Census Commission, Statistical Report, Part I and II. Addis Ababa: Central Statistical Agency.

Demissie, A. and Legesse, B. 2013. Determinants of income diversification among rural households: The case of smallholder farmers in Fedis district, Eastern Hararghe zone, Ethiopia. *Journal of Development and Agricultural Economics* 5(March), 120-128. https://doi.org/10.5897/JDAE12.104

Diao, X., Hazell, P. and Thurlow, J. 2010. The role of agriculture in African development. *World Development* 38(10): 1375-1383. https://doi.org/10.1016/j.worlddev.2009.06.011

Dixon, J., Gulliver, A. and Gibbon, D. 2001. *Farming systems and poverty: Improving farmers' livelihoods in a changing world*. Rome: Food and Agriculture Organization of the United Nations; Washington, D.C.: World Bank.

Dorosh, P.A. and Rashid, S. 2012. *Food and agriculture in Ethiopia*. Philadephia: University of Pennsylvania Press. https://doi.org/10.9783/9780812208610

EPHI. 2013. *Ethiopian national food consumption survey*. Addis Ababa: Ethiopian Public Health Institute.

FAO. 2016. FAOSTAT. Retrieved 9 August 2016 from http://faostat3.fao.org/browse/R/RL/E

FAO. 2017. Food Balance Sheets. Retreived 29 June 2018 from www.fao.org/faostat/en/#data/FBS

Frelat, R., Lopez-Ridaura, S., Giller, K.E., Herrero, M., Douxchamps, S., Andersson Djurfeldt, A., Erenstein, O., Henderson, B., Kassie, M., Paul, B.K., Rigolot, C., Ritzema, R.S., Rodriguez, D., van Asten, P.J.A. and van Wijk, M.T. 2016. Drivers of household food availability in sub-Saharan Africa based on big data from small farms. PNAS, 113(2): 458-463. https://doi.org/10.1073/pnas.1518384112

Frühwirt-Schnatter, S. and Frühwirt, R. 2012. Bayesian inference in the multinomial Logit model. *Austrian Journal of Statistics* 41(1): 27-43.

Garnett, T., Appleby, M.C., Balmford, A., Bateman, I.J., Benton, T.G., Bloomer, P., Burlingame, B., Dawkins, M., Dolan, L., Fraser, D., Herrero, M., Hoffmann, I., Smith, P., Thornton, P.K., Toulmin, C., Vermeulen, S.J. and Godfray, H.C.J. 2013. Sustainable intensification in agriculture: Premises and policies. *Science* 341(6141): 33-34. https://doi.org/10.1126/science.1234485

George, E.I. and McCulloch, R.E. 1993. Variable selection via Gibbs sampling. *Journal of the American Statistical Association* 88(423): 881-889. https://doi.org/10.1080/01621459.1993.10476353

Graeub, B.E., Chappell, M.J., Wittman, H., Ledermann, S., Kerr, R.B. and Gemmill-Herren, B. 2015. The state of family farms in the world. *World Development* 87: 1-15. https://doi.org/10.1016/j.worlddev.2015.05.012

Hanjra, M.A., Ferede, T. and Gutta, D.G. 2009. Pathways to breaking the poverty trap in Ethiopia: Investments in agricultural water, education, and markets. *Agricultural Water Management* 96(11): 1596-1604. https://doi.org/10.1016/j.agwat.2009.06.008

Henderson, B., Godde, C., Medina-Hidalgo, D., van Wijk, M., Silvestri, S., Douxchamps, S., Stephenson, E., Power, B., Rigolot, C., Cacho, O. and Herrero, M. 2016. Closing system-wide yield gaps to increase food production and mitigate GHGs among mixed crop-livestock smallholders in sub-Saharan Africa. *Agricultural Systems* 143: 106-113. https://doi.org/10.1016/j.agsy.2015.12.006

Herrero, M ., Notenbaert, A., Thornton, P., Pfeifer, C., Silvestri, S., Omolo, A., Quiros, C. 2014. A framework for targeting and scaling-out interventions in agricultural systems. CCAFS Working Paper No. 62. Copenhagen: CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS). https://ccafs.cgiar.org/publications/framework-targeting-and-scaling-out-interventions-agricultural-systems

Holmes, C.C. and Held, L. 2006. Bayesian auxiliary variable models for binary and multinomial regression. *Bayesian Analysis* 1(1): 145-168. https://doi.org/10.1214/06-BA105

IES-JRC. 2000. Global Land Cover 2000 Project (GLC 2000). http://forobs.jrc.ec.europa.eu/products/glc2000/glc2000.php

Jayne, T.S., Yamano, T., Weber, M.T., Tschirley, D., Benfica, R., Chapoto, A. and Zulu, B. 2003. Smallholder income and land distribution in Africa: Implications for poverty reduction strategies. *Food Policy* 28(3): 253-275. https://doi.org/10.1016/S0306-9192(03)00046-0

Josephson, A.L., Ricker-Gilbert, J. and Florax, R.J.G.M. 2014. How does population density influence agricultural intensification and productivity? Evidence from Ethiopia. *Food Policy* 48: 142-152. https://doi.org/10.1016/j.foodpol.2014.03.004

Keola, S., Andersson, M. and Hall, O. 2015. Monitoring economic development from space: Using nighttime light and land cover data to measure economic growth. *World Development* 66: 322-334. https://doi.org/10.1016/j.worlddev.2014.08.017

Kuo, L. and Mallick, B. 1998. Variable selection for regression models. *Sankhyā Indian Journal of Statistics* 60: 65-81.

LeSage, J., Pace, R.K. 2009. *Introduction to spatial econometrics*. Boca Raton: Chapman & Hall/CRC.

Lowder, S.K., Skoet, J. and Raney, T. 2016. The number, size, and distribution of farms, smallholder farms, and family farms worldwide. *World Development* 87: 16-29. https://doi.org/10.1016/j.worlddev.2015.10.041

Mertens, B., Sunderlin, W.D., Ndoye, O. and Lambin, E.F. 2000. Impact of macroeconomic change on deforestation in south Cameroon: Integration of household survey and remotely-sensed data. *World Development* 28(6): 983-999. https://doi.org/10.1016/S0305-750X(00)00007-3

Otte, M.J. and Chilonda, P. 2002. Cattle and small ruminant production systems in sub-Saharan Africa: A systematic review. Rome: Food and Agriculture Organization of the United Nations. www.fao.org/ag/againfo/resources/en/publications/agapubs/AGAL-Y4176E.pdf

Oumer, A.M., Hjortsø, C.N. and de Neergaard, A. 2013. Understanding the relationship between livelihood strategy and soil management: Empirical insights from the central highlands of Ethiopia. *Food Security* 5(2): 143-156. https://doi.org/10.1007/s12571-013-0242-3

Paul, C.J., Weinthal, E.S., Bellemare, M.F. and Jeuland, M.A. 2016. Social capital, trust, and adaptation to climate change: Evidence from rural Ethiopia. *Global Environmental Change* 36: 124-138. https://doi.org/10.1016/j.gloenvcha.2015.12.003

Polson, N., Scott, J. and Windle, J. 2013. Bayesian inference for logistic models using Pólya–Gamma latent variables. *Journal of the American Statistical Association* 108(504): 1339-1349.

Rapsomanikis, G. 2015. *The economic lives of smallholder farmers: An analysis based on household data from nine countries*. Rome: Food and Agriculture Organization of the United Nations. www.fao.org/3/a-i5251e.pdf

Sachs, J.D. 2012. From millennium development goals to sustainable development goals. *The Lancet* 379(9832): 2206-2211. https://doi.org/10.1016/S0140-6736(12)60685-0

Seré, C. and Steinfeld, H. 1996. World livestock production systems: Current status, issues and trends. *FAO Animal Production and Health Paper* 127. Rome: Food and Agriculture Organization of the United Nations.

Sibhatu, K.T., Krishna, V.V. and Qaim, M. 2015. Production diversity and dietary diversity in smallholder farm households. *Proceedings of the National Academy of Sciences of the United States of America* 112(34): 10657-10662. https://doi.org/10.1073/pnas.1510982112

Tadese, M., Alemu, B., Bekele, G., Tebikew, T., Chamberlin, J. and Benson, T. 2006. *Atlas of the Ethiopian rural economy*. Washington, D.C.: International Food Policy Research Institute; Addis Ababa: Central Statistical Agency, Ethiopian Development Research Institute. https://doi.org/10.2499/0896291545

Tittonell, P. 2014. Livelihood strategies, resilience and transformability in African agroecosystems. *Agricultural Systems* 126: 3-14. https://doi.org/10.1016/j.agsy.2013.10.010

van Wesenbeeck, C.F., Keyzer, M.A. and Nubé, M. 2009. Estimation of undernutrition and mean calorie intake in Africa: Methodology, findings and implications. *International Journal of Health Geographics* 8: 37. https://doi.org/10.1186/1476-072X-8-37

Watmough, G.R., Atkinson, P.M., Saikia, A. and Hutton, C.W. 2016. Understanding the evidence base for poverty–environment relationships using remotely sensed satellite data: An example from Assam, India. *World Development* 78: 188-203. https://doi.org/10.1016/j.worlddev.2015.10.031

Wiggins, S., Kirsten, J., Llambí, L. 2010. The future of small farms. *World Development* 38(10): 1341-1348. https://doi.org/10.1016/j.worlddev.2009.06.013

World Bank. 2016. About LSMS. Retrieved 3 October 2016 from http://surveys.worldbank.org/lsms/about-lsms

# Annex

**Table A1** Descriptive statistics of the explanatory variables used in the regression

| Variable | Min | Max | Mean | Std dev. |
|---|---|---|---|---|
| Population density (people/km2) | 1.00 | 77 920.00 | 2 109.00 | 8 199.84 |
| Average elevation (masl) | 175.00 | 3 040.00 | 1 617.94 | 699.98 |
| Average slope (%) | 0.04 | 16.16 | 5.30 | 3.25 |
| Average rain (mm/month) | 0.00 | 9.00 | 3.78 | 2.20 |
| Road density (m/km2) | 0.00 | 1 696.52 | 54.34 | 172.85 |
| Coffee Marketing Co-op | 0.00 | 1.00 | 0.04 | 0.21 |
| Multi-purpose Co-op | 0.00 | 1.00 | 0.06 | 0.25 |
| No co-op | 0.00 | 1.00 | 0.85 | 0.36 |
| Number of cooperatives | 0.00 | 74.00 | 1.98 | 6.76 |
| Total members involved | 0.00 | 47 249.00 | 1 808.67 | 6 494.12 |
| Barley dominant | 0.00 | 1.00 | 0.08 | 0.27 |
| Maize dominant | 0.00 | 1.00 | 0.25 | 0.43 |
| Millet dominant | 0.00 | 1.00 | 0.01 | 0.12 |
| Pastoral dominant | 0.00 | 1.00 | 0.14 | 0.35 |
| Sorghum dominant | 0.00 | 1.00 | 0.13 | 0.34 |
| Teff dominant | 0.00 | 1.00 | 0.18 | 0.38 |
| Wheat dominant | 0.00 | 1.00 | 0.06 | 0.25 |
| Barley secondary | 0.00 | 1.00 | 0.10 | 0.29 |
| Maize secondary | 0.00 | 1.00 | 0.20 | 0.40 |
| Millet secondary | 0.00 | 1.00 | 0.02 | 0.15 |
| Pastoral secondary | 0.00 | 1.00 | 0.15 | 0.35 |
| Sorghum secondary | 0.00 | 1.00 | 0.15 | 0.35 |
| Teff secondary | 0.00 | 1.00 | 0.15 | 0.36 |
| Wheat secondary | 0.00 | 1.00 | 0.10 | 0.30 |
| Temperate agroecological zone | 0.00 | 1.00 | 0.46 | 0.50 |
| Humid agroecological zone | 0.00 | 1.00 | 0.16 | 0.37 |
| Area (km2) | 0.15 | 2 228.47 | 233.12 | 306.43 |
| Cropland (km2) | 0.00 | 77.16 | 12.51 | 15.05 |
| Other agricultural land (km2) | 0.00 | 16.11 | 2.13 | 3.18 |
| Grassland (km2) | 0.00 | 1 074.84 | 77.83 | 157.59 |
| Forest (km2) | 0.00 | 305.71 | 9.70 | 30.80 |
| Wetland (km2) | 0.00 | 108.09 | 2.19 | 9.43 |
| Other natural land (km2) | 0.00 | 1 341.35 | 89.18 | 135.96 |
| Other land (km2) | 0.00 | 1 334.52 | 39.56 | 144.38 |

Std dev., standard deviation.

**Table A2** Estimation results of the multinomial Logit model

| Variable | Livestock | | | Millet-Sorghum | | | |
|---|---|---|---|---|---|---|---|
| | Post. mean | Incl. prob. | Std dev. | Post. mean | Incl. prob. | Std dev. | |
| (Intercept) | 0.55 | 1.00 | 0.27 | 0.54 | 1.00 | 0.26 | |
| Population density | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | |
| Average elevation | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | |
| Average slope | 0.02 | 0.52 | 0.01 | -0.01 | 0.01 | 0.01 | |
| Average rain | -0.03 | 0.11 | 0.02 | 0.00 | 0.01 | 0.02 | |
| Road density | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | |
| Coffe Marketing Co-op | -0.02 | 0.01 | 0.04 | -0.01 | 0.00 | 0.03 | |
| Multi-purpose Co-op | 0.42 | 0.90 | 0.22 | 0.02 | 0.59 | 0.14 | |
| No co-op | 0.03 | 0.01 | 0.05 | 0.03 | 0.00 | 0.04 | |
| Number of cooperatives | -0.04 | 0.43 | 0.03 | 0.00 | 0.03 | 0.01 | |
| Total members involved | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | |
| Barley dominant | 0.01 | 0.00 | 0.07 | -0.15 | 0.02 | 0.08 | |
| Maize dominant | -0.18 | 0.10 | 0.08 | 0.19 | 0.00 | 0.06 | |
| Millet dominant | 0.03 | 0.00 | 0.22 | 0.04 | 0.00 | 0.22 | |
| Pastoral dominant | -0.17 | 0.02 | 1.07 | 0.03 | 0.00 | 0.50 | |
| Sorghum dominant | 0.24 | 0.18 | 0.07 | -0.06 | 0.00 | 0.06 | |
| Teff dominant | -0.08 | 0.00 | 0.06 | 0.02 | 0.00 | 0.06 | |
| Wheat dominant | 2.22 | 0.00 | 1.37 | -9.33 | 0.89 | 5.70 | |
| Barley secondary | -0.10 | 0.01 | 0.06 | -0.02 | 0.00 | 0.06 | |
| Maize secondary | -0.02 | 0.00 | 0.05 | 0.12 | 0.01 | 0.05 | |
| Millet secondary | 0.03 | 0.00 | 0.07 | 0.08 | 0.00 | 0.07 | |
| Pastoral secondary | -9.91 | 1.00 | 5.61 | 3.03 | 0.01 | 1.68 | |
| Sorghum secondary | 0.02 | 0.00 | 0.05 | -0.17 | 0.01 | 0.05 | |
| Teff secondary | 0.04 | 0.00 | 0.05 | 0.06 | 0.00 | 0.05 | |
| Wheat secondary | -0.08 | 0.00 | 0.07 | 0.03 | 0.00 | 0.06 | |
| Temperate agroecological zone | 0.01 | 0.01 | 0.03 | 0.03 | 0.09 | 0.05 | |
| Humid agroecological zone | 2.37 | 0.01 | 0.93 | 2.46 | 0.00 | 0.97 | |

| | Maize | | | Wheat-Barley-Teff | | | Perennial | | |
|---|---|---|---|---|---|---|---|---|---|
| | Post. mean | Incl. prob. | Std dev. | Post. mean | Incl. prob. | Std dev. | Post. mean | Incl. prob. | Std dev. |
| | -0.65 | 1.00 | 0.45 | -0.66 | 1.00 | 0.32 | 0.24 | 1.00 | 0.19 |
| | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | 0.00 | 0.10 | 0.01 | 0.00 | 0.03 | 0.00 | -0.01 | 0.52 | 0.01 |
| | 0.00 | 0.03 | 0.03 | 0.03 | 0.34 | 0.04 | -0.01 | 0.34 | 0.01 |
| | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | 0.00 | 0.02 | 0.05 | 0.03 | 0.17 | 0.06 | -0.01 | 0.17 | 0.02 |
| | 0.24 | 0.85 | 0.17 | -0.21 | 0.03 | 0.11 | -0.23 | 0.90 | 0.14 |
| | 0.06 | 0.01 | 0.06 | -0.09 | 0.48 | 0.09 | 0.03 | 0.48 | 0.03 |
| | 0.00 | 0.05 | 0.02 | 0.02 | 0.15 | 0.02 | 0.01 | 0.43 | 0.01 |
| | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | -0.01 | 0.00 | 0.10 | 0.09 | 0.01 | 0.05 | 0.00 | 0.02 | 0.03 |
| | -0.12 | 0.00 | 0.10 | 0.00 | 0.00 | 0.05 | 0.05 | 0.10 | 0.03 |
| | 0.32 | 0.02 | 0.35 | -0.20 | 0.05 | 0.57 | 0.02 | 0.05 | 0.17 |
| | 0.35 | 0.04 | 2.43 | -0.04 | 0.00 | 0.33 | -0.03 | 0.04 | 0.37 |
| | -0.20 | 0.00 | 0.11 | 0.00 | 0.00 | 0.05 | -0.05 | 0.18 | 0.03 |
| | 0.11 | 0.00 | 0.09 | 0.00 | 0.00 | 0.05 | 0.00 | 0.00 | 0.03 |
| | 3.53 | 0.00 | 2.14 | 1.55 | 0.00 | 0.95 | 1.81 | 0.89 | 1.10 |
| | 0.04 | 0.00 | 0.09 | 0.04 | 0.00 | 0.05 | 0.02 | 0.01 | 0.03 |
| | -0.06 | 0.00 | 0.08 | 0.02 | 0.00 | 0.05 | -0.07 | 0.01 | 0.03 |
| | -0.16 | 0.01 | 0.12 | -0.02 | 0.00 | 0.09 | 0.01 | 0.01 | 0.04 |
| | 1.30 | 0.34 | 5.17 | 2.13 | 0.00 | 1.12 | 2.51 | 1.00 | 1.31 |
| | -0.09 | 0.00 | 0.08 | 0.12 | 0.01 | 0.04 | 0.02 | 0.01 | 0.03 |
| | 0.01 | 0.00 | 0.08 | -0.04 | 0.00 | 0.04 | -0.03 | 0.00 | 0.03 |
| | 0.27 | 0.01 | 0.09 | -0.10 | 0.02 | 0.05 | 0.01 | 0.02 | 0.03 |
| | -0.02 | 0.07 | 0.06 | -0.02 | 0.10 | 0.05 | 0.00 | 0.10 | 0.02 |
| | -13.43 | 1.00 | 5.73 | 1.02 | 0.88 | 0.66 | 1.90 | 1.00 | 0.74 |

**Table A2 cont.d**

| Variable | Livestock | | | Millet-Sorghum | | | |
|---|---|---|---|---|---|---|---|
| | Post. mean | Incl. prob. | Std dev. | Post. mean | Incl. prob. | Std dev. | |
| Area | 0.00 | 0.00 | 0.04 | 0.00 | 0.00 | 0.04 | |
| Cropland | 0.00 | 0.00 | 0.04 | 0.01 | 0.00 | 0.04 | |
| Other agricultural land | 0.00 | 0.00 | 0.04 | -0.01 | 0.00 | 0.04 | |
| Grassland | 0.00 | 0.00 | 0.04 | 0.00 | 0.00 | 0.04 | |
| Forest | 0.00 | 0.00 | 0.04 | 0.00 | 0.00 | 0.04 | |
| Wetland | 0.00 | 0.00 | 0.04 | 0.00 | 0.00 | 0.04 | |
| Other natural land | 0.00 | 0.00 | 0.04 | 0.00 | 0.00 | 0.04 | |
| Other land | 0.00 | 0.00 | 0.04 | 0.00 | 0.00 | 0.04 | |
| W-Population density | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | |
| W-Average elevation | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | |
| W-Average slope | 0.00 | 0.01 | 0.01 | 0.00 | 0.00 | 0.01 | |
| W-Average rain | -0.06 | 0.06 | 0.03 | -0.04 | 0.00 | 0.03 | |
| W-Road density | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | |
| W-Area | 0.01 | 0.00 | 0.04 | -0.01 | 0.00 | 0.04 | |
| W-Cropland | -0.01 | 0.00 | 0.04 | 0.02 | 0.00 | 0.04 | |
| W-Other agricultural land | 0.05 | 0.00 | 0.04 | -0.06 | 0.00 | 0.04 | |
| W-Grassland | -0.01 | 0.00 | 0.04 | 0.01 | 0.00 | 0.04 | |
| W-Forest | 0.00 | 0.00 | 0.04 | 0.00 | 0.00 | 0.04 | |
| W-Wetland | -0.01 | 0.00 | 0.04 | 0.00 | 0.00 | 0.04 | |
| W-Other natural land | -0.01 | 0.00 | 0.04 | 0.01 | 0.00 | 0.04 | |
| W-Other land | -0.01 | 0.00 | 0.04 | 0.01 | 0.00 | 0.04 | |
| Residual deviance | 3 425 | | | | | | |
| Log-likelihood | $8\times10^{6}$ | | | | | | |
| BIC | 1.256 | | | | | | |

BIC, Baysian information criterion; incl., inclusion; post., posterior; prob., probability; std dev., standard deviation; W, spatial weight.

| | Maize | | | Wheat-Barley-Teff | | | Perennial | | |
|---|---|---|---|---|---|---|---|---|---|
| | Post. mean | Incl. prob. | Std dev. | Post. mean | Incl. prob. | Std dev. | Post. mean | Incl. prob. | Std dev. |
| | 0.00 | 0.00 | 0.06 | 0.00 | 0.00 | 0.02 | 0.00 | 0.00 | 0.01 |
| | 0.01 | 0.00 | 0.06 | -0.01 | 0.00 | 0.03 | 0.00 | 0.00 | 0.01 |
| | -0.02 | 0.00 | 0.06 | 0.00 | 0.00 | 0.03 | 0.01 | 0.00 | 0.02 |
| | 0.00 | 0.00 | 0.06 | 0.00 | 0.00 | 0.02 | 0.00 | 0.00 | 0.01 |
| | 0.00 | 0.00 | 0.06 | 0.00 | 0.00 | 0.02 | 0.00 | 0.00 | 0.01 |
| | -0.01 | 0.00 | 0.06 | 0.00 | 0.00 | 0.02 | 0.00 | 0.00 | 0.01 |
| | 0.00 | 0.00 | 0.06 | 0.00 | 0.00 | 0.02 | 0.00 | 0.00 | 0.01 |
| | 0.01 | 0.00 | 0.06 | 0.00 | 0.00 | 0.02 | 0.00 | 0.00 | 0.01 |
| | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | 0.00 | 0.00 | 0.01 | 0.00 | 0.00 | 0.01 | 0.00 | 0.01 | 0.00 |
| | 0.11 | 0.89 | 0.05 | 0.05 | 0.63 | 0.05 | -0.04 | 0.89 | 0.02 |
| | 0.00 | 0.00 | 0.00 | 0.01 | 0.22 | 0.00 | 0.00 | 0.22 | 0.00 |
| | -0.01 | 0.00 | 0.06 | 0.00 | 0.00 | 0.03 | 0.00 | 0.00 | 0.01 |
| | 0.04 | 0.00 | 0.06 | -0.02 | 0.00 | 0.03 | -0.01 | 0.00 | 0.02 |
| | -0.13 | 0.00 | 0.07 | 0.04 | 0.00 | 0.04 | 0.03 | 0.00 | 0.02 |
| | 0.01 | 0.00 | 0.06 | 0.00 | 0.00 | 0.03 | 0.00 | 0.00 | 0.01 |
| | 0.01 | 0.00 | 0.06 | 0.00 | 0.00 | 0.03 | -0.01 | 0.00 | 0.01 |
| | 0.03 | 0.00 | 0.06 | -0.01 | 0.00 | 0.03 | 0.00 | 0.00 | 0.01 |
| | 0.01 | 0.00 | 0.06 | 0.00 | 0.00 | 0.03 | 0.00 | 0.00 | 0.01 |
| | 0.01 | 0.00 | 0.06 | 0.00 | 0.00 | 0.03 | 0.00 | 0.00 | 0.01 |

**Figure A1** Extrapolated smallholder activity per *woreda*
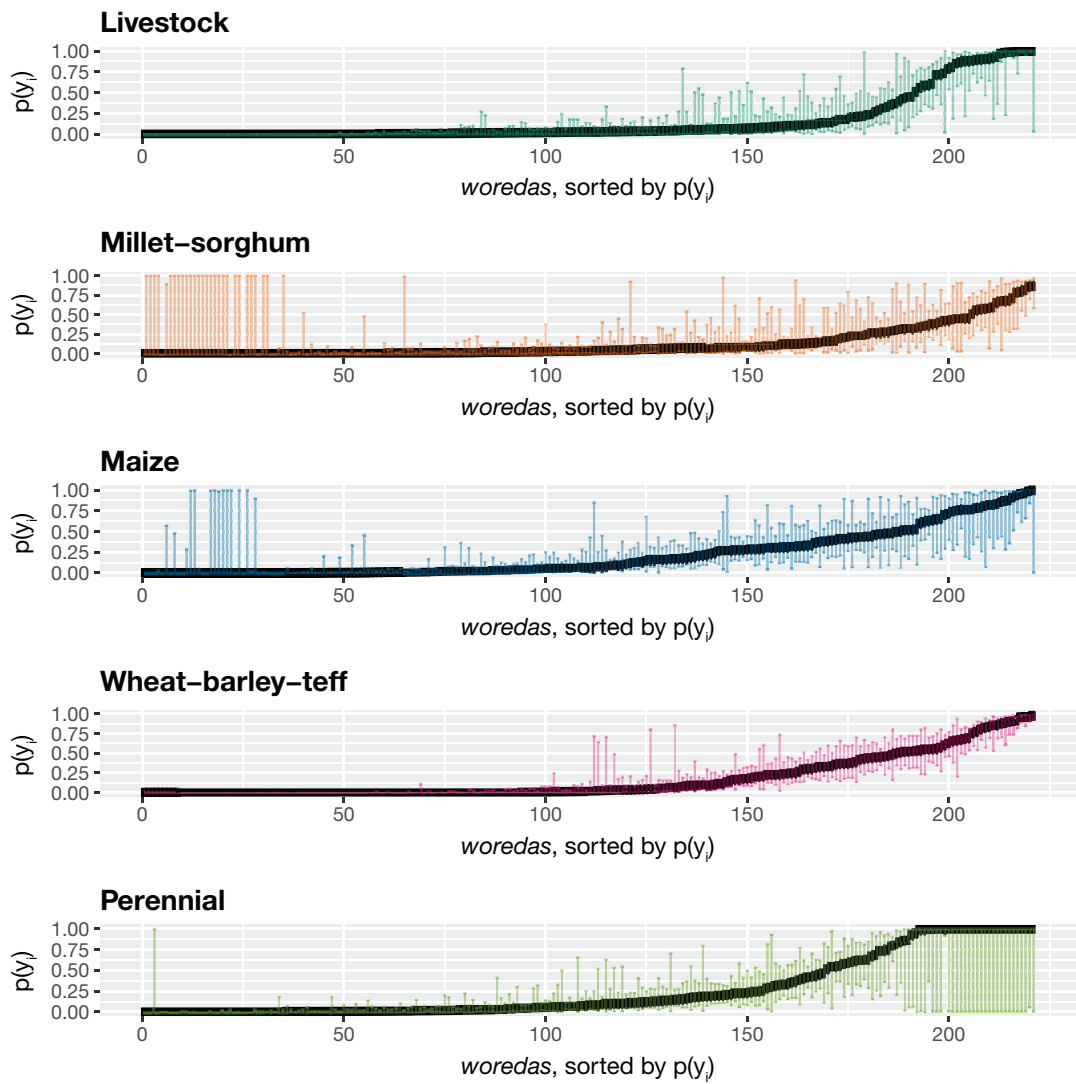


Points represent posterior median, while the error bars denote the lower 5th and the upper 95th quantiles

**The IFAD Research Series**

01.  Agricultural and rural development reconsidered
     A guide to issues and debates
     By Steve Wiggins

02.  Migration and transformative pathways
     A rural perspective
     By David Suttie, Rosemary Vargas-Lundius

03.  Fostering inclusive outcomes in sub-Saharan African agriculture
     Improving agricultural productivity and expanding agribusiness opportunities
     By David Suttie, Rui Benfica

04.  The effects of smallholder agricultural involvement on household food consumption and
     dietary diversity
     Evidence from Malawi
     By Rui Benfica, Talip Kilic

05.  Rural-urban linkages and food systems in sub-Saharan Africa
     The rural dimension
     By Karim Hussein, David Suttie

06.  Why food and nutrition security matters for inclusive structural and rural transformation
     By Steven Were Omamo

07.  Measuring IFAD's impact
     Background paper to the IFAD9 Impact Assessment Initiative
     By Alessandra Garbero

08.  Fostering inclusive rural transformation in fragile states and situations
     By Karim Hussein

09.  Social protection and inclusive rural transformation
     By Carolina Trivelli, Silvana Vargas, Jhonatan Clausen

10.  Inclusive Finance and Inclusive Rural Transformation
     By Calum G. Turvey

11.  Food safety, trade, standards and the integration of smallholders into value chains
     A review of the literature
     By John Humphrey

12.  An evidence-based assessment of IFAD's end of project reporting
     By Bia Carneiro, Alessandra Garbero

13.  Graduation models for rural financial inclusion
     By Khalid El Harizi, Xinjia Yan

14.  Disbursement performance of the International Fund for Agricultural Development (IFAD)
     An in-depth analysis of drivers and trends
     By Tim Balint, Daniel Higgins, Paola Mallia, Silvana Scalzo, Paul Winters

15.  Remittances, growth and poverty reduction in Asia
     A critical review of the literature and new evidence from cross-country panel data
     By Katsushi S. Imai, Bilal Malaeb, Fabrizio Bresciani

16.  Getting the most out of impact evaluation for learning, reporting and influence
     Insights from piloting a Participatory Impact Assessment and Learning Approach (PIALA) with IFAD
     By Edward Heinemann, Adinda Van Hemelrijck, Irene Guijt

17.  Population age structure and sex composition in sub-Saharan Africa
     A rural-urban perspective
     By Ashira Menashe-Oren, Guy Stecklov