Dynamic Agricultural Household Bio-Economic Simulator (DAHBSIM) Model Description[a]

Guillermo Flichman[1]
Hatem Belhouchette[1]
Adam M. Komarek[2]
Sophie Drogue[3]
James Hawkins[2]
Roza Chenoune[1]
Siwa Msangi[2,*]

---

[1] Mediterranean Agronomic Institute of Montpellier (CIHEAM-IAMM)
[2] International Food Policy Research Institute (IFPRI-CGIAR)
[3] National Institute for Agricultural Research, France (INRA)
[*] Project Manager Contact Info: International Food Policy Research Institute, Washington D.C. , S.Msangi@cgiar.org

[a] Please cite this document as follows: Flichman, G., Belhouchette, H., Komarek, A.M., Drogue, S., Chenoune, R., Hawkins, J., Msangi, S.. 2016. Dynamic Agricultural Household Bio-economic Simulator (DAHBSIM) Model Description. *BioSight* project Technical Paper. International Food Policy Research Institute, Washington, D.C.

TABLE OF CONTENTS

LISTING OF TABLES

Listing of Figures

ACKNOWLEDGEMENTS

The CGIAR Research Program on Policies, Institutions, and Markets funded part of this research. The United States Agency for International Development funded the collection of the household survey data in Malawi, as part of the Africa Research in Sustainable Intensification for the Next Generation (AFRICA RISING) project. This research benefited from the help of Professor Maria Blanco (University of Madrid) and Dr. Wajid Nasim Jatoi (CIHEAM-IAMM) during the intitial model design phase.

1. INTRODUCTION

This document presents the architecture of a components-based bio economic model developed for the assessment of research into the sustainable intensification of agricultural production systems in Africa south of the Sahara. The goal of the document is to describe the basic model framework.

The Dynamic Agricultural Household Bio-Economic Simulator (DAHBSIM) was developed out of a joint project effort between researchers at the Mediterranean Agronomic Institute of Montpellier (IAMM), the University of Madrid, the French National Institute for Agricultural Research (INRA), and the International Food Policy Research Institute (IFPRI). The origins of DAHBSIM come from the earlier FSSIM (Farm Systems SIMulator) model (Louhichi et al, 2010). The following sections describe the theoretical and technical aspects of the model. The next section, Model Overview, provides a general background on the modelling framework. Next, the cropping module, and the biophysical module within it are described, followed by the livestock module, and the socieoeconomic components of the model. At the end of each section describing each individual module, the parameters, variables, and equations for that particular module are listed. The reader is referred to Appendix A where a concise description of how the model runs can be found.

2. MODEL OVERVIEW

2.1 Overview of the DAHBSIM Model Structure

DAHBSIM is a dynamic, bio-economic model of agricultural households that was designed to be applied to a rural, developing country-setting, for the purpose of addressing questions around the biophysical constraints to on-farm agricultural productivity, and the whole-farm implications of alternative strategies to sustainable agricultural intensification. The model links socio-economic and biophysical aspects, in order to better illustrate the environmental and human welfare implications of different agricultural production practices, as they are influenced by policy-driven changes in prices of inputs or outputs, or by changes in the physical environment.

Modelling Framework: Recursive Inter-Temporal Optimization

DAHBSIM is a recursive inter-temporal model:

-Inter-temporal because equations are indexed over years and the decisions of farmers are optimized given a discounted utility function

-Recursive because the results obtained at the end of a specific simulation are used as starting values for the next simulation. Particularly, the yields obtained at the end of the first (inter-temporal) simulation are multiplied by the biophysical stress coefficients which increase or decrease the yields in the next (inter-temporal) simulation depending on the precedent crop and the next year's precipitation

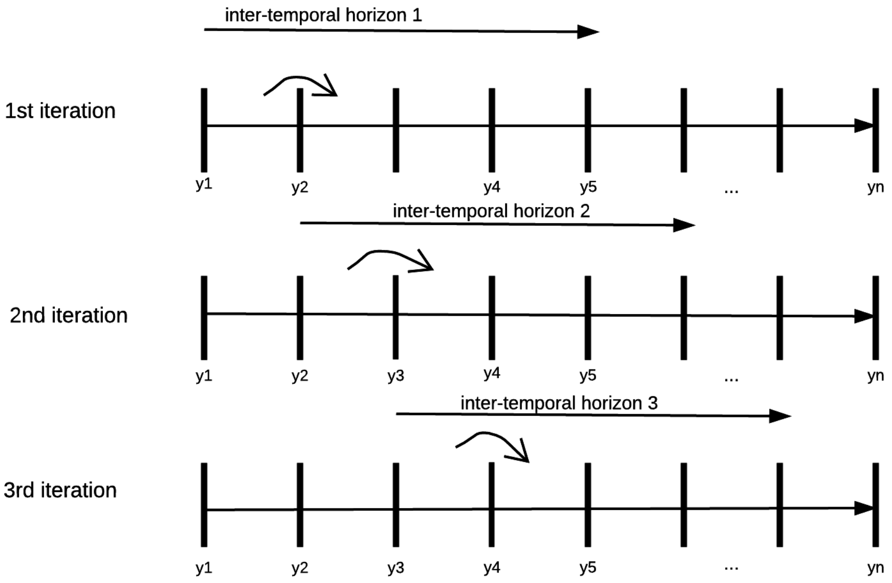
For a previous implementation of a similar this modelling framework, see the model description for FSSIM-DEV (Louhichi and Gomez Y Paloma, 2014). Recursive updates occur on a yearly time step in DAHBSIM. However, particular processes are accounted for on a sub-annual basis (e.g. livestock numbers, cropping activities). The objective function is defined over a time period (the inter-temporal horizon) which can be pre-specified by the user. The time frame over which the model operates is also pre-specified by the user. The model objective function is run at each recursive iteration (year, in this case) and production parameters are updated each year based on the previous year's decisions.

The essential components of the model are as follows:

- In our model the decision maker has a multi-year planning horizon. The exact number of years can be specified by the model user
- The dynamics of the model are designed to capture long term changes in soil productivity, based on farm management decisions, which are updated on an annual basis
- We account for price and yield variability in the agent's decision making process (via the risk module)

- The expected crop yield in future periods is based on the current year value, but each year actual crop yields are updated based on previous periods' management decisions

Fig**ure 1: Recursive inter-temporal optimization**



Note: n is the total number of years in the model horizon, and y1,y2,… represent specific years within the model horizon

Objective Function

The model objective involves the maximization of a risk adjusted net present value term. Net present value is first defined as the present value of the future income (revenues minus costs) from crop and animal sales, plus the value of consumption:

$$\text{NPV} = \sum_{y=1}^{n} \text{(full income + value of consumption)} \times (1/(1+i)^{y}) \qquad [1]$$

Where

NPV is net present value (USD hh^{-1} yr^{-1})

n is the number of years in the model horizon (hereafter referred to as the inter temporal horizon)
y represents a specific year within the model horizon
i is the discount rate (taking a value of 0.04 as default)
full income is defined as off farm income plus revenues minus costs for crop and livestock production, and is measured in USD hh^{-1} yr^{-1}
value of food consumption is defined as the value of food consumed from on-farm production (USD hh^{-1} yr^{-1})

The Price and yield variability enters negatively into the objective function under the assumption of a risk averse decision maker. Variability in output prices and yields are calculated based on data of price variability for the study region (FAO, 2016a, 2016b). The standard deviation of net present value, σ, is calculated assuming 50 different states of nature, each one with a random market price and crop yield. These are based on the actual variability of market prices and yields, respectively. The standard deviation is calculated in the Risk module (Section 4.5). The model objective is thus defined as follows:

$$\text{Maximize NPV} - \Phi \times \sigma \qquad [2]$$

Where

NPV is as defined above, and

σ is the standard deviation of net present value

Φ, phi is a risk aversion coefficient

The modularity of DAHBSIM

DAHBSIM is designed to be run as a full household model, including both production and consumption, as well as a production only model. Further, the livestock module can be turned on and off. The result is four different model types which can be interpreted as simulating different types of households, or be used to conduct policy analysis while controlling specifically for feedbacks between consumption and production, as well as crops and livestock. The following describes each of these models.

With both the consumption module and livestock module on, the 'full' model (Model 1) simulates all the essential components of the farm household, specifically how consumption is obtained between self-produced versus market purchases goods, and including the biophysical interactions between the on farm livestock and relevant linkages with soil and crops.

With the consumption module off and livestock module on, Model 2 simulates all the relevant biophysical interactions between crop and livestock production, but does not include household expenditure and production allocated to consumption.

Without livestock, Model 3 is a simplified version of the farm household, which simulates a farm household only producing crops. Model 4 simulates a farm household only producing crops and without consumption.

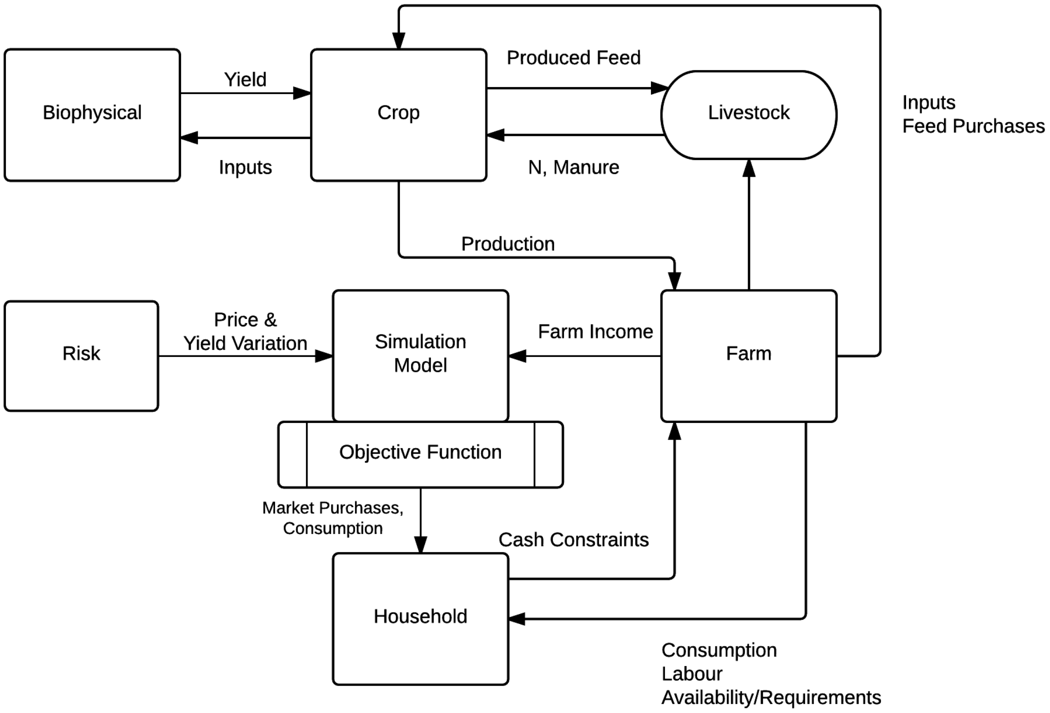
Table 1: Overview of Model Modularity

Model 1: All	Model 2: Supply Only	Model 3: No Livestock	Model 4: No Livestock and Consumption
Crop Livestock Risk Farm Household Consumption	Crop Risk Farm Livestock Household	Crop Risk Farm Household Consumption	Crop Farm Risk Household

Overview of key modules

The crop module contains the equations which describe the cropland allocation, the labor use, the rotation constraints, and the production of crops and crop residues. The farm module contains the equations which describe the resource constraints. The household module contains the equations which define household demand and time allocation.

Figure 2: Overview of model structure



In order to help the reader to understand better how the principal socio-economic and biophysical components of the model are put together – we now present the details on each of the main components of the DAHBSIM model in the following sections.

2.2 The modeling of crop production in DAHBSIM

Crop Module

Starting with the cropping model, we now describe the key biophysical components that represent the constraints that farmers face in on-farm productivity, which is captured by DAHBSIM's biophysical module. The following sections describe the structure of the crop and biophysical modules. The results of the calibration of the biophysical module with observed crop yields in the case study in Malawi are presented.

Cropland Accounting

Cropping activities in DAHBSIM are differentiated by soil type, crop, preceding crop and management intensity. The preceding crop has an influence on soil fertility and, therefore, in input requirements and yields for the current crop. The influence of the preceding crop in this year's crop yield is taken into account by integrating a biophysical module, described in the next section. The following equations form the core part of the crop module. The crop land accounting constraint specifies the breakdown of total cropland to individual cropping activities by soil and is defined as follows:

$$\sum_{c} v\_croparea_{hh,c,s,y} = v\_cropland_{hh,s,y} \qquad [3]$$

Where

hh is the index for household type
c is the index for crop
s is the index for soil type
y is the index for year
v_cropArea is the crop area by household and soil type
v_cropLand is the total cultivated land

Crop Area by Household and Soil Type

The crop area over all preceding crops and intensities must sum to the total crop area for a specific crop and soil, as defined by the following equation:

$$\sum_{cp} \sum_{t} v\_cactLev_{hh,c,p,s,t,y} = v\_cropArea_{hh,c,s,n} \qquad [4]$$

Where

cp is the index for preceding crop

t is the index for intensity level

v_cactLev is the crop activity level (ha)

Rotation Constraints

Each cropping activity, input-output relationships (i.e. seeds, labor requirements, agrochemicals, yields, externalities) is defined based on survey data and expert knowledge. Rotation constraints are defined endogenously in a dynamic way. They express that total land allocated to a particular crop in a particular soil type this year cannot exceed the land allocated to preceding crops in the rotation last year. The crop rotation constraint is thus defined as follows:

$$v\_cropArea_{hh,cp,s,n} \leq \sum_{cp} \sum_{t} v\_cactLev_{hh,c,cp,s,t,y-1} \qquad [5]$$

Labor Constraints

Labor constraints are indexed by month of the year and specify that total labor requirements of the production plan cannot exceed labor availability. Family labor availability is highly detailed, differentiating between male and female labor. Apart from family labor, hired labor and communal labor are considered. These constraints express that total labor requirements have to be fulfilled either by family or hired labor. A further differentiation between male and female labor is also included. The labour constraint is thus defined as follows:

$$\sum_{c} \sum_{cp} \sum_{t} p\_labreq_{c,t,m} \times v\_cactLev_{hh,c,p,s,t,y} \leq v\_cropFlab_{hh,l,m} \qquad [6]$$

Where

p_labreq is the labor requirements by crop activity, intensity, and month

v_cropFlab is the family labor used for cropping activities

v_cropHlab is the hired labor used for cropping activities

l is the labour type

m is the month

Agricultural land is differentiated between cropland and permanent grassland. For each land type and soil type, the total endowment of the household is equal to the initial endowment. Total cropped land is defined as the land occupied by the different crops, and cannot be greater than the land endowment. For a complete listing of the parameters, variables, and equations in the crop module, see Tables 2, 3, and 4.

Table 2: Crop module variables

Variable Name	Units	Description
v_cactLev	ha	crop activity level by cp
v_cactPrd	ha	crop activity total
v_cactYld	t ha^{-1}	crop activity yield
v_prodQuant	t	production quantity
v_cropArea	ha	crop area by soil type
v_cropLand	ha	cropland used
v_cropLabor	person-days	labor used for cropping activities
v_inputUse	kg y^{-1}	input use
v_seedQuant	kg ha^{-1}	total seed quantity
v_residuesfeedm	kg m^{-1}	crop residues allocated for potential livestock feed intake or for feed balance each month
v_residuesmulch	kg y^{-1}	crop residues allocated to crops for mulch
v_residuesfeed	kg y^{-1}	crop residues allocated for potential livestock use or for feed balance
v_residuessell	kg y^{-1}	crop residues sold
v_residuessellm	kg y^{-1}	crop residues sold per month

Table 3: Crop module parameters

Parameter Name	Units	Description	Source File
p_cropCoef; other	nc ha^{-1}	Value of other inputs per ha	cropcoef.xlsx
p_cropCoef; nitr	kg N ha^{-1}	Nitrogen applied per ha	cropcoef.xlsx
p_cropCoef; seed	kg ha^{-1}	Seed applied per ha	cropcoef.xlsx
p_cropCoef; labor	Labor days ha^{-1}	Labor days per ha	cropcoef.xlsx
p_cropCoef; yield	kg ha^{-1}	Yield per ha	cropcoef.xlsx
p_cropCoef; straw yield	kg ha^{-1}	Straw yield per ha	cropcoef.xlsx
p_cropCoef; phyto	Phyto nc ha^{-1}	Value of phytochemicals per ha	cropcoef.xlsx

p_landReq	ha	Land requirement (growth period)	cropcoef.xlsx
p_laborReq	person-days ha^{-1}	Labor requirement (person-day per ha)	cropcoef.xlsx
p_inputReq	kg ha^{-1}, nc ha^{-1}	Direct input requirements	cropcoef.xlsx
p_cropData; area	ha	Area of crop	cropcoef.xlsx
p_cropData; yield	kg ha^{-1}	Yield per ha	cropdata.xlsx
p_cropData; yield straw	kg ha^{-1}	Straw yield per ha	cropdata.xlsx
p_cropData; seed	kg ha^{-1}	Seed applied per ha	cropdata.xlsx
p_cropData; cprd	kg ha^{-1}	Seed production per ha	cropdata.xlsx
p_cropData; labor	person-days	Person days of labor	cropdata.xlsx
p_cropData; nitrogen	kg ha^{-1}	Nitrogen applied per ha	cropdata.xlsx
p_cropData; fertilizer	nc ha^{-1}	Value of fertilizer per ha	cropdata.xlsx
p_cropData; phyto	nc ha^{-1}	Value of phytochemicals per ha	cropdata.xlsx
p_cropData; other	nc ha^{-1}	Observed baseyear data	cropdata.xlsx
p_seedData	kg	Observed seed data	cropdata.xlsx
p_perresmulch	percentage	% of crop residue production allocated to field as mulch	cropcoef.xlsx
p_residuedep	percentage	% of feed not wasted between months, then can be carried forward from month m to month m+1	cropcoef.xlsx
v0_cropYld	kg ha^{-1}	Crop yield	Determined endogenously after first iteration
v0_cactLev	ha	Observed activity level	cropdata.xlsx
v0_cropArea	ha	Observed crop area by soil type	cropdata.xlsx
v0_cropLand	ha	Observed crop land	cropdata.xlsx
v0_cactPrd	t	Observed crop activity production	cropdata.xlsx
v0_prodQuant	t	Observed crop output	cropdata.xlsx
v0_cactYld	kg ha^{-1}	Crop activity yield	cropcoef.xlsx
v0_inputUse	kg	Input use	cropdata.xlsx, cropcoef.xlsx
v0_inputCost	nc	Input cost	cropcoef.xlsx
v0_seedUse	kg	Seed use	cropdata.xlsx

Note: ‘nc’ refers to the native currency of the region for the specific model case study

Table 4: Crop module equations

Equation	Definition
E_CROPLAND	$v\_cropLand_{hh,s,y} = \sum_{caen} v\_cropArea_{hh,caen,s,y}$
E_CROPLABOR	$v\_cropLabor_{hh,y,m} = \sum_{caen,cp,s,tt} v\_cactLev_{hh,caen,cp,s,tt,y} \times p\_laborReq_{hh,caen,tt,m}$
E_CROPAREA_EN	$v\_cropArea(hh,caen,s,y) = \sum_{c\_c(caen,cp),tt} v\_cactLev_{hh,caen,cp,s,tt,y}$
E_ROTATION	$\sum_{c\_c(caen,cp),tt} v\_cactLev_{hh,caen,cp,s,tt,y} < v\_cropArea_{hh,cp,s,y-1} + v0\_cropArea_{hh,cp,s}$ (if y = 1)
E_CACTPRD_EN	$v\_cactPrd_{hh,caen,s,tt,y} = \sum_{c\_c(caen,cp)} v\_cactLev_{hh,caen,cp,s,tt,y} \times v\_cactYld_{hh,caen,cp,s,tt,y}$
E_CACTYLD_EN	$v\_cactYld_{hh,caen,cp,s,tt,y} = v0\_cactYld_{hh,caen,cp,s,tt}$
E_CROPPRD_CJ	$v\_prodQuant_{hh,cjen,j} = \sum_{a\_j(caen,cjen),s,tt} v\_cactPrd(hh,caen,s,tt,y)$
E_CROPPRD_CK	$v\_prodQuant_{hh,cken,y} = \sum_{a\_k(caen,cken),cp,s,tt}\$c\_c(caen,cp)\ v\_cactLev_{hh,caen,cp,s,tt,y} \times p\_cropcoef_{hh,caen,s,tt,'ystraw'}$
E_INPUTUSE_EN	$v\_inputUse_{hh,caen,cj.y} = \sum_{c\_c(caen,cp),s,tt} v\_cactLev_{hh,caen.cp,s,tt,y} \times p\_cropCoef_{hh,caen,s,tt,ci}$ v_inputUse(hh,caen,ci,y) = sum((c_c(caen,cp),s,tt), v_cactLev(hh,caen,cp,s,tt,y)*p_cropcoef(hh,caen,s,tt,ci))
E_SEEDUSE	$v\_seedQuant_{hh,caen,y} = v\_inputUse_{hh,caen,'seed',y}$ _seedQuant(hh,caen,y) = v_inputUse(hh,caen,'seed',y)
E_RESIDUES	$v\_residuesfeed_{hh,cken,y} + v\_residuessell_{hh,cken,y} + v\_residuesmulch_{hh,cken,y} = v\_prodQuant_{hh,cken,y}$

E_RESIDUESMULCH	$v\_residuesmulch_{hh,cken,y} = v\_prodQuant_{hh,cken,y} \times p\_perresmulch_{cken}$
E_RESIDUESFEED	$v\_residuesfeed_{hh,cken,y} = \sum_{m} v\_residuesfeedm_{hh,cken,y,m} \times flagm_{cken,m}$ v_residuesfeed(hh,cken,y)=sum(m,v_residuesfeedm(hh,cken,y,m)*flagm(cken,m))
E_RESIDUESSELL	$v\_residuessell_{hh,cken,y} = \sum_{m} v\_residuessellm_{hh,cken,y,m} \times flagm_{cken,m}$

2.3 Biophysical Module of DAHBSIM

The biophysical model is a summary model aimed at estimating the effect of farm management and climate on the evolution of yields over time. It is a multi-year, multi-crop, monthly time step cropping systems simulation model. It was developed to serve as an analytical tool to study the effect of climate, soils, and management on cropping systems productivity and the environment. The cropping system model simulates, in a summary way, the soil water (including water use and drainage) and nitrogen budgets (including residue production and decomposition) and crop yield. The model has a generic crop simulator that enables the simulation of different type of crops and both yearly and multi-year crops and crop rotations via a single set of parameters. Simulations can be a fraction of a year (at least for some variables such water drainage and yield) to several years. These processes are affected by weather, soil properties, crop characteristics, and cropping system management options including crop rotation, cultivar (variety) selection, irrigation and nitrogen fertilization.

Overall, the summary model is intended for crop cycle simulation over a single land block fragment with uniform soil, weather, crop rotation and management. The crop cycle is described at the level of whole plant. Figure 1 shows a flowchart describing the approach used in the summary model to calculate actual yield. The core of these calculations is the determination of unstressed (potential) yield based on crop potential evapotranspiration. This potential yield is then corrected by water and nitrogen limitations, if any, to determine actual yield. Consequently, during each crop cycle simulation, the actual yield for the year (Ya) is taken as the minimum of the yield limited by water (Yw) and by nitrogen (Yn).

Actual yield limited by water (Yw)

The actual activity yield limited by water in the summary cropping system model is calculated as follows (Doorenbos *et al,* 1986):

$$1 + \frac{\mathrm{Ym}}{\mathrm{Ym}} = Ky\ (1 - \frac{ETa}{ETm}) \qquad [7]$$

Where

Yw is activity yield limited by water (kg ha^{-1})

Ym is crop maximum yield (kg ha^{-1}). This parameter depends on cultivar properties

Eta is activity actual evapotranspiration (mm day^{-1})

ETm is crop potential evapotranspiration (mm day^{-1})

Ky is yield response factor to water stress (can be negative, positive and greater than 1).

The actual activity yield limited by nitrogen in the summary cropping system model is calculated as follows:

$$\mathrm{Gn} = \mathrm{PNG}\ [1 - \left(\frac{\mathrm{NCcrit} - \mathrm{NCONCa}}{\mathrm{NCcrit} - \mathrm{NCmin}}\right)] \qquad [8]$$

Where

G_N is the Nitrogen dependent growth (kg ha^{-1}).

PNG is Potential growth after other limiting factors have been accounted for (kg ha^{-1})

$NCONC_a$ is the crop nitrogen concentration after new growth (kg ha^{-1})

NC_{crit} is the critical nitrogen concentration required by the crop to grow potential rate (kg ha^{-1})

NC_{min} is the crop minimum nitrogen concentration at which growth stops (kg ha^{-1})

As the ratio between actual nitrogen absorbed (NCcrit – NCONCa) and nitrogen absorbed for potential growth (NCcrit – NCmin) are correlated with the ratio actual yield to potential to potential yield, the dependent nitrogen-growth is calculated as following:

$$\mathrm{Yn} = \mathrm{Yw}\left(1 - \frac{\mathrm{Nab}}{\mathrm{Npot}}\right) \qquad [9]$$

Where

Y_n is the Actual yield after water stress (kg ha^{-1}).

$N_{pot} = Y_{pot}/k$, Where N_{pot} is the amount of N to grow at potential level (kg ha^{-1})

Y_{pot} is the potential yield (without N and Water stress) (kg ha^{-1})

KN = coefficient for N conversion to crop biomass

Equation [9] then becomes:

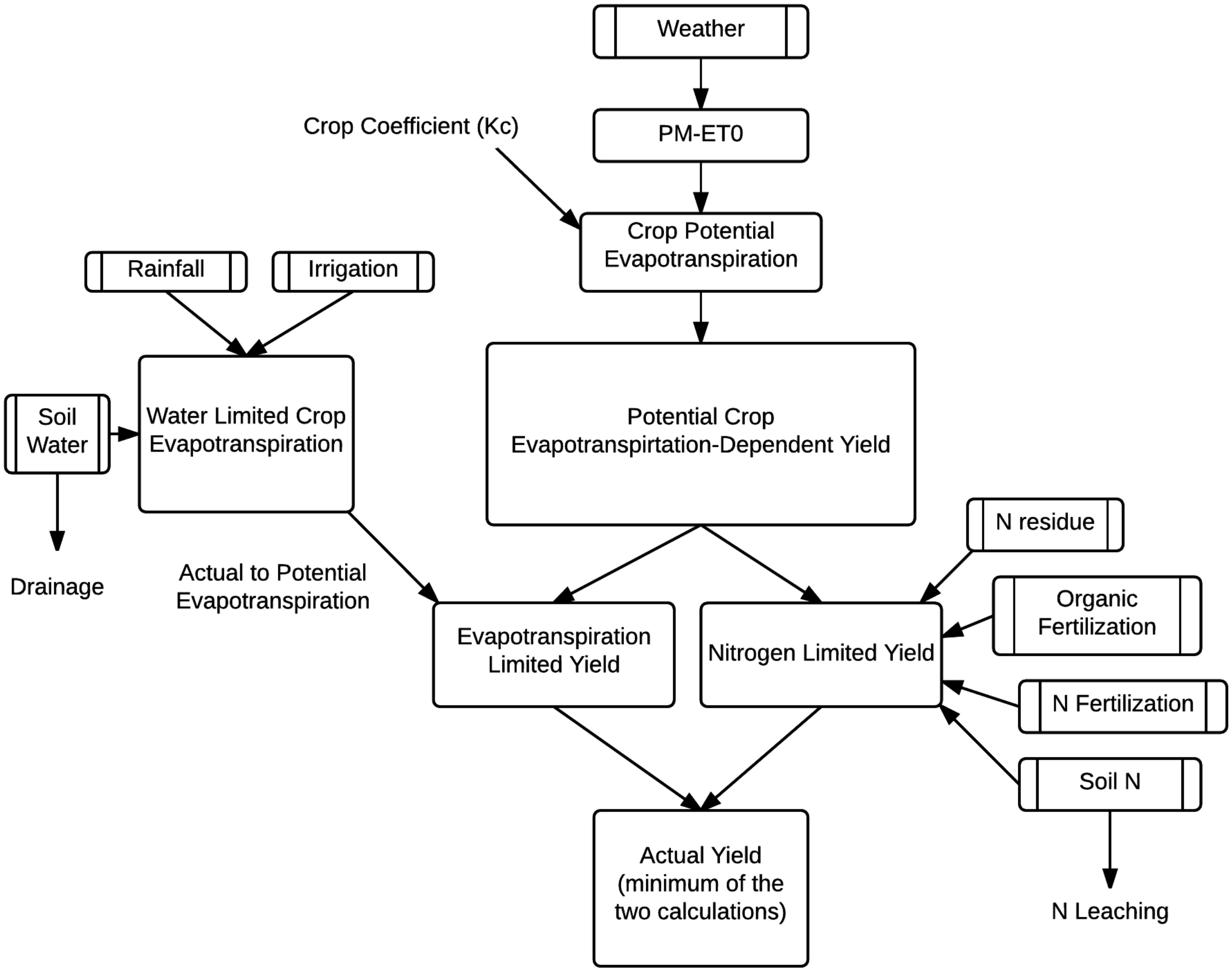
$$YN = YW\left(1 - \frac{Nab}{\left[\frac{Ypot}{k}\right]}\right) \qquad [10]$$

Model parameterization

In order to assess model performance in a large range of cropping conditions such as in our case study, two steps are usually followed (Belhouchette et al., 2010, Belhouchette et al., 2008, Oreskes et al., 1994). First (i) the crop model is calibrated by using experimental data with several dynamic and cumulative variables (yield, biomass, leaf area index, N-leaching) but under limited cropping conditions, and then (ii) the crop model is validated for a wider range of cropping systems but usually only for crop yield which is the common variable measured in all crop experiments (Therond et al, 2010; Faivre et al., 2004; Van Ittersum et al., 2003; Jagtap and Jones, 2002; Bouman et al., 1996).

These sources of information have at least two main drawbacks: i) observed data needed for model evaluation mainly require observations that are usually time consuming and costly and so performed under limited soil, crop management and climate conditions, and ii) they do not take into account the interactive effects of soil and management on output data for different climate (years) conditions.

Figure 3: Diagrammatic Overview of Crop Yield Calculation



In our study, to evaluate the performance of the DAHBSIM summary cropping system model at regional scale, i.e. covering a large range of crop, management, soil, and climate conditions (see for example Belhouchette et al., 2010), the input-output data used to describe and evaluate the cropping systems performance are obtained through regional farmers surveys and existing regional databases as suggested by Clavel et al., (2011); Therond et al., (2010) and Faivre et al., (2004) (see Data section for description of household survey).

For each activity, the crop summary cropping system model was evaluated in two steps. In the first phase it was parameterized by calibrating the model for each activity cultivated with the extensive technique in clay soil (e.g. maize, extensive, clay). In the second phase it was evaluated for the same activity but for more contrasted crop management (intensive technique and loam, sand, and other soil types).

Table 5: Crop Activities by Soil and Intensity

Crop Variety	Cultivar	Intensity	Soil Types
Maize	Improved Local	Intensive, Medium, Extensive	Clay, Sand, Loam, Other
Soybean	Improved Local	Intensive, Medium, Extensive	Clay, Sand, Loam, Other
Groundnut	Improved Local	Intensive, Medium, Extensive	Clay, Sand, Loam, Other
Beans	Improved Local	Intensive, Medium, Extensive	Clay, Sand, Loam, Other
Cowpea	Improved Local	Intensive, Medium, Extensive	Clay, Sand, Loam, Other

Note: Intensity level refers to level of external inputs (and yield)

By doing so, only the conversion nitrate to crop yield coefficient (KN) and the yield response factor to water stress (ky), were determined by calibration since the model were sensitive to these parameters under rainfed conditions (Belhouchette et al., 2012). Values of Kn and ky were adjusted within a reasonable range of variation (Donatelli *et al.*, 1997) based on previous research, knowledge or experience in order to have the best model estimation of the yield observed for each activity from the survey. In order to ensure a good correlation between observed and simulated data sets, the adjustment process was stopped when further modification of crop parameters values generate little or no change on the basis of the relative root mean square error (RRMSE) (Loague and Green, 1991). The simulation is considered excellent with RRMSE < 10%, good if 10 – 20%, fair if 20 – 30%, poor >30% (Jamieson *et al.,* 1991).

$$RRMSE = \frac{\sqrt{\sum \frac{(Oi-Si)^2}{n}}}{\bar{O}} \text{ x } 100 \qquad [11]$$

Where

S_i is the simulated value for observation i

O_i is the observed value for observation i

$\bar{O}$ is the average observed value, and

n is the number of observations.

Model Inputs

Table 2 summarizes the crop input parameters which can be either, i) available in the literature, ii) calibrated, or calculated from field data to compare model output against observed field data. The summary inputs were set based on:

- Soil parameters: The bulk density, the soil water holding capacity, the soil depth, the mineralization rates from humus and the initial soil water, nitrate and organic matter contents for each soil type were determined using experimental data (Table 6).
- Weather: The monthly precipitation was available at the study area (Harris et al, 2014). Potential evapo-transpiration was calculated using the Priestley-Taylor method (Priestley and Taylor, 1972).
- Management: The amounts and timing of irrigation and nitrogen fertilization and planting and harvest dates were collected from farmers (see data section).

- Crop: The phenological stages, growth and morphologic characteristics such as maximum rooting depth were compiled from literature or from farmer's interviews (Table 1).

Table 6: Crop input parameters used in the summary model simulation.

	Source	**bean**	**maize**	**cowpea**	**groundnut**	**soybean**
Crop Phenology						
Crop cycle (days)	*Survey*	120	210	150	180	90
Water component						
Maximum yield (kg ha^{-1})	*Literature*	3500	10000	4000	1500	8500
Crop Coefficient Kc	*Literature*	1.15	1.15	1.15	1.05	1.1
Yield response factor to water stress (Ky)	*Calibrated*	1.15	1.25	1.15	0.7	0.85
ET0 (mm day^{-1})	*Calculated*	4.5	4.5	4.5	4.5	4.5
Maximum root depth (m)	*Literature*	0.9	0.9	0.7	0.8	0.9
Nitrate component						
Nitrate to crop yield coefficient (KN) (kg NO_3 kg^{-1} biomass)	*Calibrated*	3.5	20	2	5	10
N by kg of crop biomass (kg ha^{-1})	*Literature*	0.45	0.3	0.5	0.5	0.5

Sources: Parameters were extracted from the literature (FAO, 1986), calculated, from the survey or from calibration.

Table 7: Soil organic matter, soil water holding capacity and initial soil N and water contents in the four soil types of the study area.

Soils	**Organic matter (%)**	**Soil N content (kg ha^{-1})**	**Soil water content (m^3 ha^{-1})**	**Soil water holding capacity (mm m^{-1})**
Clay	0.05	15.5	80	120
Loam	0.025	8.5	70	100
Sand	0.02	6.5	45	50
Loamy-sand	0.01	6.5	70	80

Source: Ollenburger and Snapp (2014); Ollenburger (2012). The initial soil water content was estimated by considering that sowing should be achieved at least when soil content is at 2/3 of its total water holding capacity which is a common planting condition for the area (Stern and Cooper, 2011).

Biophysical module calibration

Table 8 presents a comparison between measured and simulated yield for the 5 crops. For all crops, except for cowpea, mean simulated yields were close to the mean measured yield. For groundnut the model gave a good estimation of yield, with RRMSE lower than 20%. The results were fair for bean, maize and soybean. However, the results were less satisfactory for cowpea with an RRMSE of 45%. For all crops, except for cowpea, and to a lesser degree for soybean, the slopes and intercepts of the regression equations for the measured and simulated yields followed the 1:1 line closely (figure 4).

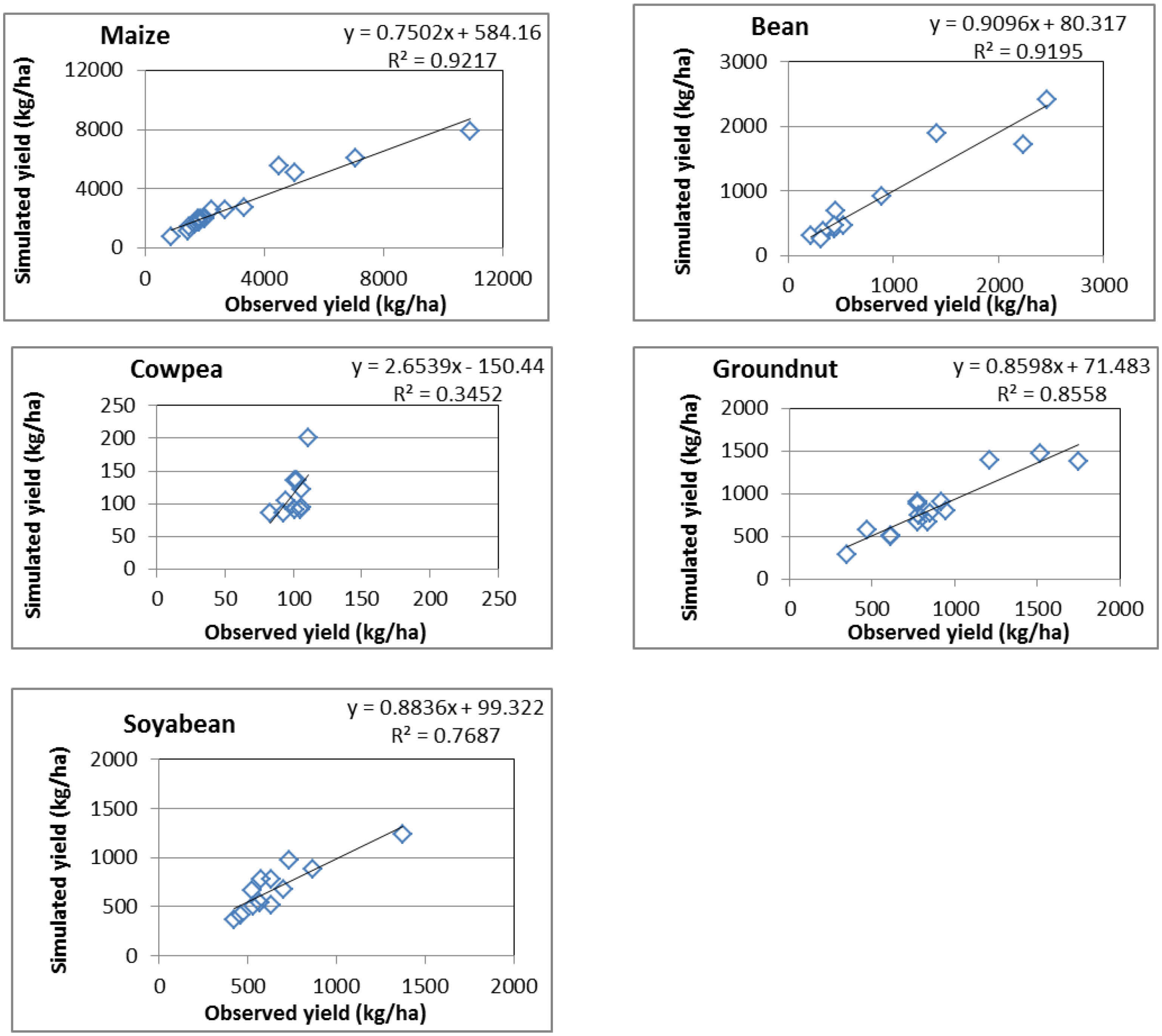
Concerning the high value of the cowpea RRMSE value, farm-survey data of cowpea yield were affected by a large variability, and this increased the uncertainty of model evaluation. In fact, at least for two activities among 8, the average standard deviation was about 1000 kg/ha (for an average yield of 1736 kg/ha) and 1500 kg/ha (for an average yield of 3600 kg/ha). By deleting these activities, the RMMSE and the R^2 become respectively 8% and 95.

Table 8: Simulated vs. observed yields and RRMSE for each crop.

Crops	Simulated yield (kg/ha)	Observed yield (kg/ha)	RRMSE (%)
Bean	785	774	26
Maize	2957	3164	27
Soybean	692	671	18
Cowpea	193	496	45
Groundnut	829	881	16

Notes: Each value represents an average by crop for the different soil types (clay, loam, loamy-sand, sand) and crop management (intensive, extensive).

Figure 4: Simulated versus observed yield for model crops



Note: Each point represents a combination of crop management (intensive, extensive) and soil type

Table 9: Water Module Parameters

Parameter Name	Units	Description
ym	t ha^{-1}	Maximum yield
p_ky	t ha^{-1}	Yield response factor specified by crop
p_hw	-	Water stress factor
p_etm	mm day^{-1}	Maximum evapotransporation
p_kc	-	Crop coefficient
p_et0	mm day^{-1}	Reference evapotransporation
p_eta_m	mm	Monthly actual evapotransporation
p_etm_m	mm	Monthly maximum evapotransporation
p_etm_t	mm	Annual maximum evapotransporation
p_eta_t	mm	Annual actual evapotransporation
p_idat;luse	-	Irrigation parameters for crops;
p_idat; nirr	-	Irrigation parameters for crops;
p_luse	-	Crop cycle - land use coefficient
p_nirr	mm $month^{-1}$	Monthly net irrigation
p_rain	mm $month^{-1}$	Monthly effective rainfall
p_meteo	mm $month^{-1}$	Monthly effective rainfall
p_asi	mm $month^{-1}$	Available soil water index
p_swd0	mm $month^{-1}$	Initial soil water depth
p_swd	mm $month^{-1}$	Actual soil water depth at sowing
p_wdf	-	Soil water depletion fraction for crop groups and Etm
p_factor	-	Actual soil water depletion fraction
p_rdm	m	Maximum rooting depth by crop
p_swa	mm m^{-1} soil depth	Soil water available by crop
p_swr	mm	Remaining soil water
p_swa_m	mm m^{-1} soil depth	Soil water available by crop
p_eta_tab	-	Table to calculate eta as function of asi and swr
p_swpar	-	Soil water parameters
p_swm	mm m^{-1} soil depth	Maximum soil water available
p_cropCoef	-	Technical coefficient (see crop module)
v0_cactYld	t ha^{-1}	Crop activity yield

Table 10: Nitrate Module Parameters

Parameter Name	Units	Description
p_cropCoef	-	Technical coefficient (see crop module)
p_Qfert	kg ha^{-1}	Quantity of fertilizer by activity
p_Qres	kg ha^{-1}	Quantity of residues from precedent crop
p_Qcomp	kg ha^{-1}	Quantity of compost
p_humus	kg ha^{-1}	Quantity of humus
P_Npot	kg ha^{-1}	N necessary for optimal growth
p_Nmin	kg ha^{-1}	N mineralized from humus
p_Nab	kg ha^{-1}	Current absorption of Nitrate
p_Nav	kg ha^{-1}	Available Nitrogen
p_Nini	kg ha^{-1}	Initial amount of Nitrogen
p_NI	kg ha^{-1}	N leaching (fixed at 10% of N_fert)
p_Nw	-	Coefficient of Nitrogen stress
p_Nitr	kg ha^{-1}	Total Nitrate requirements by activity
p_Nres	kg ha^{-1}	N from precedent crop residue
p_Norg	kg ha^{-1}	N from organic fertilization
p_Nfert	kg ha^{-1}	N from mineral fertilization
p_Ncomp	kg ha^{-1}	N from compost
p_Nfin	kg ha^{-1}	N at the end of the growth period
p_Hfin	%	Final amount of humus
p_Hini	%	Amount of inorganic matter (initial amount of humus)
p_Mscomp	%	Amount of biomass from compost
K3	-	Humification rate
yn	-	Parameter
v0_cactLev	ha	Crop activity level by cp
p_cropData	-	Crop Data (see crop module)
p_hini	%	Amount of inorganic matter (initial amount of humus)
K1	-	Coefficient for Nitrogen conversion
K2	-	Mineralization rate
da	kg m^{-3}	Soil bulk density
prof	m	Ploughed layer
p_MSres	%	% of dry matter from Qres
p_effr	kg of biomass	Nitrate by kg of biomass
p_Nw2	-	Coefficient of nitrogen stress
p_OrgMat	%	Amount of organic matter

3. THE LIVESTOCK MODEL OF DAHBSIM

Overview of the livestock module

Livestock play a significant role in rural livelihoods around the world, contribute to food security and global environmental change, and consume natural resources (Herrero *et al.*, 2013). To account adequately for agricultural household livestock systems and their interactions with other components of the household within DAHBSIM, we developed a livestock module. This module describes the processes influencing the livestock component of the household, and its relationship with other components. The model focuses on how feeding decisions and opportunities alter both animal productivity and livestock numbers. Rufino *et al.* (2009) show that livestock feed management effectively increases animal productivity. The objective of the livestock module is to identify the effect of feeding strategies on herd productivity and production, while accounting for the resource-use, environmental impact, and demographics of these decisions. The model accounts for livestock enterprises with different animal types (beef cattle, dairy cattle, goats, and sheep), and different outputs (milk and meat) for three levels of intensification.

The livestock module predicts the nutrient (protein, energy and dry matter) demand for different types and classes of livestock based on obtaining a specified production level. The module calculates daily nutrient demands for animal maintenance, lactation, pregnancy, and growth. The module tracks herd or flock dynamics and feed supply and demand balances over time in a dynamic recursive manner. The two main caveats of the model are that 1) monthly demand for nutrients is constant every month, and 2) every month a constant fraction of animal's transition to a different class and pregnancy occurs in every month at the same rate.

Accounting framework for livestock nutrient requirements

The nutrient module calculates the metabolizable protein and net energy requirements associated with different levels of output for different types and classes of livestock (Figure 5). The model represents these calculations using a series of equations derived primarily from National Research Council nutrient requirements (NRC, 2000; NRC, 2001; NRC, 2007) and Jarrige (1989). Energy and protein are the two diet properties that affect livestock outputs in our model. We specify discrete levels of milk production, live weight gain, and live weight, and then we calculate the associated energy and protein requirements for the different livestock.

For a specific type and class of animal its monthly demand for net energy and protein is constant every month of every year. The model uses the following assumptions:

- A weight to volume conversion factor for milk of 1.04 kg l^{-1} (USDA, 2011).
- For dairy and beef cattle, milk has 3.5% fat and 3.5% protein (Wong *et al.*, 1988; NRC, 2001). For sheep, milk has 4.5% protein (CSIRO, 2007), and 6.5% fat. For goats, milk 3.2% protein (CSIRO, 2007), and 6.5% fat.
- Live weight means the average live weight of an animal in that specific age category.
- For the equations that calculate nutrient requirements based on equations from Jarrige (1989), we use a correction factor of 7.11 to convert units of feed for lactation (UFL) to megajoules.

The following sections describe the procedure used to calculate animal nutrient requirements.

3.1.1 Livestock net energy requirements

Requirements for energy are determined on a net energy basis, distinguishing between gross energy intake and the energy actually used for maintenance and productive purposes. Energy requirements are broken into five categories: maintenance, lactation, pregnancy, growth, and activity (confinement, pasture, or grazing). The total daily energy requirement of the animal is the sum of these five categories. Depending on the animal's life stage, and the production system, these requirements may take a value of zero.

Maintenance

The equation used to calculate maintenance requirements for cattle is based on the following equation from NRC (1996):

$$NE_m = C_m \times \text{live weight}^{.75} \qquad [12]$$

Where NE_M is the net energy for maintenance (MJ hd^{-1} d^{-1}), C_m is the coefficient of maintenance, and live weight is the live weight (kg hd^{-1}) of the animal. C_m takes a value of .322 and .386 for young/weaners and mature animals respectively (NRC, 1996).

The equation used to calculate maintenance requirements for sheep is based on the following equation from Jarrige (1989):

$$NE_m = 7.11 \times 0.033 \times \text{live weight}^{.75} \qquad [13]$$

Where NE_m is the net energy required for maintenance (MJ hd^{-1} d^{-1}), 7.11 is a conversion factor (from UFL to MJ) and live weight is the animal‘s live weight (kg hd^{-1}).

The equation used to calculate maintenance requirements for goats is based on the following equation from Jarrige (1989):

$$NE_m = 7.11 \times 0.039 \times \text{live weight}^{.75} \qquad [14]$$

Where NE_m is the net energy required for maintenance (MJ hd^{-1} d^{-1}), 7.11 is a correction factor (from UFL to MJ), and live weight is the animal’s live weight (kg hd^{-1}).

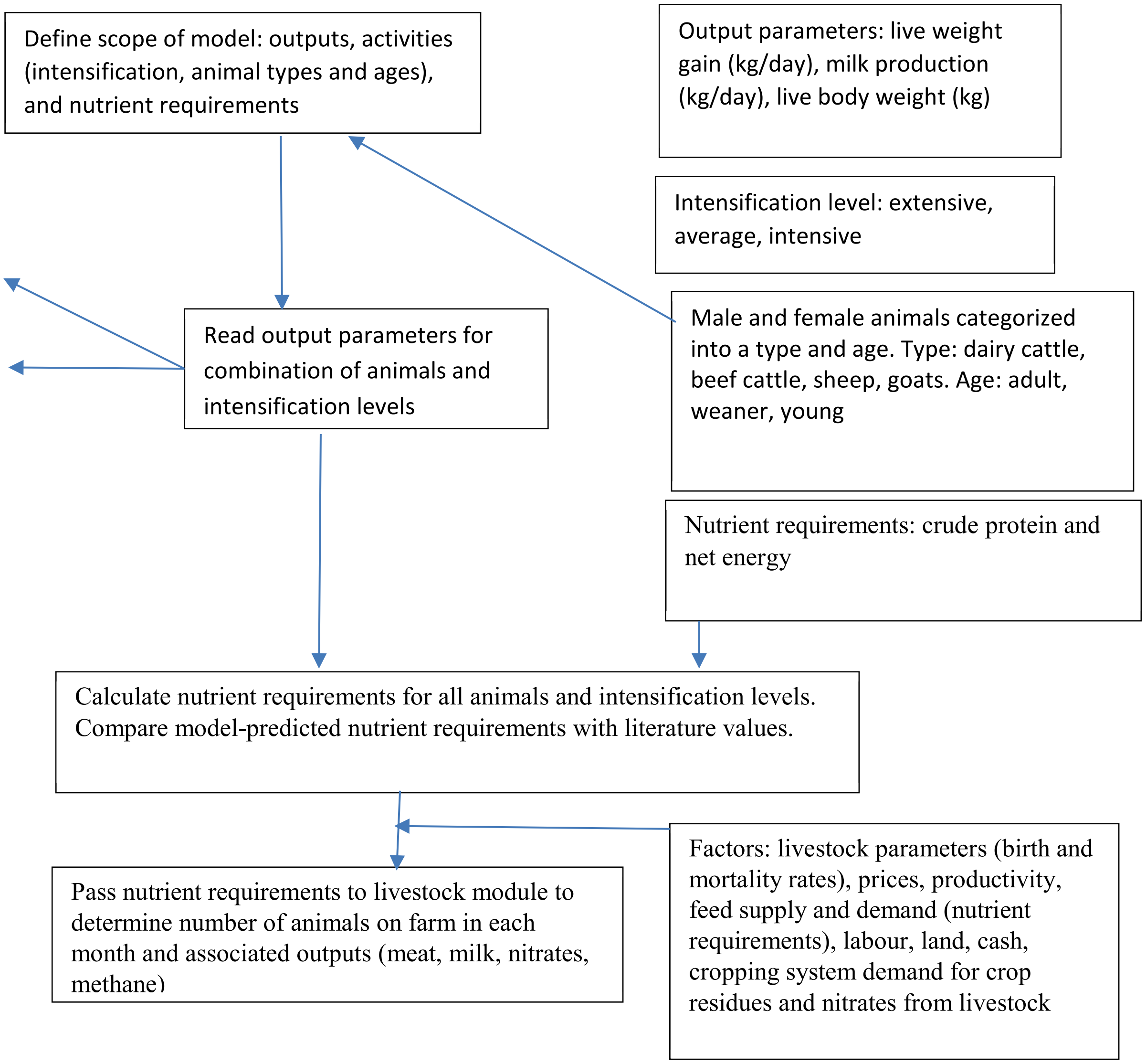
Lactation

Energy requirements for lactation for cattle are based on the following NRC (1989) equation:

$$NE_l = Y_m\,(1.47 + 0.40 \times \text{milk fat}) \qquad [15]$$

Where NE_l is net energy for lactation (MJ hd^{-1} d^{-1}), Y_m is milk yield (kg hd^{-1} d^{-1}). Milk fat is the fat percentage of milk (%) expressed as a number, for example, 3 not 3% (0.03).

Figure 5: Flow diagram illustrating the structure of the nutrient and livestock modules



Notes: The nutrient module calculates daily nutrient requirements for individual livestock types and ages and with different levels of output, and the livestock module determines the number of animals based on whole-farm trade-offs, resource demands, and economic factors.

Energy requirements for lactation for sheep are based on the following Jarrige (1989) equation:

$$NE_l = 7.11 \times Y_m \times (0.265 + 0.00588 \times \text{milk fat}) \qquad [16]$$

Where NE_l is net energy for lactation (MJ hd^{-1} d^{-1}), Y_m is milk yield (l hd^{-1} d^{-1}), and milk fat is the fat percentage of milk (%).

Our model follows Jarrige (1989) for lactating goat energy requirements:

$$NE_l = 7.11 \times Y_m \times 0.385 \qquad [17]$$

Where NE_l is net energy for lactation (MJ hd^{-1} d^{-1}), and Y_m is milk yield (l hd^{-1} d^{-1}).

Pregnancy

Requirements for pregnancy for cattle are obtained from the following NRC (1996) equation:

$$NE_p = C_p \times NE_m \qquad [18]$$

Where NE_p is the energy required for pregnancy (MJ hd^{-1} d^{-1}), and C_p is the coefficient for pregnancy. C_p takes a value of 0.1. For sheep and goats, requirements for pregnancy are obtained with the following equation from Cannas *et al.* (2004):

$$NE_p = .037 \times \text{live weight} \qquad [19]$$

Where live weight is the live weight of the animal (kg hd^{-1}).

Growth

Energy requirements for cattle growth are obtained with the following equation based on NRC (1996):

$$NE_g = 22.02 \times (\text{live weight} / (C_d \times \text{final weight}))^{0.75} \times ADG^{1.097} \qquad [20]$$

Where NE_G is the energy required for growth (MJ hd^{-1} d^{-1}), live weight is the live weight of the animal at any given point in time (kg hd^{-1}), final weight is the final weight of the animal over the period (kg hd^{-1}), and ADG is the average daily live weight gain of the animal (kg hd^{-1} d^{-1}). C_d is a coefficient that takes the value of 0.8 for females, 1.0 for castrates, and 1.2 for bulls (NRC, 1996).

Gibbs *et al.,* (2002) provides the energy requirements for growth of sheep:

$$NE_g = ADG \times (a + b\,(\text{live weight})) \qquad [21]$$

Where NE_G is the energy required for growth by the animal (MJ hd^{-1} d^{-1}), ADG is the average daily live weight gain of the animal (kg hd^{-1} d^{-1}), BW_i and BW_f are the animal's body weight at the beginning and end of the period respectively, and parameters a and b depend on the animal class and type. The constant "a" takes a value of 2.5, 4.4, and 2.1 for intact males, castrates, and females respectively (AFRC, 1993). The constant "b" takes a value of 0.35, 0.32, and 0.45 for intact males, castrates, and females respectively (AFRC, 1993).

We obtain energy requirements for the growth of goats from the following NRC (2007) equation:

$$NE_g = C_g \times ADG \qquad [22]$$

Where NE_g is the energy required for growth by the animal (MJ hd^{-1} d^{-1}), C_g is the coefficient of growth (MJ kg^{-1}), ADG is the average daily gain (kg hd^{-1} d^{-1}). C_g takes a value of 13.4 and 23.1 for young and weaning animals respectively (NRC, 2007).

Activity

We use NRC (1996) to determine requirements for activity for cattle:

$$NE_a = C_a \times NE_m \qquad [23]$$

Where NE_a is the energy required activity (MJ hd^{-1} d^{-1}), and C_a is the coefficient of activity. For cattle, it takes a value of 0, .17 and .36 for confinement, pasture and grazing animals respectively, based off.

We use the following equation to obtain requirements for activity for sheep and goats:

$$NE_a = C_a \times \text{live weight} \qquad [24]$$

Where C_A takes a value of 0.008, 0.01, and 0.02 for animals in confinement, pasture, and grazing respectively (AFRC, 1993), and live weight is the live weight of the animal (kg hd^{-1}).

3.1.2 Livestock metabolizable protein requirements

Maintenance

Maintenance requirements for metabolisable protein for cattle is obtained from the following Jarrige (1989) equation:

$$P_m = (95 + 0.5 \times \text{live weight})/1000 \qquad [25]$$

Where P_M is metabolisable protein required for maintenance (kg hd^{-1} d^{-1}), and live weight is the live weight of the animal (kg hd^{-1}). We obtain maintenance requirements for metabolisable protein for sheep and goats from the Jarrige (1989) equation below:

$$P_m = (2.5 \times \text{live weight}^{.75})/1000 \qquad [26]$$

Where P_M is metabolisable protein required for maintenance (kg hd^{-1} d^{-1}), and live weight is the live weight of the animal (kg hd^{-1}). Based on these two equations above, we obtain reduced form generic equations expressing maintenance requirements for metabolisable protein.

Lactation

We obtain lactation requirements for metabolisable protein for cattle from the following Jarrige (1989) equation:

$$P_l = (C \times Y_m)/1000 \qquad [27]$$

Where P_l is metabolisable protein required for lactation (kg hd^{-1} d^{-1}), and Y_m is the milk yield of the animal (l hd^{-1} d^{-1}) and C=48 for dairy cattle, C=53 for beef cattle, and C=45 for goats. The value of C can be decomposed into the gram of protein per kg milk and a coefficient. An alternative method to calculate metabolizable protein requirements for goats is 1.45g metabolizable protein per gram of milk protein (NRC, 2007). Using the method of C=45 implies that the protein content of milk is 3.1%, CSIRO (2007) suggest using a value of 3.2%.

An alternative equation for calculating metabolizable protein for dairy cattle is $P_l = (Y_m \times (\text{milk true protein}/100))/0.67$ (NRC, 2001 p.68). Y_m is the milk yield of the animal (kg hd^{-1} d^{-1}) .

Lactation requirements for metabolisable protein for sheep are obtained from the following Jarrige (1989) equation:

$$P_l = (1.72 \times \text{protein} \times Y_m)/1000 \qquad [28]$$

Where P_l is metabolisable protein required for lactation (l hd^{-1} d^{-1}), protein is the crude protein content of the milk (g l^{-1}), and Y_m is the milk yield (l hd^{-1} d^{-1}).

Growth

We obtained requirements for growth for cattle from NRC (2000):

$$P_g = ADG \times (268-(29.4 \times (RE/ADG)))/1000 \qquad [29]$$

Where P_g is the net protein required for growth (kg hd^{-1} d^{-1}), ADG is the average daily gain of the animal (kg hd^{-1} d^{-1}), and RE is retained energy (Mcal hd^{-1} d^{-1}), which we define as follows:

$$RE = 0.0635 \times EBW^{.75} \times EBG^{1.097} \qquad [30]$$

Where EBW is empty body weight and EBG is empty body gain. EBW is defined as live weight (kg hd^{-1}) $\times 0.891 \times 96\%$ (conversion factor for shrunk body weight to full body weight), and EBG is defined as ADG (kg hd^{-1} d^{-1}) $\times 0.956$. To convert net protein to metabolizable protein for dairy cattle we use the calculations from NRC (2001 pp. 68-69) and for beef cattle we use the calculations from NRC (2000 p. 116).

Metabolisable protein requirements for growth for goats and sheep are based on the following NRC (2007) equation:

$$P_g = C_g^p \times \text{ADG} \qquad [31]$$

Where P_g is the protein requirement for growth (kg hd^{-1} d^{-1}), C_g^p is the coefficient of growth for protein, and ADG is the animal's average daily gain (kg hd^{-1} d^{-1}). The coefficient of growth takes a value of 0.290 for dairy and indigenous animals, and 0.404 for meat-producing animals. While these coefficients are specific to goats, we use the same procedure for sheep, as there is no direct way to formulate protein requirements for sheep.

3.1.3 Nutrient model evaluation

We compared the net energy requirements and metabolizable protein requirements associated with model-specified levels of milk and meat production and associated milk protein and fat content, average daily gain live weight gain and full body weight with 109 different scenarios taken from the literature (Figs. 6 and 7). Scenarios included all types and classes of livestock. The model has an overall normalized root of the square mean errors (NRMSE) of 15% for net energy requirements and overall NRMSE of 19% for metabolizable protein requirements.

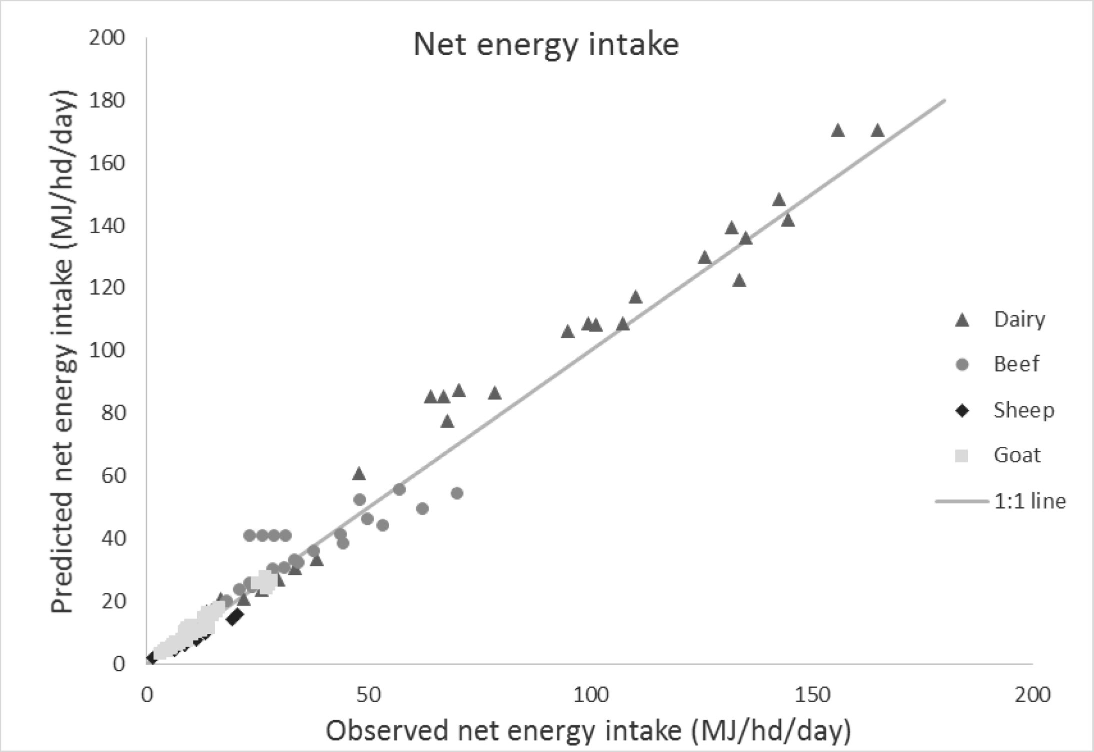
Our model calculates metabolizable protein. To distinguish between crude protein intake and metabolizable protein, crude protein requirements can be obtained by dividing metabolizable protein amounts by a value between 0.64 and 0.80 (NRC, 2000, p. 16). When the literature reported crude protein, we multiplied these values by 0.72 to obtain metabolizable protein.

Some sources, for example NRC (2007), reported energy in terms of metabolizable energy. When this was the case, to obtain net energy requirements we converted 60% of any reported metabolizable energy to net energy. This conversion follows the work of Tolkamp

(2010). AFRC (1993) discuss how the conversion ratio will change depending on livestock weight and diet (and its digestibility); however, because we do not know diet composition before our simulations we use 60%.

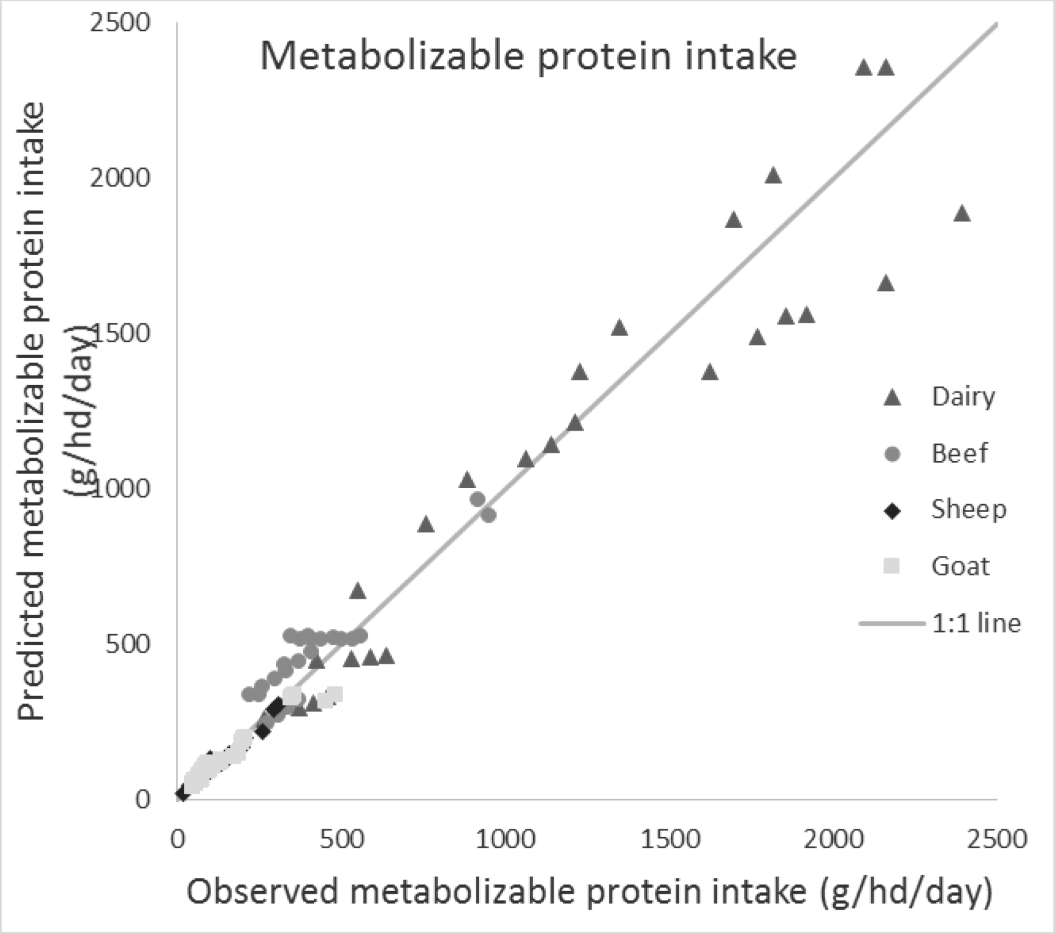
The evaluation procedure used a value of 0.29 in the equations for the metabolisable protein requirement for the growth of goats and sheep.

Figure 6: Relationship between observed net energy intake (x-axis) and predicted net energy intake (y-axis) by DAHBSIM's nutrient module.



Sources: Observed data from NRC guidelines (NRC, 2000; NRC, 2001; NRC, 2007), AFRC guidelines (AFRC, 1993), LIVSIM-modelled outputs (Rufino *et al.*, 2009), a summary of 37 published modeling studies (Tedeschi *et al.*, 2014), and feeding experiment data (Philp *et al.*, 2015).

Figure 7: Relationship between observed metabolizable protein intake (x-axis) and predicted metabolizable protein intake (y-axis) by DAHBSIM's nutrient module.



Sources: Observed data from NRC guidelines (NRC, 2000; NRC, 2001; NRC, 2007), AFRC guidelines (AFRC, 1993), LIVSIM-modelled outputs (Rufino *et al.*, 2009), a summary of 37 published modeling studies (Tedeschi *et al.*, 2014), and feeding experiment data (Philp *et al.*, 2015).

3.2.1 Modeling of livestock herd dynamics and feed management

The nutrient module calculates the nutrient demands for animals and the livestock module determines system-level outcomes given the household objective, household resources, prices, and other external factors. The system-level outcomes are influenced by other components of the whole-farm household model and include livestock production at the animal and farm level, animal demographics over time, feed system balances, labour demands, and nitrate and methane excretion.

3.2.1 Livestock Production

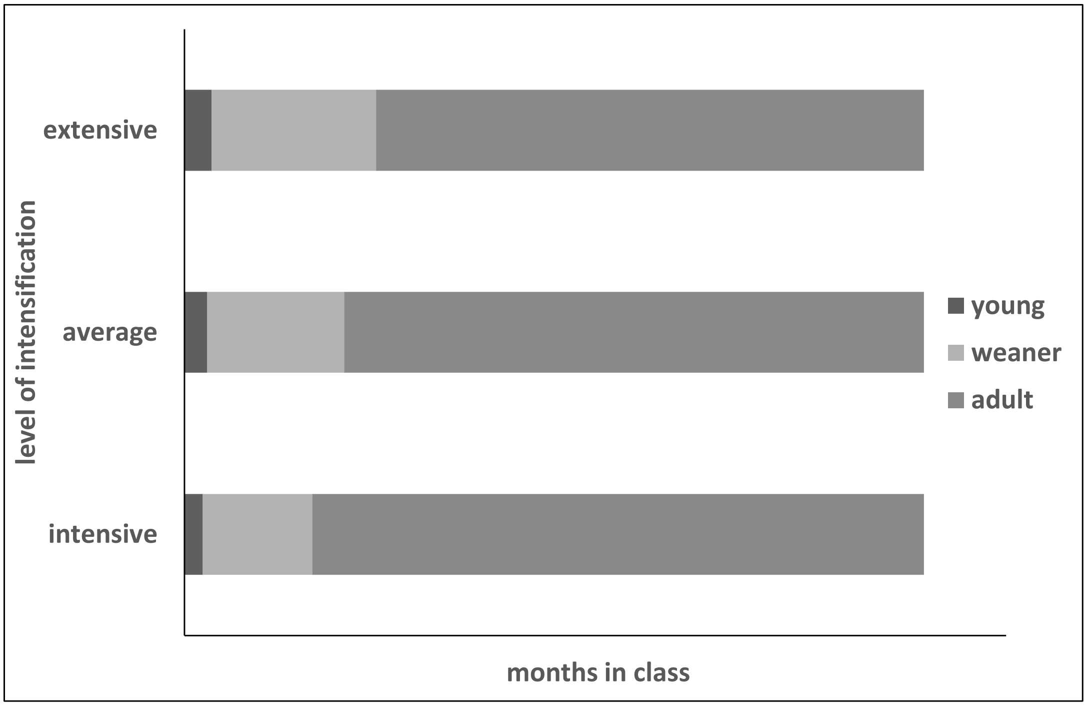
For different levels of intensification, we specify parameters for live weight at maturity for each age class, daily live weight gain, and milk yield. The model then chooses the number of animals for a specific level of intensification. Two factors determine the quantity of meat and milk produced in every month and year. First, the number of animals in that period (a decision variable), multiplied by their respective live weights (associated with its intensification level). Second, the quantity of lactating animals in that period, multiplied by each individual animal's milk yield (associated with its intensification level). The farm module determines how the household allocates production to either sales or consumption.

3.2.2 Herd & flock dynamics

The herd and flock dynamics system used for accounting for demographic transitions of livestock depends jointly on the model's parameters, as well as decision variables (levels of animal sales and purchases, as well as livestock growth rates). Activity levels (the number of animals) for each age class are determined on a monthly time step for all animal types.

The rate at which an animal progresses through the different age classes (months in a class) depends on its growth rate (average daily gain), which is determined by its nutrient requirements and associated feed intake. Animals transition between classes at the same weight across all intensification levels. Each intensification level has a different growth rate and therefore a different time spent in the class to reach the standard weight to transfer class, for example, from a weaner to an adult or from a young animal to a weaner (Fig. 8). Survey data or secondary data sources provide information on the other parameters related to herd demographics, which include the interval between adult breeder births, fertility rate, mortality rate, and newborn sex ratio. Each animal has a uniform life length across the intensification levels; however, because growth rates differ across the intensification systems (based on different diet requirements) the time spent as an adult breeder will be greater in the more intensive system. Thus, lifetime productivity will be higher. Rufino *et al.* (2009) provided the motivation for using this concept of changing feed mixes to explore lifetime productivity.

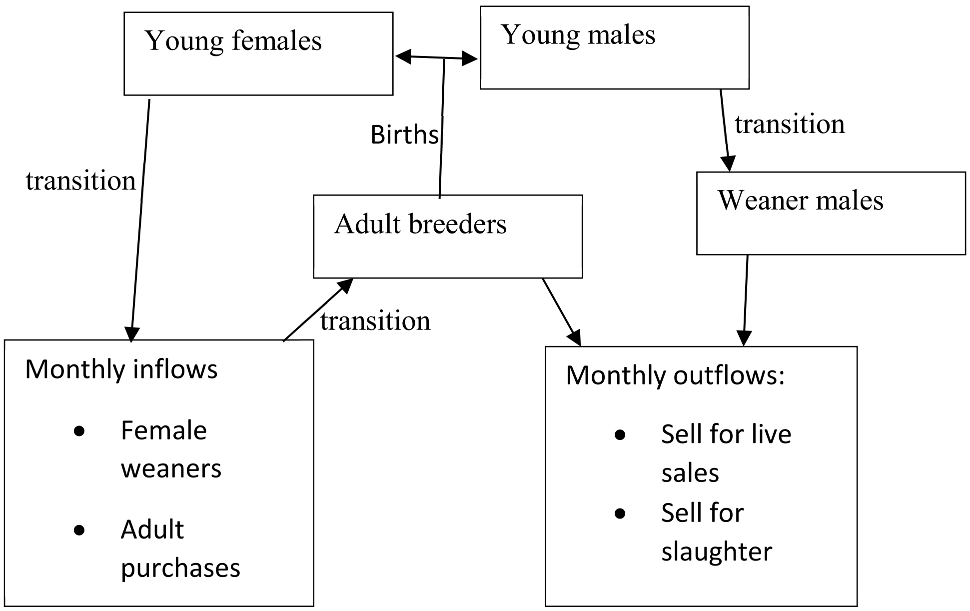
Figure 8: Outline of relationship between animal progression and intensification level.



Notes: Intensified feeding regimes are associated with faster live weight gains, and therefore more rapid transition into the weaner and adult animal categories, thus influencing lifetime productivity.

Figure 9 provides an overview of the livestock demographics. The number of young animals each month equals to the number in the previous month that survived, minus those that became weaners, plus newborns. The growth rate of young animals (average daily live weight gain) determines the rate at which young transfer into the weaner category. We specify the time taken for the animal to grow from its birth weight to its weaner weight, and this determines its average daily gain. Alternatively, we can specify average daily gain or final weights in the class. We calculate the number of newborns based on the number of breeders and breeder fertility rates. Young births occur at the same rate in every month of the model. Young have an equal probability of being male or female. Every month a constant fraction of animals transition to a different age class, because the transition rate is a parameter and the number of animals is a variable. Births occur in every month at the same fertility rate. We monitor an average herd or flock not individual animals. The number of weaners each month equals the number of young in the previous month that reach the transition weight, plus weaners in the previous month that have not transitioned, corrected for the mortality rate, plus net purchases (purchases minus sales) minus home consumption. The number of adults equals the number of adults in the previous period that survived, adjusted for the replacement rate, net purchases, plus the female weaners in the previous period that transition to adult breeders minus home consumption.

Figure 9: Basic overview of monthly livestock dynamics



Notes: Transitions to a different category are model parameters, and sales and purchases variables are decision variables.

We fix the number of reproductive animals to a percentage of adult breeders. For biological reasons, the household cannot buy, sell, or consume young animals. The time between breeder births is not equal to one year, therefore the month each year that pregnancy occurs changes.

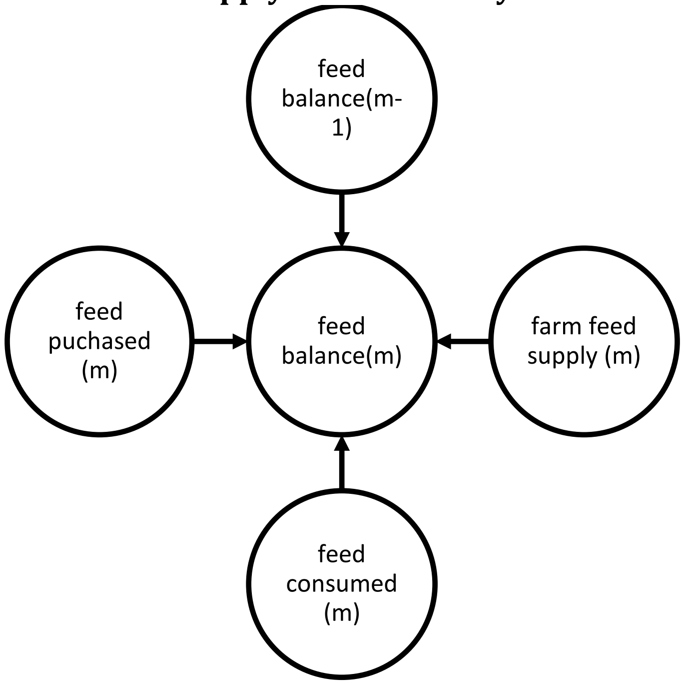
3.2.3 Feeding system

The model requires data on the dry matter content, net and gross energy, and crude (or metabolizable) protein content of potential feed sources.

We specify a condition that amount of protein consumed by an animal must be at least equal to the animal's monthly protein requirement (kg crude protein/month), as defined in the nutrient module. We separately specify a condition that amount of energy consumed by an animal must be at least equal to the animal's monthly net energy requirement (MJ/month), as defined in the nutrient module. Then we calculate the amount of protein (kg crude protein/month) and energy (MJ/month) an animal could consume, based on the feed types consumed and their nutrient properties. We also calculate total feed consumed across all feed types by each animal (kg/month). We also calculate the actual percentage protein in the diet, based on actual modelled consumption for each animal type and age (kg/month). We also calculate the amount of a specific feed consumed by multiplying the amount of a feed consumed per animal by the number of animals present.

We include a feed balance equation to ensure feed supply and demand are consistent (Fig. 6). This balance equation holds for every specific feed in every month and year. The feed stockpile in the previous month plus feed available for livestock to consume from on-farm feed production, plus feed purchases equals total feed consumed by livestock plus the stockpile. The stockpile is the amount of feed not consumed and the household can carried this stockpile over to the following month, correcting for a specific amount lost (spoiled).

Figure 10: Schematic of feed supply and demand system.



Notes: Feed balance in every month (m) equals the previous month's balance (m-1) plus feed supplied from farm in m + feed purchased in m minus feed consumed by livestock in m.

The livestock model requires the specification of different feed baskets for different animals of different intensification levels. The model user specifies these feed baskets and the feed baskets should reflect local contexts, for example pasture or concentrate availability. The feedactivity_module.gms calculates the physical quantity of different feeds required to meet a level of total energy and total protein. This level of total energy and total protein intake has an associated level of meat and milk production, based on the nutrient module calculations. The feedactivity_module.gms module specifies local context conditions, for example, the minimum or maximum physical quantities of a feed to consume. The productivitycalculationmodule.gms module allocates the total energy and protein available from the feed basket to different beneficial uses of energy and protein. For example, energy or protein required to growth, maintenance or reproduce are all beneficial uses. The productivitycalculationmodule.gms uses a cross entropy method to allocate the energy and protein to the different sub categories of energy and protein so that the actual intake of the nutrients matches as close as possible to the required amounts (based on the level of pre-determined productivity).

3.2.4 Labour use in livestock production

The model calculates total labour required by the livestock enterprise by summing up labour requirements for each animal unit in days/month. The model specifies labour requirements per animal based on the type and age of the animal, the season, and the intensity level. These labour requirements are based on Grandin *et al.* (1993).

3.3 Modeling the environmental impact of livestock

3.3.1 Manure and Nitrate excretion from livestock

We calculate N excretion endogenously in our model based on the characteristics of the diet, which determine total N intake as well as N produced endogenously, and N retained by the animal, which is related to the animal's protein requirement. This method has proven to be highly accurate when compared to other ruminant models, especially the Large and Small Ruminant Nutrition Systems (LRNS and SRNS; http://nutritionmodels.com/models.html). Under this system, N excretion for large ruminants is estimated as

$$\text{N excretion} = \text{Feed N intake} + \text{Microbial N Production} - \text{N Retained} \qquad [32]$$

Where N excretion is N excretion per animal (kg d^{-1}), Feed N intake is crude protein intake (kg d^{-1}) divided by 6.25 to convert from kg of dietary protein to kg of dietary N (IPCC, 2006). N retained is the metabolizable protein intake/requirement of the animal divided by 6.25, and Microbial N Production is N produced in the rumen from bacteria (kg d^{-1}). This value is based on the following equation from AFRC (1993):

$$MCP = C_a * FME / 1000 \qquad [33]$$

Where MCP is microbial crude protein production (kg), C_a is an animal specific coefficient, taking the value of 9,10, and 11 for mature animals, growing sheep and cattle, and lactating ewes or dairy cows respectively (AFRC, 1993), and FME is fermentable metabolisable energy (MJ) in the diet. Since we are lacking data on the FME content of feeds, we simply assume a standard content of 44% of the gross energy content of the feed. After obtaining MCP, we can then convert to Microbial N Production by dividing by 6.25, and 1000 to convert from g to kg.

For small ruminants:

$$\text{N excretion} = (\text{Feed N intake} + \text{Microbial N Production}) \times 0.20 \qquad [34]$$

Where Feed N Intake and Microbial N Production are both calculated as described above.

We convert nitrogen excreted into nitrate excreted, based on nitrate nitrogen being 22.6% nitrate (Rasby *et al.*, 1988); for each kg of nitrogen there is 4.43 kg of nitrate. The model uses total nitrate excreted as inputs into the crop and biophysical modules, which then determine soil nitrate balances, and subsequently choices for fertilizer and crop residue usage.

Total volume of manure produced from cattle is calculated based on empirical equations estimated by Nennich *et al.* (2005). The model expresses manure production per animal as a function of the dry matter intake (DMI) of the animal. The following equation estimates manure production:

$$DM_e = DMI \times 0.3256 + 0.8 \qquad [35]$$

Where DM_e is dry matter excretion of manure (kg hd d^{-1}).

3.3.2 Methane emissions

We follow the IPCC (2006) methodology to calculate methane production from livestock, and this occurs post-optimization:

$$CH_4 = (1/55.65) \times GEI \times Y_m / 100 \qquad [36]$$

Where CH_4 is the methane produced per animal (kg CH_4 hd^{-1} d^{-1}) and Y_m is the methane conversion factor. We estimate Y_m using the following equation, from Opio et al. (2013).

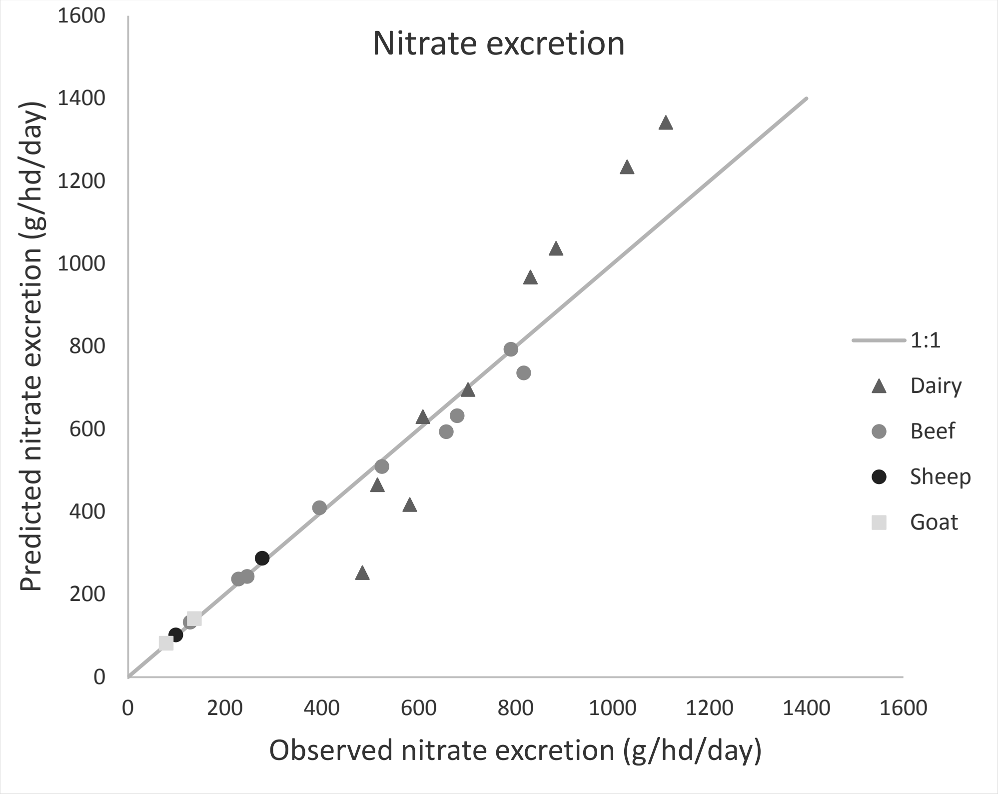
$$Y_m = 9.75 - 0.05 \times DE \qquad [37]$$

Where DE is digestible energy of the diet (%). The model sums methane production across the animals into a monthly value.

3.3.3 Nitrogen excretion evaluation

We compare estimates of nitrogen excretion to the RNS Models (http://nutritionmodels.com/models.html). The LRNS (Large Ruminant Nutrition System) and SRNS (Small Ruminant Nutrition System) provide an overview of N cycling through the animal. This includes intake from feed, production within the animal, and excretion of urinary and faecal N. We compare our model's estimated nitrogen output for large and small ruminants to the estimated excretion for an animal of equivalent characteristics, and consuming a closely matched diet, with the LRNS and SRNS respectively.

Figure 11: Relationship between observed nitrate excretion (x-axis) and predicted nitrate excretion (y-axis) by DAHBSIM's nutrient module.



Source: Observed data from Ruminant Nutrition System estimates.

The model has an overall normalized root of the square mean errors (NRMSE) of 10% for nitrate excretion.

Table 11: Livestock Module Equations

Name	Equation
E_MILKPRODUCTION	$\text{v\_prodQmilk}_{\text{hh,adultf,typec,inte,y,m}} = \text{p\_milkprod}_{\text{adultf,typec,inten}} \times \text{v\_aactLev}_{\text{hh,adultf,typec,inten,y,m}}$
E_MEATPRODUCTION	$\text{v\_prodQmeat}_{\text{hh,aaact,typec,inten,y,m}} = \text{p\_salesw}_{\text{aaact,typec,inten}} \times \text{v\_aactLevslaughter}_{\text{hh,aaact,typec,inten,y,m}}$
E_LEGUMECONSUMPTION1	$\sum_{\text{legumeresidues}} \text{p\_nutcontent}_{\text{'metabolisableprotein',legumeresidues}} \times \text{v\_fdcons}_{\text{hh,aaact,typec,inten,y,m,legumeresidues}} > \text{p\_fdnutrients}_{\text{hh,aaact,typec,inten,y,m,'proteinconsumptionlegumes'}}$
E_LEGUMECONSUMPTION2	$\sum_{\text{legumeresidues}} \text{p\_nutcontent}_{\text{'netenergy',legumeresidues}} \times \text{v\_fdcons}_{\text{hh,aaact,typec,inten,y,m,legumeresidues}} = \text{p\_fdnutrients}_{\text{hh,aaact,typec,inten,y,m,'energyconsumptionlegumes'}}$
E_MAXPROTEIN	$\sum_{\text{feedc}} \text{p\_nutcontent}_{\text{'metabolisableprotein',feedc}} \times \text{v\_fdcons}_{\text{hh,aaact,typec,inten,y,m,feedc}} < \text{p\_fdnutrients}_{\text{hh,aaact,typec,inten,y,m,'proteinconsumption''}}$
E_MAXENERGY	$\sum_{\text{feedc}} \text{p\_nutcontent}_{\text{'netenergy',feedc}} \times \text{v\_fdcons}_{\text{hh,aaact,typec,inten,y,m,feedc}} < \text{p\_fdnutrients}_{\text{hh,aaact,typec,inten,y,m,'energyconsumption''}}$
E_FEEDING	$\text{v\_aactLev}_{\text{hh,aaact,typec,inten,y,m}} \times \text{p\_feedreq}_{\text{nut,aaact,typec,inten,m}} < \text{v\_totalnutdem}_{\text{hh,aaact,typec,inten,nut,y,m}}$
E_FEEDBAL	$\text{v\_feedbalance}_{\text{hh,feedc,y,m-1}} \times \text{p\_transferrate} + \text{v\_feedbalance}_{\text{hh,feedc,y-1,'m12'}} \times \text{p\_transferrate} + \text{p\_s}_{\text{hh,feedc}} + \text{v\_residuesfeedm}_{\text{hh,feedc,y,m}} + \text{v\_residuesbuy}_{\text{hh,feedc,y,m}} = \text{v\_feedcons}_{\text{hh,feedc,y,m}} + \text{v\_feedbalance}_{\text{hh,feedc,y,m}}$
E_ENERGYCONSUMED	$\text{v\_energyconsumed}_{\text{hh,aaact,inten,y,m,typec}} = \sum_{\text{feedc}} \text{v\_fdcons}_{\text{hh,aaact,typec,inten,y,m,feedc}} \times \text{p\_nutcontent}_{\text{'netenergy',feedc}}$
E_PROTEINCONSUMED	$\text{v\_proteinconsumed}_{\text{hh,aaact,inten,y,m,typec}} = \sum_{\text{feedc}} \text{v\_fdcons}_{\text{hh,aaact,typec,inten,y,m,feedc}} \times \text{p\_nutcontent}_{\text{'metabolisableprotein',feedc,}}$
E_CRPROTEINCONSUMED	$\text{v\_cproteinconsumed}_{\text{hh,aaact,inten,y,m,typec}} = \sum_{\text{feedc}} \text{v\_fdcons}_{\text{hh,aaact,typec,inten,y,m,feedc}} \times \text{p\_nutcontent}_{\text{'crudeprotein',feedc}}$

E_GROSSENERGYCONSUMED	$\text{v\_grossenergyconsumed}_{\text{hh,aaact,typec,inten,y,m}} = \sum_{\text{feedc}} (\text{v\_fdcons}_{\text{hh,aaact,typec,inten,y,m,feedc}} \times \text{p\_nutcontent}_{\text{’grossenergy’,feedc}})/30$
E_FEEDCONS	$\text{v\_feedcons}_{\text{hh,feedc,y,m}} = \sum_{\text{aaact,typec,inten}} (\text{v\_aactLev}_{\text{hh,aaact,typec,inten,y,m}} \times \text{v\_fdcons}_{\text{hh,aaact,typec,inten,y,m,feedc}} / \text{p\_nutcontent}_{\text{’drymatter’,feedc}})$
E_FEEDCONSA	$\text{v\_DMI}_{\text{hh,aaact,typec,inten,y,m}} = \sum_{\text{feedc}} \text{v\_fdcons}_{\text{hh,aaact,typec,inten,y,m,feedc}}$
E_LABORAA	$\sum_{\text{aaact,typec,inten}} \text{v\_aactLev}_{\text{hh,aaact,typec,inten,y,m}} \times \text{p\_labour}_{\text{m,inten,aaact,typec}} = \text{v\_totallabourdemlive}_{\text{hh,y,m}}$
E_NINTAKE	$\text{v\_nintake}_{\text{hh,aaact,typec,inten,y,m}} = \sum_{\text{feedc}} (\text{v\_fdcons}_{\text{hh,aaact,typec,inten,y,m,feedc}} \times \text{p\_nutcontent}_{\text{’grossenergy’,feedc’}}/30) \times \text{v\_dietprotein}_{\text{hh,aaact,typec,inten,y,m}} \times (1/6.25) \times (1/18.45)$
E_NEC[a]	$\text{v\_nec}_{\text{hh,aaact,’dairy’,inten,y,m}} = \text{p\_nce}_{\text{aaact,’dairy’,inten}} \times (\text{v\_nintake}_{\text{hh,aaact,’dairy’,inten,y,m}} + 0.44 \times \text{v\_grossenergyconsumed}_{\text{hh,aaact,’dairy’,inten,y,m}} \times \text{p\_ca}_{\text{aaact,sr,inten}}/(1000*6.25))$
E_NESR[a]	$\text{v\_nesr}_{\text{hh,aaact,sr,inten,y,m}} = \text{p\_nce}_{\text{aaact,sr,inten}} \times (\text{v\_nintake}_{\text{hh,aaact,sr,inten,y,m}} + 0.44*\text{v\_grossenergyconsumed}_{\text{hh,aaact,’dairy’,inten,y,m}} \times \text{p\_ca}_{\text{aact,sr,inten}}) \times (1000 \times 6.25)$
E_NEC2	$\text{v\_nec}_{\text{hh,aaact,’beef”,inten,y,m}} = \text{p\_nce}_{\text{aaact,’beef’,inten}} \times (\text{v\_nintake}_{\text{hh,aaact,’beef’,inten,y,m}} + (.44 \times \text{v\_grossenergyconsumed}_{\text{hh,aaact,’dairy’,inten,y,m}} \times \text{p\_ca}_{\text{aaact,’beef’,inten}})/1000 \times 6.25 - \text{p\_nr}_{\text{hh,aaact,’beef’,inten,y,m}})$
E_NEC3	$\text{v\_nec}_{\text{hh,aaact,sr,inten,y,m}} = 0$
E_NESR2	$\text{v\_nesr}_{\text{hh,aaact,cattle,inten,y,m}} = 0$
E_NLIVEM	$\text{v\_nitlivem}_{\text{hh,y,m}} = \sum_{\text{aaact,typec,inten}} ((30*\text{v\_nec}_{\text{hh,aaact,typec,inten,y,m}} \times \text{v\_aactLev}_{\text{hh,aaact,typec,inten,y,m}}) +$

	$v\_nesr_{hh,aaact,typec,inten,y,m} \times v\_aactLev_{hh,aaact,typec,inten,y,m}))$
E_NITLIVE	$v\_nitlive_{hh,y} = p\_nconv \times \sum_{m} v\_nitlivem_{hh,y,m}$
E_DIETPROTEIN	$v\_dietprotein_{hh,aaact,inten,y,m,typec} \times v\_DMI_{hh,aaact,typec,inten,y,m} = v\_cproteinconsumed_{hh,aaact,inten,y,m,typec}$

Notes: [a] These equations are specific to cattle (beef and dairy) and small ruminants respectively.

Table 12: Livestock Module Variables

Variable	Units	Description
v_aactLev	head	Livestock numbers
v_aactLevsell	head	Livestock sales
v_aactLevbuy	head	Livestock purchases
v_aactLevslaughter	head	Livestock slaughtered
v_aactselladult	head	Adults sold not for old age
v_totallabourdemlive	person-days month^{-1}	Total labour demand per month for the entire herd
v_prodQmilk	l month^{-1} animal^{-1}	Milk produced
v_prodQmeat	kg month^{-1} animal^{-1}	Meat produced
v_feedbalance	kg	Feed carried over to next month
v_grossenergyconsumed	MJ GE hd^{-1} month^{-1}	Gross energy consumed per animal per month
v_eneryconsumed	MJ^{-1} hd^{-1} month^{-1}	Net energy consumption per animal in MJ
v_proteinconsumed	kg MP animal^{-1} month^{-1}	Metabolisable protein consumption per animal

v_cproteinconsumed	kg CP hd^{-1} month^{-1}	Crude protein consumption per animal
v_fdcons	kg hd^{-1} month^{-1}	Amount of a specific feed consumed per animal per month
v_feedcons	kg month^{-1}	Total kg of a feed consumed per month
v_totalnutdem	MJ net energy hd^{-1} month^{-1} and kg MP hd^{-1} month^{-1}	Total nutrient demand of net energy or metabolisable protein per animal per month
v_totalnutdemt	MJ net energy month^{-1} and kg MP hd^{-1} month^{-1}	Total nutrient demand of net energy or metabolisable protein for the entire herd per month
v_DMI	kg DMI hd^{-1} month^{-1}	DMI per animal per month
v_residuesbuy	kg m^{-1}	kg of residue consumed from market purchases
v_nintake	kg hd^{-1} month^{-1}	Total nitrogen intake
v_ManureDMc	kg month^{-1}	Manure produced from cattle
v_ManureDMsr	kg month^{-1}	Manure produced from small ruminants
v_ManureDM	kg month^{-1}	Manure produced from all animals
v_dietprotein	kg CP kg DMI^{-1} hd^{-1} month^{-1}	Percentage crude protein in diet
v_nesr	kg N hd^{-1} month^{-1}	Total nitrogen excreted per head of small ruminant per month
v_nec	kg N hd^{-1} month^{-1}	Total nitrogen excreted per head of cattle per month
v_nitlivem	kg N month^{-1}	Nitrogen excreted from all livestock per month
v_nitlive	kg NO$_3$ year^{-1}	Total yearly nitrate excretion

Table 13: Livestock Module Parameters

Parameter Name	Units	Description	Data Source
v0_aactLev	head	initial animal activity levels	livestockdata.xlsx
p_nutcontent	MJ, %	net and gross energy content per kg feed in MJ, % metabolisable and % crude protein of feed, %dry matter of feed	livestockdata.xlsx
p_rateFert	ratio	fertility rate, young born per female adult per month	livestockdata.xlsx
p_mreprodf	head	number of males	livestockdata.xlsx
p_sexratio	ratio	birth ratio male and female	livestockdata.xlsx
p_survivalrate	%	survival rate, % of livestock that survive each year	livestockdata.xlsx
p_grad	months	months in class, used for livestock demographics	livestockdata.xlsx
p_startvalue	head	starting livestock numbers	livestockdata.xlsx
p_s	kg	starting feed balance	livestockdata.xlsx
p_labour	person-days head^{-1} month^{-1}	labour demand for each activity in each season and intensification level across households	livestockdata.xlsx

4. THE MODELLING OF SOCIO-ECONOMICS IN DAHBSIM

Given the description of biophysical components of both crops and livestock, that were discussed in the previous sections – now we turn towards the description of how DAHBSIM models the economic decisions of farm production and consumption, that must take into account socio-economic and biophysical constraints. The key modules that capture the farm management and household consumption decisions are described in the following sub-sections.

4.1 Farm Management Module

The farm module depicts the interactions between cropping and livestock activities, as well as the linkages between on-farm and off-farm labor. The main equations making part of this module are:

- land balance: land allocation between cropland and grassland
- labor balance: allocation of family labor
- supply balance: for each product and year, production should equal use (divided into self-consumption, on-farm seed use, on-farm feed use and market sales).
- farm income: from market sales

We differentiate buying and selling prices.

Agricultural Land Balance

For each household type, total cropland plus grassland cannot exceed total agricultural land.

$$\text{v\_cropLand}_{hh,s,n} + \text{v\_grassLand}_{hh,s,n} \leq \text{p\_agriLand}_{hh,s,n}$$

[38]

Where

v_grassLand is the total grassland area

v_agriLand is the total agricultural area

Labor Balance

For each household type, total family labor used per month cannot exceed total labor available:

$$\text{v\_cropFlabh}_{h,l,n,m} + \text{v\_livesFlab}_{hh,l,n,m} \leq \text{p\_totalFlab}_{hh,l,n,m}$$

[39]

Where

l = index for labor type

v_cropFlab = family labor used for cropping activities

v_livesFlab = family labor used for livestock activities

Supply Balance

For each product, total production can be used on-farm (as seed or feed); or can be self-consumed; or can be sold in the market.

$$\text{v\_prodQuant}_{hh,j,n} \leq \text{v\_seedOnfarm}_{hh,j,n} + \text{v\_feedOnfarm}_{hh,j,n} + \text{v\_selfCons}_{hh,j,n} + \text{v\_markSales}_{hh,j,n}$$

[40]

Where

j is the index for products

v_prodQuant is the production quantity

v_seedOnfarm is the on-farm seed use

v_feedOnfarm is the on-farm feed use

v_selfCons is household self-consumption

Table 14: Farm Module Equations

Label	Definition
E_ALANDBALANCE	$\text{v\_cropLand}_{hh,s,y} + \text{v\_grasLand}_{hh,s,y} < \text{p\_landEnd}_{hh,s,\text{'agriland'}}$
E_CLANDBALANCE	$\text{v\_cropLand}_{hh,s,y} < \text{p\_landEnd}_{hh,s,\text{'cropland'}}$
E_LABORBALANCE	$\text{v\_farmLabor}_{hh,y,m} = \text{v\_cropLabor}_{hh,y,m} + \text{v\_totallabourdemlive}_{hh,y,m}$
E_SEEDBALANCE	$\text{v\_seedquant}_{hh,caen,y} = \text{v\_seedonfarm}_{hh,caen,y-1}\ (\text{if } y > 1) + \text{v0\_seedUse}_{hh,caen,\text{'seedOnFarm'}}\ (\text{if } y = 1) + \text{v\_seedPurch}_{hh,caen}$
E_SEEDBALANCE_1	$\text{v\_seedQuant}_{hh,caen,y} = \text{v0\_seedUse}_{hh,caen,\text{'seedOnFarm'}} \times \text{v0\_seedUse}_{hh,caen,\text{'seedTotal'}}$
E_NITRBAL	$\sum_{caen} \text{v\_inputUse}_{hh,caen,\text{'nitr'},y} = \text{v\_Nfert}_{hh} + \sum_{cken} (\text{p\_MSres} \times \text{p\_effr}_{cken} \times \text{v\_residuesmulch}_{hh,cken,y-1}) + \text{v\_nitlive}_{hh,y-1}$
E_SBALANCE_CJ	$\text{v\_prodQuant}_{hh,cjen,y} = \sum_{\text{a\_j(caen,cjen)}} \text{v\_seedOnfarm}_{hh,caen,y} + \text{v\_feedOnfarm}_{hh,cjen,y} + \text{v\_selfCons}_{hh,cjen,y} + \text{v\_markSales}_{hh,cjen,y}$
E_SBALANCE_AK	$\text{v\_prodQuant}_{hh,ak,y} = \text{v\_selfCons}_{hh,ak,y} + \text{v\_markSales}_{hh,ak,y}$
E_DBALANCE_GD	$\text{v\_hconQuant}_{hh,gd,y} = \sum_{\text{o\_g(cj,gd)}} \text{v\_selfCons}_{hh,cj,y} + \text{v\_markPurch}_{hh,gd,y} + \sum_{\text{o\_g(ak,gd)}} \text{v\_selfCons}_{hh,ak,y}$
E_INCOME_FARM	$\text{v\_farmIncome}_{hh,y} = \text{v\_cactIncome}_{hh,y} + \text{v\_aactIncome}_{hh,y}$
E_INCOME_CA	$\text{v\_cactIncome}_{hh,y} = \text{v\_cactSrev}_{hh,y} - \text{v\_cactCost}_{hh,y}$
E_COST_CA	$\text{v\_cactCost}_{hh,y} = \sum_{caen,inpv} \text{v\_inputUse}_{hh,caen,inpv,y} + \sum_{caen,} \text{v\_seedPurch}_{hh,caen,y} \times \text{p\_seedbuyPri}_{hh,caen} +$ $\sum_{caen,} \text{v\_inputUse}_{hh,caen,\text{'nitr'},y} \times \text{p\_buyPrice}_{hh,\text{'nitr'}} + \text{v\_Nfert}_{hh} \times \text{p\_buyPrice}_{hh,\text{'nitr'}}$
E_SREV_CA	$\text{v\_cactSrev}_{hh,y} = \sum_{cj} \text{v\_markSales}_{hh,cj,y} \times \text{p\_selPrice}_{hh,cj}$
E_INCOME_AA	$\text{v\_aactIncome}_{hh,y} = \text{v\_aactSrev}_{hh,y} - \text{v\_aactCost}_{hh,y}$

E_COST_AA	$\text{v\_aactCost}_{hh,y} = \sum_{aj,,m} (\text{v\_residuesbuy}_{hh,aj,y,m} \times \text{p\_buyPrice}_{hh,ai}) + \sum_{aj} (\text{v\_animalpurch}_{hh,aj,y} \times \text{p\_buyPrice}_{hh,aj})$
E_SREV_AA	$\text{v\_aactSrev}_{hh,y} = \sum_{\text{a\_j}} (\text{v\_MarkSales}_{hh,aj,y} \times \text{p\_selPrice}_{hh,aj}) + \sum_{ak} (\text{v\_markSales}_{hh,ak,y} \times \text{p\_selPrice}_{hh,ak}) + \sum_{cken} (\text{v\_residuessell}_{hh,cken,y} \times \text{p\_selPrice}_{hh,cken})$
E_PRODQUANTMAPMEAT	$\text{v\_prodQuant}_{hh,akmeat,y} = \sum_{\text{a\_k(aa,akmeat)}} \sum_{\text{m,inten,aaact\_act(aa,aaact,typec)}} \text{v\_prodQmeat}_{hh,aaact,typec,inten,y,m}$
E_PRODQUANTMAPMILK	$\text{v\_prodQuant}_{hh,akmilk,y} = \sum_{\text{a\_k(aa,akmilk)}} \sum_{\text{m,inten,aaact\_act(aa,aaact,typec)}} \text{v\_prodQmilk}_{hh,aaact,typec,inten,y,m}$
E_ANIMALPURCHMAP	$\text{v\_animalpurch}_{hh,aj,y} = \sum_{\text{a\_j(aa,aj)}} \sum_{\text{m,inten,aaact\_act(aa,aaact,typec)}} \text{v\_aactLevbuy}_{hh,aaact,typec,inten,y,m} \times \text{p\_salesw}_{aaact,typec,inten}$
E_ANIMALSALEMAP	$\text{v\_markSales}_{hh,aj,y} = \sum_{\text{a\_j(aa,aj)}} \sum_{\text{m,inten,aaact\_act(aa,aaact,typec)}} \text{v\_aactLevsell}_{hh,aaact,typec,inten,y,m} \times \text{p\_salesw}_{aaact,typec,inten}$

Table 15: Farm Module Parameters

Parameter Name	Units	Description	Source File
p_farmData; seed on farm	kg ha^{-1}	Farm baseyear data	farmdata
p_farmData; Seed exchange	kg ha^{-1}	Seed use coming from exchange	farmdata
p_farmData; Seed purchase	kg ha^{-1}	Seed use coming from purchases (total)	farmdata
p_farmData; Seed purchase local	kg ha^{-1}	Seed use coming from purchases (local)	farmdata
p_farmData; Seed purchase improved	kg ha^{-1}	Seed use coming from purchases (improved)	farmdata
p_farmData; Total Seed	kg ha^{-1}	Total seed use	farmdata
p_landEnd	ha	initial land endowment by soil type	farmdata
p_waterEnd[a]	th.m3	water endowment	-
v0_agriLand[a]	ha	agricultural land by soil type	-
v0_grasland[a]	ha	grass land by soil type	-
v0_irriLand[a]	ha	irrigable land by soil type	-
v0_laborEnd	person-days $month^{-1}$	labor endowment by month	Household data
p_outData; self consumption	kg	Self consumption of a specific crop	farmdata
p_outData; market sales	kg	Market sales of a specific crop	farmdata
p_outData; straw yield	kg	Straw yield of a specific crop	farmdata
p_outData; straw	kg	Straw produced of a specific crop	farmdata
p_outData; feed	kg	Feed use by a specific crop	farmdata

p_outData; residue	kg	Residue use by a specific crop	farmdata
p_cpriData	nc ton^{-1}	price data for cropping activities	pricedata
p_apriData	nc kg^{-1}	price data for livestock activities	pricedata
p_spriData	nc kg^{-1}	seed price data	pricedata
p_hLabWage	nc d^{-1}	wage rate for hired labor	pricedata
p_fLabWage	nc d^{-1}	reservation wage for family labor	pricedata
p_selPrice	nc ton^{-1}	crop selling price (nc/ton)	pricedata
p_buyPrice	nc ton^{-1}	crop buying price (nc/ton)	pricedata
p_seedbuypri	nc kg^{-1}	seed buying price	pricedata

Note: Model has these data sources built into the structure, but they are not being used for the present case study. ‘nc’ refers to the native currency of the specific model case study

Table 16: Farm module variables

Variable Name	Units	Description
v_farmIncome	nc	farm Income
v_cactIncome	nc	income from cropping activities
v_aactIncome	nc	income from livestock activities
v_farmLabor	person-days	farm labor
v_cactSrev	nc	sales revenue from crops
v_aactCost	nc	direct crop costs excluding labor
v_aactSrev	nc	sales revenue from livestock
v_aactCost	nc	direct livestock costs excluding labor
v_seedOnfarm	kg	on farm seed use
v_seedPurch	kg	input purchases
v_nitrQuant	kg	total fertilizer use quantity
v_nitrChem	kg	chemical fertilizer
v_nitrOrgan	kg	organic fertilizer
v_Nfert	kg	chemical fertilizer
v_farmLabor	person-days	farm labor
v_laborUse	nc	labor use
v_inputCost	nc	production costs
v_prodQuant	t	production quantity
v_feedOnfarm	t	on farm feed use
v_selfCons	t	self-consumption quantity
v_markSales	t	market sales quantity (tons)
v_markPurch	t	market purchases quantity (tons)
v_hconQuant	t	household consumption quantity
v_animalpurch	kg	weight of all live animals purchased
v_feedendval	nc	value of total feed in end period
v_liveend	kg	weight of all live animals in final period
v_liveendval	nc	value of all live animals in final period

v_seedStore	kg	storage of grain into stock of feed
v_seedWdraw	kg	withdrawal of grain from stock of seed
v_seedStock	kg	quantity of seed in stock

Note: ‘nc’ refers to the native currency of the specific model case study

4.2 Household Economic Decision Module

The household module describes the interactions between household decisions, in particular the simultaneous decisions on production, consumption and time allocation.

The main equations in this module are:

- demand balance: total consumption by product and year should equal self-consumption plus market purchases
- time allocation balance: time allocation between on-farm labor and off-farm labor
- household income: from market sales plus off-farm labor

Demand balance

For each product, total consumption cannot exceed self-consumption plus market purchases.

$$\text{v\_consQuant}_{\text{hh,j,n}} \leq \text{v\_selfCons}_{\text{hh,j,n}} + \text{v\_markPurch}_{\text{hh,j,n}} \qquad [41]$$

Where

v_consQuant is the consumption quantity (kg yr^{-1})

v_selfCons is the self-consumption quantity (kg yr^{-1})

v_markPurch is the market purchases (kg yr^{-1})

Table 17: Household Module Equations

Name	Definition
E_TIMEBALANCE	$\text{v\_farmLabor}_{hh,y,m} < \text{p\_workTimeMax}_{hh,m} \times \text{v0\_hholdSize}_{hh} + \text{v\_hiredLabor}_{hh,y,m}$
E_MAXHIREDLABOR	$\text{v\_hiredLabor}_{hh,y,m} < \text{p\_workTimeMax}_{hh,m} \times \text{v0\_hholdSize}_{hh}$
E_INCOMEOFF	$\text{v\_offFarmInc}_{hh,y} = \text{v0\_offFarmInc}_{hh}$
E_INCOMETOT	$\text{v\_fullIncome}_{hh,y} = \text{v\_FarmIncome}_{hh,y} = \text{v\_offFarmInc}_{hh,y} - \text{sum(m,v\_hiredLabor(hh,y,m)*p\_buyPrice(hh,'labor'))}$
E_HHCON	$\text{v\_hconQuant}_{hh,gd,y} = \text{p\_gamma}_{hh,gd} + (\text{p\_beta}_{hh,gd} \times \text{v\_fullincome}_{hh,y} - \text{sum(gd2, p\_gamma}_{hh,gd2} \times \text{p\_goodprice}_{hh,gd2})/\text{p\_goodprice}_{hh,gd}$
E_CASH	$\text{v\_fullincome}_{hh,y} > \text{sum(gd,p\_goodprice}_{hh,gd} \times \text{v\_markPurch}_{hh,gd,y}$

Table 18: Household Module Parameters

Parameter Name	Units	Description
p_hholdData; hhsize	number of people	Number of people in the household
p_hholdData; farm size	ha	Farm size
p_hholdData; labor_family	person-days	Labor availability from family
p_hholdData; labor hired	person-days	Hired labor availability
p_hholdData; labor community	person-days	Community labor availability
p_hholdData; labor total	person-days	Total labor availability
p_hholdData; aa bovine	head	Quantity of bovine animals
p_hholdData; aa chicken	head	Quantity of chickens
p_hholdData; aa goatsheep	head	Quantity of goats & sheep
p_hholdData; aa pig	head	Quantity of pigs
p_hholdData; aa other	head	Quantity of other animals
p_hholdData; hhcon_nonfood	nc yr^{-1}	Household non-food consumption
p_hholdData; inc_offfarm	nc yr^{-1}	Off farm income
p_hholdData; hh nbr	number	Number of households of this type in survey
p_consoData	kg $person^{-1}$ $year^{-1}$	Household consumption data
p_hholdNbr	Number of households	household number
p_gpridata	nc	goods prices
p_workTimeMax	days	maximum number of working days per month
p_elasIncome	-	Income elasticity
p_Frisch	-	Frisch parameter
p_goodPrice	nc kg^{-1}	Price of consumption goods
p_beta	-	Les parameter beta
p_gamma	-	Les parameter gamma
v0_hholdSize	number of people	Household size (number of members of household)
v0_hconNonfood	nc	non food consumption

v0_offFarmInc	nc	off-farm income
v0_selfcons	kg year^{-1}	self-consumption data
v0_hconquant	kg year^{-1}	Quantity consumed
v0_farmIncome	nc	farm income
v0_ConsShare	-	Budget share for LES

Note: ‘nc’ refers to the native currency of the specific model case study

Table 19: Household Module Variables

Variable Name	Units	Description
v_offFarmInc	nc	off-farm income
v_fullIncome	nc	full income
v_offFarmInc	nc	off-farm income
v_hcon_min	kg	minimum consumption
v_hcon_nonfood	nc yr^{-1}	non food consumption
v_hiredLabor	person-days	hired labor

4.3 Household Food and Nutrition

DAHBSIM accounts for non-separability between production and consumption and calculates human food consumption using a Linear Expenditure System (LES), as used in Louhichi and Gomez y Paloma (2014). In this system food and non-food expenditures are increasing in income and decreasing in own price. The system decribes household expenditures for a set of 32 food items and a non-food bundle.

$$p_i q_{h,i} = \gamma_{h,i} + \beta_{h,i}(y_h - \sum_j \gamma_{h,j} p_{h,j}) \qquad [42]$$

With

$$\begin{cases} 0 < \beta_{h,i} < 1 \\ \sum_{j} \beta_{h,i} = 1 \\ q_{h,i} - \gamma_{h,i} > 0 \end{cases}$$

Where $p_i$is the price of good *i,* $q_{h,i}$ is the quantity of good *i* consumed by household h; y_h is the household full income. $\beta_{h,i}$ and $\gamma_{h,i}$ are the LES parameters. In this system $\sum_j \gamma_{h,j} p_{h,j}$ is considered as the subsistence expenditure and the term $\left(y_h - \sum_j \gamma_{h,j} p_{h,j}\right)$ is generally interpreted as the *supernumerary* income (Sadoulet and de Janvry, 1995, 42).

We adapt the income elasticities of food demand from Ecker and Qaim (2011) and the Frisch parameter for Subsaharan Africa from the GTAP database to compute the parameters γ and β. Based on household food consumption, a series of nutrition indicators are calculated. These report nutrition intake of macro and micronutrients per individual in the household. Macro nutrients include calories and protein, and micronutrients are for vitamin A, vitamin B12, vitamin C, iron, zinc, and folate. These nutrition intakes are calculated and compared to observed data for the study region as well as reference values representing the recommended intake (recommended daily allowance – RDA). The nutrition content of food items are obtained from Ecker and Qaim (2011) and FAO (2012).

4.4 Model Calibration

Our model uses positive mathematical programming to calibrate the cropping areas to the observations obtained in household surveys. This is required due to the poor quality of the data related to costs of crop production obtained from survey data for smallholder farming in Africa south of the Sahara. This method has been used previously for calibrating farm production models under uncertainty in data quality (Louhichi *et al*, 2013). For an in depth overview of this calibration method, see Howitt (1995).

The PMP technique involves a two stage procedure that introduces a cost term in the objective function that represents the implicit cost associated with a given crop activity. In the first stage, the model is run with modelled crop areas fixed at observed levels. The shadow price of this constraint is calculated and used to calculate a PMP cost term. Then, the model is re-run with an additional term in the objective function which represents the implicit cost of a given crop activity that isn't observed in the data.

An alternative or complementary method of calibration is to introduce a non-linearity in the model objective that is related to the preferences of the small-holder farmers, who are assumed to be risk-averse. This is a plausible assumption, given the characteristics of farming in many rural, developing settings, and the body of literature that points towards the risk preference of agricultural household farms (Binswanger, 1981; Antle, 1987; Chavas, 2004). For this purpose, we introduce a risk module into DAHBSIM, that is described below.

4.5 Risk Module

The risk module contains the equations for introducing the mean-standard deviation approach for risk analysis (Hazel & Norton, 1984). This approach is similar to that used by which Semaan *et al*. (2007) and Blanco-Gutiérrez *et al*. (2011). We consider the variation in both prices and yields, which together influence production and consumption decisions. The historical price and yield data are obtained from FAOStat (FAO 2016a, FAO 2016b).

An equation calculates the utility for each state of nature, or world state (WS), emerging from the price and yield variability. The standard deviation of utility over all possible world states is calculated and introduced in the objective function as a negative term, affected by a risk aversion coefficient (see the definition of the objective function in the model overview section). The random utility under a given world state is defined as follows:

$$v\_utrd_{WS} = v\_MarkSales \times p\_sellPrice_{MS} - v\_inputUse \times p\_buyPrices_{WS} \qquad [43]$$

Where

v_utrd_{WS} is the random utility under a specific world state v_MarkSales is the quantity of market sales

p_sellPrice$_{WS}$ is the selling price under a given world state

v_inputUse is input use related to crop and livestock production

p_buyPrices$_{WS}$ are the buying prices under a given world state

The random net present value is defined as follows:

$$\text{NPV}_{WS} = \sum_{y} (1/(1+\text{i})^{\text{y-1}}) \times \text{v\_utrd}_{\text{ws}} \qquad [44]$$

Where

NPV$_{ws}$ is the net present value under a given world state

i is the discount factor (taking a default value of 0.04)

y represents the number of years in the inter temporal horizon

Finally, the standard deviation of net present value is defined as follows:

$$\text{v\_StdDev} = \frac{\sqrt{\sum(NPV - NPV(WS))\hat{}2}}{n} \qquad [45]$$

Where

v_StdDev is the standard deviation

NPV is the real (inflation adjusted) net present value

n is the number of world states

Louhichi *et al* (2010) used a similar approach to risk. The differences concern two aspects:

- DAHBSIM is a household model, taking into account consumption, whereas the model developed by Louhichi *et al* (2010) is a supply-side model

- The introduction of a biophysical module in the bio-economic model, allowing a re-initialization of yields in relation with previous crop and weather, defining the crop rotations in an endogenous way

REFERENCES

AFRC (Agriculture and Food Research Council). 1993. *Energy and Protein Requirements of Ruminants: An Advisory Manual*. Prepared by the AFRC Technical Committee on Responses to Nutrients.

Antle, J.M. 1987. Econometric Estimation of Producers' Risk Attitudes, *American Journal of Agricultural Economics* 69, 509-522. http://dx.doi.org/10.2307/1241687

Binswanger, H.P. 1981. Attitudes towards risk: theoretical implications of an experiment in rural India, *The Economic Journal* 91, 867–890. http://dx.doi.org/10.2307/2232497

Belhouchette H., Braudeau, E., Donatelli, M., Mohtar, R.H., Wery, J.. 2008. Integrating spatial soil organization data with a regional agricultural management simulation model: a case study in northern Tunisia. *Transactions of the ASABE* 51(3): 1099-1109.

Belhouchette, H., Blanco, M., Wery, J., and Flichman, G. 2012. Sustainability of irrigated farming systems in a Tunisian region: A recursive stochastic programming analysis. *Computers and Electronics in Agriculture* 86, 100-110. http://dx.doi.org/10.1016/j.compag.2012.02.016

Belhouchette, H., Louhichi, K., Therond, O., Mouratiadou, I., Wery, J., Ittersum, M. v., and Flichman, G. 2011. Assessing the impact of the Nitrate Directive on farming systems using a bio-economic modelling chain. *Agricultural Systems* 104, 135-145. http://dx.doi.org/10.1016/j.agsy.2010.09.003

Blanco-Gutiérrez, I., Varela-Ortega, C. and Flichman, G.. 2011. Cost-effectiveness of groundwater conservation measures: A multi-level analysis with policy implications. *Agricultural Water Management* 98, 639-652. http://dx.doi.org/10.1016/j.agwat.2010.10.013

Bouman, B. A. M., van Keulen, H., van Laar, H. H., and Rabbinge, R.. 1996. The 'School of de Wit' crop growth simulation models: A pedigree and historical overview. *Agricultural Systems* 52, 171-198. http://dx.doi.org/10.1016/0308-521X(96)00011-X

Cannas, A., Tedeschi, L.O., Fox, D.G., Pell, A.N. and Van Soest, P.J.. 2004. A mechanistic model for predicting the nutrient requirements and feed biological values for sheep. *Journal of Animal Science* 82, 149-169.

Chavas, J.P.. 2004. *Risk Analysis in Theory and Practice*. Elsevier Academic Press, San Diego.

Clavel, L., Soudias J., Baudet, D. and Leenhardt, D.. 2011. Integrating expert knowledge and quantitative information for mapping cropping systems. *Land Use Policy* 28, 57-65. http://dx.doi.org/10.1016/j.landusepol.2010.05.001

CSIRO (Commonwealth Scientific and Industrial Organization). 2007. *Nutrient Requirements of Domesticated Ruminants*. CSIRO Publishing, Collingwood.

Donatelli, M., Stockle, C.O., Ceotto, E., Rinaldi, M.. 1997. Evaluation of CropSyst for cropping systems at two locations of northern and southern Italy. *European Journal of Agronomy* 6: 35–45. http://dx.doi.org/10.1016/S1161-0301(96)02029-1

Doorenbos, J. and Kassam, A.H.. 1979. Yield response to water. FAO Irrigation and Drainage Paper No. 33. Rome, FAO.

Doorenbos, J. and Pruitt, W.O.. 1977. Crop water requirements. FAO Irrigation and Drainage Paper No. 24. Rome, FAO

Ecker, O. and Qaim, M.. 2011. Analyzing Nutritional Impacts of Policies: An Empirical Study for Malawi. *World Development* 39, 412-428. http://dx.doi.org/10.1016/j.worlddev.2010.08.002

FAO. 2016a. Annual Producer Prices for Malawi: 2000-2014. Food and Agriculture Organization of the United Nations Statistics Division.

FAO. 2016b. Crop Yields for Malawi: 2000-2014. Food and Agriculture Organization of the United Nations Statistics Division.

FAO. 2012. West African Food Composition Table. Food and Agriculture Organization of the United Nations. Rome, Italy.

FAO. 1986. Yield response to water: the original FAO water production function. FAO irrigation and drainage paper, Rome, Italy.

Faivre, R., Leenhardt, D., Voltz, M., Benoît, M., Papy, F., Dedieu, G., Wallach, D.. 2004. Spatialising crop models. *Agronomie* 24: 205–217

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, 33–34. http://dx.doi.org/10.1126/science.1234485

Gibbs, M.J., Conneely, D., Johnson, D., Lassey, K.R. and Ulyatt, M.J. 2002. CH_4 emissions from enteric fermentation. in Eggleston, H.S., Buendia, L., Miwa, K., Ngara, T. and Tanabe, K. (eds.), *Background Papers: IPCC Expert Meetings on Good Practice Guidance and Uncertainty Management in National Greenhouse Gas Inventories*. IPCC-NGGIP, IGES, Japan, http://www.ipccnggip.iges.or.jp/public/gp/bgp/4_1_CH4_Enteric_Fermentation.pdf [Accessed May 5, 2015]

Grandin, B. E.. 1983. Labour data collection. In: Stewart R A (ed.), *Pastoral systems research in sub-Saharan Africa.* Proceedings of the workshop held at ILCA, Addis Ababa, Ethiopia, 21 to 24 March, 1983. ILCA (International Livestock Centre for Africa), Addis Ababa, Ethiopia. pp. 305-319.

Harris, I., Jones, P.D., Osborn, T.J. and Lister, D.H. 2014. Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset. *International Journal of Climatology* 34, 623-642.

Herrero, M., Grace, D., Njuki, J., Johnson, N., Enahoro, D., Silvestri, S. and Rufino, M.C.. 2013. The roles of livestock in developing countries. *Animal* 7: 3-18. http://dx.doi.org/10.1017/S1751731112001954

Howitt, R.. 1995. Positive Mathematical Programing. American Journal of Agricultural Economics 77(2): 329-342. http://www.jstor.org/stable/1243543.

International Food Policy Research Institute (IFPRI), 2015. Malawi Africa Research in Sustainable Intensification for the Next Generation (Africa RISING) Baseline Evaluation Survey. http://dx.doi.org/10.7910/DVN/28557. Harvard Dataverse, V1.

IPCC (Inter-governmental panel on climate change). 2006. Emissions from Livestock and Manure Management. in Eggleston, H.S., Buendia, L., Miwa, K., Ngara, T. and Tanabe, K.. (eds.), *IPCC Guidelines for National Greenhouse Gas Inventories* IGES. Japan, http://www.ipccnggip.iges.or.jp/public/2006gl/pdf/4_Volume4/V4_10_Ch10_Livestock.pdf [Accessed May 5, 2015].

Jagtap, S. S., and Jones, J. W.. 2002. Adaptation and evaluation of the CROPGRO-soybean model to predict regional yield and production. *Agriculture, Ecosystems & Environment* 93, 73-85. http://dx.doi.org/10.1016/S0167-8809(01)00358-9

Jamieson, P.D., Porter, J.R., Wilson, D.R.. 1991. A test of computer simulation model ARC-WHEAT1 on wheat crops grown in New Zealand. *Field Crops Research* 27, 337–350. http://dx.doi.org/10.1016/0378-4290(91)90040-3

Jarrige, R.. 1989. Ruminant nutrition: recommended allowances and feed tables. INRA, Paris.

Louague, K., Green, R.E.. 1991. Statistic and graphical methods for evaluating solute transport models: Overview and application. *Journal of Contaminant Hydrology.* 7: 51-73. http://dx.doi.org/10.1016/0169-7722(91)90038-3

Louhichi, K., Kanellopoulos, A., Janssen, S., Flichman, G., Blanco, M., Hangsdijk, H., Heckele, T., Berentsen, P., Oude Lansink, A., and Van Ittersum, M.. 2010. FSSIM, a bio-economic farm model for simulating the response of EU farming systems to agricultural and environmental policies. *Agricultural Systems* 103(8):585-597. http://dx.doi.org/10.1016/j.agsy.2010.06.006

Nennich, T.D., Harrison, J.H., VanWieringen, L.M., Meyer, D., Heinrichs, A.J., Weiss, W.P., St-Pierre, N.R., Kincaid, R.L., Davidson, D.L. and Block, E.. 2005. Prediction of Manure and Nutrient Excretion from Dairy Cattle. *Journal of Dairy Science* 88: 3721-3733. http://dx.doi.org/10.3168/jds.S0022-0302(05)73058-7

NRC. (National Research Council). 1989. *Nutrient Requirements of Dairy Cattle*. 6th Edition. National Academy Press, Washington, D.C.

NRC. (National Research Council). 1996. *Nutrient Requirements of Beef Cattle*. National Academy Press, Washington, D.C.

NRC. (National Research Council). 2000. *Nutrient Requirements of Beef Cattle*. Revised 7th Edition. National Academy Press, Washington, D.C.

NRC. (National Research Council). 2001. *Nutrient Requirements of Dairy Cattle*. Revised 7th Edition. National Academy Press, Washington, D.C.

NRC (National Research Council). 2007. *Nutrient Requirements of Small Ruminants*. National Academy Press, Washington, D.C.

Ollenburger, M., Snapp, S.. 2014. Model applications for Sustainable Intensification of Maize-based smallholder cropping in a changing world. Advances in Agricultural Systems Modeling. 5. ASA, CSSA, and SSSA, Madison, WI. p. 375-398. doi:10.2134/advagricsystmodel5.c14

Ollenburger M., 2012. Modeling integrated soil fertility management options in Malawi under variable soil and climate conditions. Master of Science: crop and soil sciences, Michigan State University, pp 107.

Opio, C., Gerber, P., Mottet, A., Falcucci, A., Tempio, G., MacLeod, M., Vellinga, T., Henderson, B. and Steinfeld, H.. 2013. Greenhouse gas emissions from ruminant supply chains – A global life cycle assessment. FAO, Rome. http://www.fao.org/docrep/018/i3461e/i3461e.pdf [Accessed May 5, 2015]

Oreskes, N., Shrader-Frechette, K., Belitz, K.. 1994. Verification, validation, and confirmation of numerical models in the earth sciences. *Science*, 263:641-646. http://dx.doi.org/10.1126/science.263.5147.641

Philp, J., Komarek, A.M., Pain, S.J., Li, X. and Bellotti, W.. 2015. Improving the use of available feed resources to overcome sheep feeding deficits in western China, *Animal Production Science* 56(9): 1545-1550. http://dx.doi.org/10.1071/AN14694

Priestley, C.H.B., and Taylor, R.J., 1972. On the assessment of surface heat flux and evaporation using large-scale parameters. *Monthly Weather Review*. 100:81-82. http://dx.doi.org/10.1175/1520-0493(1972)100

Rasby, R., Stock, R., Anderson, B. and Schneider, N.. 1988. Nitrates in Livestock Feeding, *NebGuide G74-170. Cooperative Extension Service Institute of Agriculture and Natural Resources University of Nebraska-Lincoln*. http://ianrpubs.unl.edu/live/g1779/build/g1779.pdf [Accessed May 5, 2015].

Rufino, M.C., Herrero, M., Van Wijk, M.T., Hemerik, L., De Ridder, N. and Giller, K.E.. 2009. Lifetime productivity of dairy cows in smallholder farming systems of the Central highlands of Kenya. *Animal* 3: 1044-1056. http://dx.doi.org/10.1017/S1751731109004248

Sadoulet, E., de Janvry, A.. 1995. Quantitative Development Policy Analysis. American Journal of Agricultural Economics 78(1): 251-253.

Semaan, J., Flichman, G., Scardigno, A. and Steduto, P. 2007. Analysis of nitrate pollution control policies in the irrigated agriculture of Apulia Region (Southern Italy): A bio-economic modelling approach, *Agricultural Systems* 94: 357-367. http://dx.doi.org/10.1016/j.agsy.2006.10.003

Stern, R.D., Cooper, P.J.M.. 2011. Assessing climate risk and climate change using rainfall data - A case study from Zambia. *Experimental Agriculture* 47:241-266.

Tedeschi, L.O., Cavalcanti, L.F.L., Fonseca, M.A., Herrero, M. and Thornton, P.K.. 2014. The evolution and evaluation of dairy cattle models for predicting milk production: an agricultural model intercomparison and improvement project (AgMIP) for livestock. *Animal Production Science* 54: 2052-2067. http://dx.doi.org/DOI: 10.1071/AN14620

Therond, O., Belhouchette, H., Janssen, S., Louhichi, K., Ewert, F., Bergez, J.-E., Wery, J., Heckelei, T., Olsson, J. A., Leenhardt, D., and Van Ittersum, M.. 2009. Methodology to translate policy assessment problems into scenarios: the example of the SEAMLESS integrated framework. *Environmental Science & Policy* 12: 619-630. http://dx.doi.org/10.1016/j.envsci.2009.01.013

Tilman, D., Balzer, C., Hill, J. and Befort, B.L.. 2011. Global food demand and the sustainable intensification of agriculture, *Proceedings of the National Academy of Sciences* 108: 20260–20264. http://dx.doi.org/10.1073/pnas.1116437108

Tolkamp, B.. 2010. Efficiency of energy utilisation and voluntary feed intake in ruminants. *Animal* 4: 1084-1092. http://dx.doi.org/10.1017/S1751731110000455

USDA (United States Department of Agriculture). 2011. National Nutrient Database for Standard Reference. United States Department of Agriculture, Agricultural Research Service. http://ndb.nal.usda.gov/ [Accessed May 5, 2015].

Van Ittersum, M.K., Leffelaar, P.A., van Keulen, H., Kropff, M.J., Bastiaans, L. 2003. On approaches and applications of the Wageningen crop models. *European Journal of Agronomy* 18: 201-234. http://dx.doi.org/10.1016/S1161-0301(02)00106-5

Wong, N.P., Jenness, R., Keeney, M. and Marth, E.H.. 1988. *Fundamentals of Dairy Chemistry*. AVI Books, New York.

APPENDIX 1 HOW DAHBSIM WORKS

DAHBSIM is written in the General Algebraic Modelling System (GAMS). The following sections are designed to aid the user in understanding and working with the model in GAMS.

RUNNING DAHBSIM IN GAMS

The DAHBSIM modelling approach relies on the modularity of different sub components, which are then integrated into a larger global model. These modules can be turned on or off in the file "settings.gms". Furthermore, the model years, including the quantity of years in the inter-temporal horizon, and the number of recursive iterations are defined here. The number of years in the inter-temporal horizon is defined as the set 'y', whereas the quantity of recursive iterations are defined as 'y2'. Before running the model, the file "set_database.gms" must be run in order to load in the raw data from the 'raw_data' folder and create the necessary parameters.

After this, the model can be run using the file "gen_baseline.gms". This file first runs all the modules that are included in the model, and then passes control to "simulation_model", in which the global model and objective function are defined based on all the modules that were included in the settings file. The objective function is then run consecutively over the years of the recursive period. At the end of each iteration, parameters are re-initialized in the file "reset_iniyear.gms". In particular, after each iteration, the biophysical coefficients are re-computed. Both "Nitrate_module.gms" and "Water_module.gms" are designed to be re-run after each recursive iteration in order to re-calculate the Nitrogen and Water stress coefficients (in the sections 'reset'). These are then used to update next year's yields based on the previous period's management and climate outcomes.

After each recursive iteration is executed, the results from the specific year are saved in a series of parameters, which are then stored in the folder 'output_data'. A specific file is created for each module, in the format of 'report_module name.xls'. These describe the model results over the entire span of the model years.

Running the Evaluation Files in GAMS

The file "nutrient_moduleevaluation.gms" is a stand-alone module that evaluates how the module predicts nutrient requirements compared to similar scenarios found in the literature. Data to run nutrient_moduleevaluation.gms are located in %rawdir%\livestockevaluationdata.xlsx. Results from the nutrient module evaluation are located in results\simu\livestockevaluationresults.xlsx. Similarly, the files "Nitrate_moduleevaluation.gms" and "Methane_moduleevaluation.gms" evaluate the model estimated nitrogen and methane production from livestock and compare these values to values obtained from the literature.

MODULE LINKAGES

The following sections describe the linkages between the different sub components of DAHBSIM. Variables that are defined in specific modules can act as variables or parameters in other modules.

Biophysical Module (BIOPH)

Nitrate Module

Inputs:

v_nitlive – Nitrogen from livestock

v_residuesmulch – Mulch from residues

v_Nfert – Nitrogen from Fertilizer

Outputs:

p_Nw – N stress coefficient (used to calculate next year's yields)

Water Module

Inputs:

p_rain – Monthly rainfall

Outputs:

p_hw – Water Stress Coefficient (used to calculate next year's yields)

CROP_MODULE (.GMS)

Inputs:

v_cactYld -- CropYields

Outputs:

v_prodQuant – Crop Production

v_cropLabor – Crop Labor

v_residuesfeed – Residues for Livestock Feed

v_residuessell -- Residues Sold

v_residuesmulch – Residues for Mulch

FARM MODULE (.GMS)

Inputs:

v_markSales – Market Sales of Crop and Animal Products

v_animalPurch – Animal Purchases

v_residuesBuy – Residues Bought

v_prodQmeat -- Meat Produced

v_prodQmilk – Milk Produced

v_seedQuant – Seed Quantity by Cropping Activity

v_aactLevBuy – Animal Purchases

v_aactLevSell – Animal Sales

Outputs:

v_seedPurch – Seed Purchases by Cropping Activity

v_farmLabor – Farm Labor

v_Nfert -- Nitrogen Fertilizer

v_nitlive – Nitrogen from Livestock

v_residuesmulch – Residues from Mulch

v_selfCons – Self Consumption

v_markPurch – Market Purchases

v_farmIncome – Farm Income

v_aactIncome – Animal Activity Income

v_cactIncome – Crop Activity Income

HOUSEHOLD_MODULE (.GMS)

Inputs:

v_farmIncome – Farm Income

v_hiredLabour – Hired Labour

Outputs:

v_hconQuant – Household Consumption

LIVESTOCK MODULE (.GMS)

Nutrient Module

Outputs:

p_feedreq – monthly livestock feed requirements

Livestock Module

Inputs:

v_residuesfeedm – crop residue availability for livestock feed

Outputs:

v_prodQmilk - milk production

v_prodQmeat - meat produced from slaughtered animals

v_aactLevbuy - purchased animals

v_aactLevsell - sold animals

v_aactLev - animals owned by the household

v_nitlive - represents total nitrate excretion from livestock per household per year in kg

v_ManureDM - manure production

residuesbuy - crop residues purchased

The feedactivity, and productivitycalculation modules are used to define input/output relationships for the livestock module.

Risk module (.GMS)

Inputs:

v_inputUse – Input Use

v_seedPurch – Seed Purchases

p_buyPrice – Buy Prices
v_markSales – Market Sales
v_animalpurch – Animal Purchases
p_selPrice – Sell Prices

Outputs:
v_npvrd_tot – Random Net Present Value

Running the Livestock and Nutrient Modules in GAMS

Running gen_baseline.gms runs nutrient_module.gms and livestock_module.gms. Both modules rely on information contained in sets_generic.gms and sets_specific.gms, and sets_database.gms generates the associated files. In addition, nutrient_module.gms calls in data from livestockdata.xlsx.

nutrient_moduleevaluation.gms is a stand-alone module that evaluates how the module predicts nutrient requirements compared to similar scenarios found in the literature. Data to run nutrient_moduleevaluation.gms are located in %rawdir%\livestockevaluationdata.xlsx. Results from the nutrient module evaluation are located in results\simu\livestockevaluationresults.xlsx.