Agricultural Systems 153 (2017) 32–42

Contents lists available at ScienceDirect

Agricultural Systems

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

To mulch or to munch? Big modelling of big data

D Rodriguez [a,*], P de Voil [a], MC Rufino [b], M Odendo [c], MT van Wijk [d]

[a] Queensland Alliance for Agriculture and Food Innovation (QAAFI), The University of Queensland, PO Box 102, Toowoomba, Qld 4350, Australia
[b] Lancaster Environment Centre, Lancaster University, UK
[c] Kenya Agricultural and Livestock Research Organization, (KALRO), Kakamega, Kenya
[d] ILRI, Costa Rica

ARTICLE INFO

Article history:
Received 29 June 2016
Received in revised form 17 January 2017
Accepted 17 January 2017
Available online xxxx

Keywords:
APSIM
Whole farm modelling
Integrative analyses
Farm diversity

ABSTRACT

African farmers are poorly resourced, highly diverse and aground by poverty traps making them rather impervious to change. As a consequence R4D efforts usually result in benefits but also trade-offs that constraint adoption and change. A typical case is the use of crop residues as mulches or as feedstock. Here we linked a database of household surveys with a dynamic whole farm simulation model, to quantify the diversity of trade-offs from the alternative use of crop residues. Simulating all the households in the survey ($n = 613$) over 99 years of synthetic climate data, showed that benefits and trade-offs from “mulching or munching” differ across agro-ecologies, and within agro-ecologies across typologies of households. Even though trade-offs between household production or income and environmental outcomes could be managed; the magnitude of the simulated benefits from the sustainable intensification of maize-livestock systems were small. Our modelling framework shows the benefits from the integration of socio-economic and biophysical approaches to support the design of development programs. Our results support the argument that a greater focus is required on the development and diversification of farmers' livelihoods within the framework of an improved understanding of the interconnectedness between biophysical, socio-economic and market factors.

1. Introduction

Across Sub Saharan Africa (SSA), crop residue biomass is a valuable and scarce household resource (Tittonell et al., 2015). Crop residues, containing Carbon (C) and Nitrogen (N) nutrients, are used either as livestock feed, a source of energy, building materials, source of cash, re-cycled back into the cropping system as mulches, or just burnt in the field. A key practice of conservation agriculture is the use of crop residues as mulches so that soil erosion is prevented and rainfall infiltration increased. However the appropriateness of the practice in SSA, widely adopted elsewhere, has been contentious (Derpsch et al., 2014; Giller et al., 2009), and calls for caution (Pittelkow et al., 2015) and pragmatism have been made (Giller et al., 2015; Mafongoya et al., 2016). Sources of concern relate to the availability of crop residues for mulching, the intertwined responses between crop responses across environments and time scales (Pittelkow et al., 2015), and the myriad of biophysical, market, and socio-economic conditions (Giller et al., 2009) that prevail across the region making it difficult to identify ‘one-size fits-all’ strategies.

Improving our understanding of the differences and similarities among households, in terms of constraints and opportunities for farmers to increase income and protect the soil capital, has helped better-target options among poorly resourced smallholder farmers (Giller et al., 2015; Tittonell et al., 2009a). Household surveys, visioning exercises (Tui et al., 2015), and ex-ante modelling exercises (Roxburgh and Rodriguez, 2016) have been all useful to narrow down the “basket” of options (Giller et al., 2015). Even though in general important assumptions and simplifications are needed for simulation modelling, household modelling has shown potential to quantify the more tractable benefits and trade-offs from alternative decisions, investments, farming systems designs, and intensification options in smallholder farming (Holzworth et al., 2014; Rodriguez and Sadras, 2011). Examples can be found in the evaluation of case study farms on soil nutrient and carbon dynamics (Tittonell et al., 2009b), to the quantification of interactions and synergisms between components within the farm system (van Wijk et al., 2009), such as alternative livestock diets (Rufino et al., 2009), irrigation strategies (Power et al., 2011), or farming systems designs (Rodriguez et al., 2014, 2011).

Despite the significant improvements in the understanding of poorly resourced smallholder households, a rather fundamental challenge remains: How to deal with the large variability in the population of farms and farmers? How to represent such diversity and quantify benefits and trade-offs from alternative pathways for development? The standard approach has been to develop a household typology, select a ‘typical’ or ‘representative’ farm from each of the farm typologies and

* Corresponding author.
E-mail address: d.rodriguez@uq.edu.au (D. Rodriguez).

http://dx.doi.org/10.1016/j.agsy.2017.01.010

perform analyses using the characteristics and management structure for this small set of contrasting farms (Herrero et al., 2014; Rodriguez et al., 2014; Rufino et al., 2011; Tittonell et al., 2009c). However, this approach ignores the large variability that is inherently present in the typologies (van Wijk, 2014). The problem was previously identified (van der Ploeg et al., 2009) who showed the large diversity of development pathways over time from an initial rather similar set of households. New analyses try to move away from the approach to first aggregate and then simulate, by applying modelling and intervention analyses across populations of farm households, and then explore and aggregate the results (Frelat et al., 2015). In statistical analyses it has been shown conclusively that 'first aggregation then simulation' can lead to different results from the 'first simulation then aggregation' approach in non-linear, complex systems. Here we explore this idea further by linking a large database of household survey data with a new whole farm model (APSFarm-LivSim). Interfacing the model with a database of a household survey allowed us to parameterize and simulate each of the 613 households in the survey, thereby retaining the base variations in farm characteristics and management throughout the assessment. Understanding the diversity of responses across the most vulnerable farmers' matters given the many examples of policy prescriptions and ill-informed institutionalization of technological packages across SSA (Valbuena et al., 2012). Here we propose that given the large disparity in responses i.e. benefits and trade-offs, identifying generalizable management strategies from the analysis of a few household case studies can be misleading if used to inform practice or policy at regional or national levels.

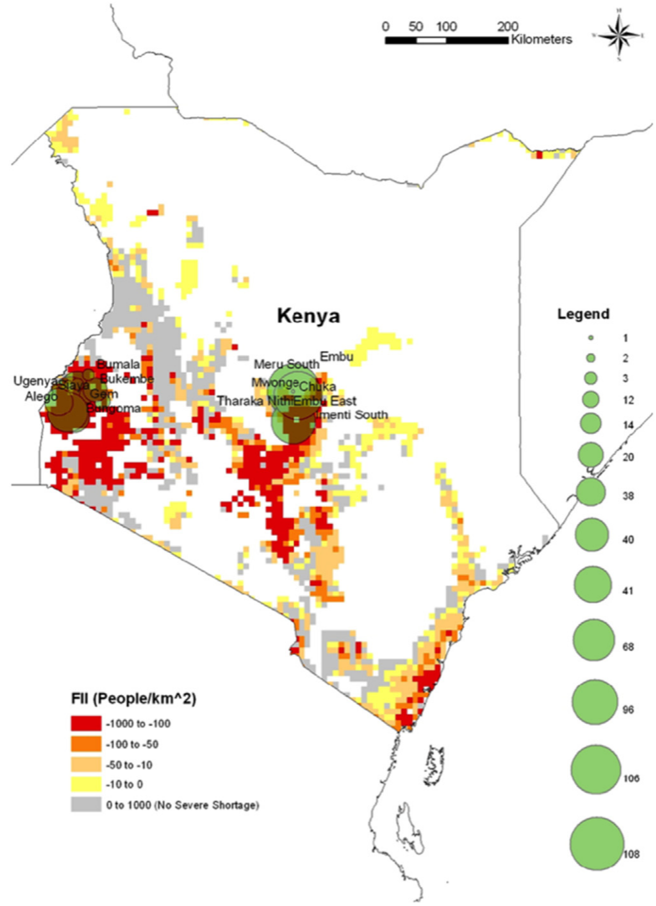

Fig. 1. Map of the distribution of the surveyed farms in Eastern and Western Kenya (n = 613), on a map showing a food insecurity index (FII, people km^{-2}) (Potgieter et al., 2013). The size of the circles indicates the number of households surveyed per village.

2. Material and methods

We used field and household level data from an extensive and homogeneous household survey (n = 613), to (i) describe the variability in household levels of endowment across Eastern and Western Kenya; and (ii) to parameterize a whole farm model (APSFarm-LivSim) that was used to quantify benefits and trade-offs in terms of changes in average ground cover, feedstock availability, heads of cattle sold, household maize production and income, and soil erosion from alternative uses of crop residues i.e. kept as mulches or fed to livestock, across all the farms in the survey. Distinctive from other studies is the dynamic coupling of whole farm models and databases of household data; and the fact that we dynamically modelled all the farms in a survey using ninety-nine years of climate records, and were able to clearly demonstrate the extent of the diversity of benefits and trade-offs across regions, and household typologies.

2.1. Baseline survey data

The survey was collected by the Sustainable Intensification of Maize-Legume Cropping Systems for Food Security in Eastern and Southern Africa (SIMLESA) program (http://aciar.gov.au/page/simlesa-program). The regions surveyed included Embu and Meru Counties in Eastern Kenya (n = 314), and Bungoma and Siaya Counties in Western Kenya (n = 299) (Fig. 1). The data was collected between January and April 2011. Survey design and data collection is described elsewhere (Frelat et al., 2015). Survey data included field and household level data. Household level data i.e. physical, financial and human capitals, was used both to describe the diversity of households by developing household structural typologies and to parameterize a dynamic whole farm model (APSFarm-LivSim, below). Briefly, factor analysis was used to extract linear combinations of the regressors that were independent (Venables and Ripley, 2000), to reduce the dimensions in the dataset. Variables showing a high correlation in the factor analysis were omitted from the cluster analysis to avoid extreme multi-co-linearity and singularity. However at a later stage, some of the variables excluded in the cluster analysis, were used to refine and help interpret the results, and to provide a more complete characterization of the household typologies. Categorical variables such as the gender of the household head, were only included in the cluster analysis. Factor analysis provides factor loadings for each variable, a measure of that variables contribution to each factor, or principal component. Variables having the largest loading values from the first most relevant principal components were examined, the first 5 (Eastern Kenya) or 9 (Western Kenya) components explained most of the variability of the total dataset. Each principal component was represented by one or two variables in the cluster analysis, and the selected variables were different between Eastern and Western Kenya. Household typologies were developed using hierarchical clustering (Ward's minimum variance linkage method) with the Euclidean distance of the normalized variables as a measure of similarity (Gong and Richman, 1995). All statistical analyses were done developing appropriate software using the R software (R Core Team, 2016).

2.2. The APSFarm-LivSim model

A combination of household and field level data collected in the survey was used to parameterize the model APSFarm-LivSim, for each farm in the database (Fig. 2). The whole farm model (APSFarm-LivSim) was derived from linking the APSFarm (Rodriguez et al., 2011 and Rodriguez et al., 2014) and LivSim (Rufino et al., 2009) models. Merging APSFarm and LivSim involved linking both mechanical and conceptual components of two distinct modelling frameworks. The simulation framework APSIM (Holzworth et al., 2014), the underlying engine of the APSFarm model, passes encoded messages between components that represent events in the system such as the transfer of resources between modules (e.g. water uptake by plants), the operation of farm level