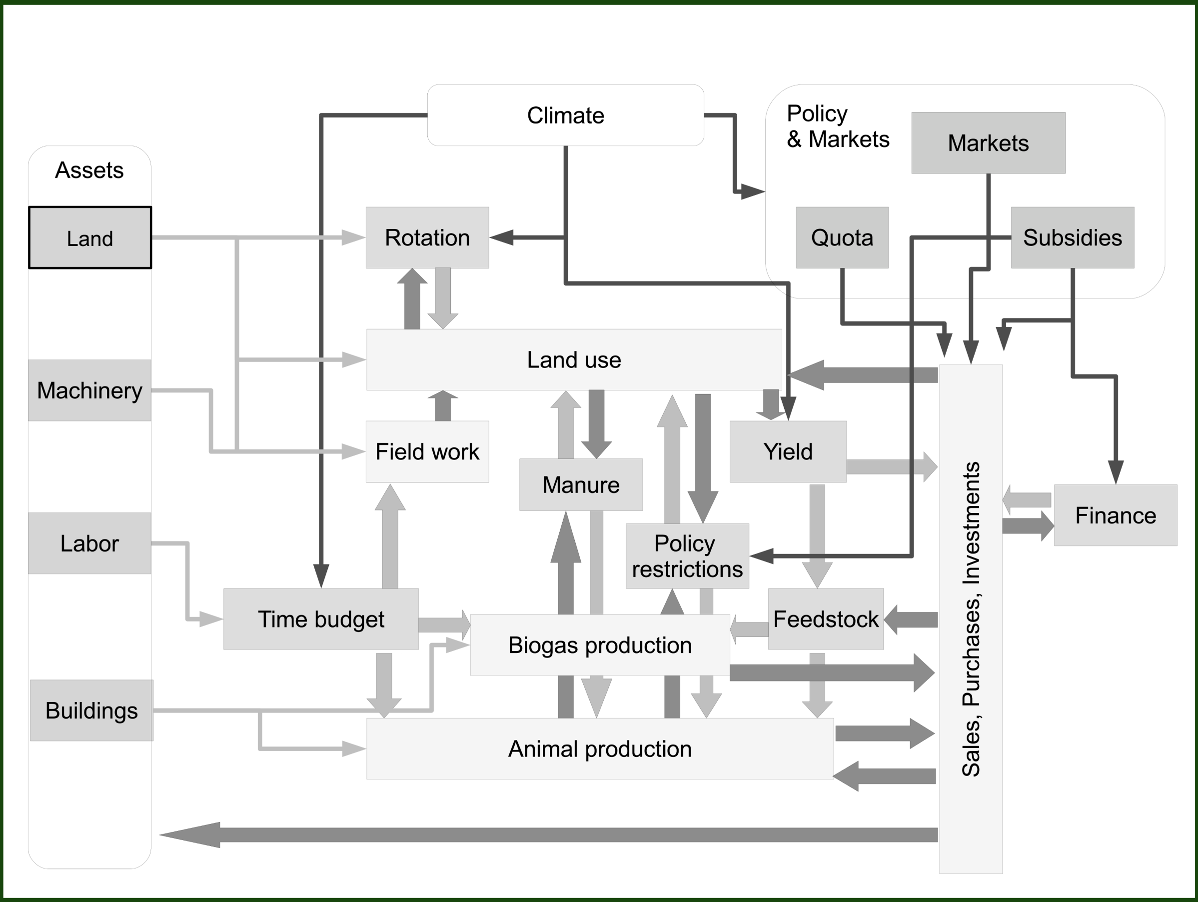
MPMAS Central Swabian Jura
Version 3.1



Christian Troost

MPMAS Central Swabian Jura (Version 3.1)

Model Documentation (v1.1)

MPMAS Central Swabian Jura

Model Documentation

Model Version 3.1

Christian Troost

Documentation Version 1.1

September 27, 2014

Contact: Christian Troost

Dept. of Land Use Economics in the
Tropics and Subtropics (490d)
University of Hohenheim, 70599 Stuttgart,
Germany

christian.troost@uni-hohenheim.de

Contents

List of acronyms

AA	agricultural area
ABM	agent-based model
AMS	automatic milking system
AOGCM	Atmosphere-Ocean General Circulation Model
AR4	the Fourth Assessment Report of the IPCC
BW	Baden-Württemberg
CAP	EU Common Agricultural Policy
CGE	computable general equilibrium model
EEG	*Erneuerbare Energien Gesetz* 'Renewable Energy Act'
EMIC	Earth System Model of Intermediate Complexity
ESU	European Size Unit (EU farm typology)
EU	European Union
FADN	Farm Accounting Data Network
GCM	General Circulation Model
GHG	greenhouse gas
GTOF	general type of farm (EU farm typology, level 1)
ha	hectare
IPCC	Intergovernmental Panel on Climate Change
LAI	leaf area index
LU	livestock unit
KWK	*Kraft-Wärme-Kopplung* 'Combined heat-power generation'
MAS	multi-agent system

MEKA	Marktentlastungs- und Kulturlandschaftsausgleich, 'Compensation Scheme for Market Easing and Landscape Protection' (an agri-environmental support scheme in BW)
ME	metabolizable energy
MF	main forage area
MIP	mixed-integer programming
MP	mathematical programming
MPMAS	mathematical programming based multi-agent system
MTR	Mid-Term Review
NaWaRo	*Nachwachsende Rohstoffe* 'Renewables'
NEL	net energy lactation
nXP	usable raw protein
PTOF	principal type of farm (EU farm typology, level 2)
py	person-year: yearly workload of a full-time employee
RCM	Regional Climate Model
RLU	roughage-consuming livestock unit
SCM	Simple Climate Model
SRES	Special Report on Emissions Scenarios
UAA	utilized agricultural area
UPCS	Unbiased Permuted Column Sample
VBA	Visual Basic for Applications
XP	raw protein

Chapter 1

Introduction

The agent-based model described in this document has been developed in order to analyze adaptation of farmers to changes in agronomic and socioeconomic conditions, which might be triggered by global climate change. Agents in the model represent farmers in a study area on the Central Swabian Jura.

When designing the model, we took the following aspects into consideration:

- The design of the decision model needs to explicitly account for the influence of yields, changes in available field working time due to meteorological conditions, changes in rotation options and market prices in order to assess the sensitivity of farmers' decisions to postulated effects of climate change.
- A dynamic modeling of adaptation requires the representation of investment decisions including the inertia caused by sunk costs and lack of liquidity, and for longer term modeling the possibility for farmers to trade land or give up farming.
- To assess vulnerability of different types of farms requires a good representation of heterogeneity, e.g. economies of scale due to indivisibility of assets and effects of soil and topographic location on yields and time for field work.
- The decision model has to be represented as a mixed integer programming model and should be solvable in a few seconds in order to allow for a large number of model evaluations with a few thousand agents.

The model has been implemented using the multi-agent modeling package MPMAS. MPMAS is an agent-based modeling framework, which stands in the agricultural economics tradition of recursive farm modeling and adaptive micro-systems using mixed integer programming to represent agent decisions [Schreinemachers and Berger, 2011]. The main – and in most applications the only – class of agents in MPMAS are farm households, each of which runs through a typical sequence of actions in each cropping season, which is depicted in figure 1.1.

Based on past experience and available information, the agents form expectations about future conditions (e.g. prices, yields). Based on these and their knowledge about

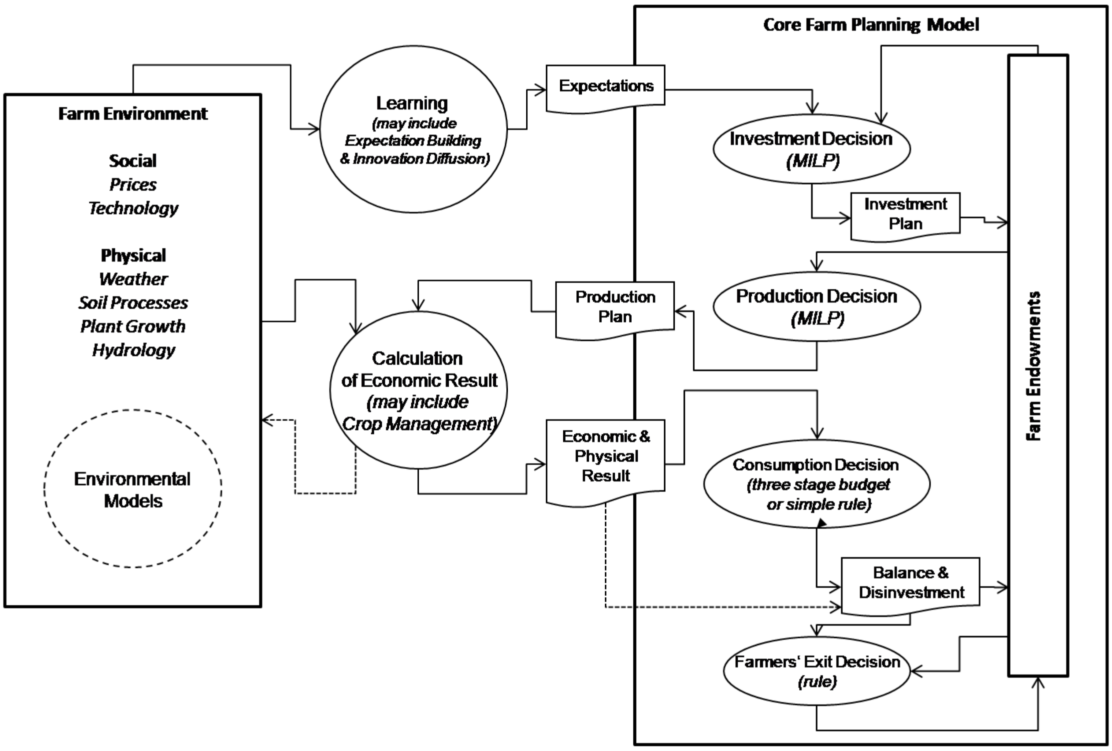


Figure 1.1: *The seasonal cycle of agent decisions and submodel invocation in MPMAS*

their current situation, the agents first decide on investments into assets (e.g. machinery, stables, etc.), and then on the production plan for the coming season. The actual physical and economic outcomes of production are determined, and the agents react to the observed outcome deciding on the usage of produce and income, whether to sell assets to retain solvency, and whether to continue farming or leave the agricultural sector.

Model equations and software architecture of the MPMAS framework are described in Schreinemachers and Berger [2011] following the ODD [Grimm et al., 2010] and will therefore not be repeated in this paper; technical documentations, software manuals, input data and executable programs can be downloaded from `https://mp-mas.uni-hohenheim.de`.

Chapter 2 contains an extensive description of the equations of the decision model used in the Central Swabian Jura application. Chapter 3 explains the generation of initial agent populations and Chapter 4 documents the choice of exogenous variables depending on scenarios, or observation years in the calibration. Chapter 5 describes the approach to calibration chosen, and compares the test simulations to observations.

Chapter 2

The model design for the Central Swabian Jura

Since MPMAS is an agent-based modeling software, we follow the ODD+D protocol [Müller et al., 2013] for the description of our model for consistency with other MPMAS applications, although our model in its current form – abstracting from interactions – is not a fully connected agent-based model, but rather a farm-level model run for every full-time farm in the study area.

2.1 Overview

2.1.1 Purpose

The model has been designed to analyze the adaptation of agricultural production decisions to potential effects of climate change. The model should provide insight into the importance of different climate related impacts, specifically the influence of yields, changes in available field working time due to meteorological conditions, changes in rotation options and market prices It should be capable of simulating the vulnerability of different types of farms and highlight the effects of climate change on the effectiveness of existing policies, specifically agri-environmental measures and biogas support. As the model is to be tested against observation data from 1999 to 2007, it needs to include relevant policy regulations valid during this time span. The model is not, however, expected to provide an accurate forecast of future development, i.e. answer a 'how will it be?' type of question, but rather improve the understanding of the influence of relevant processes, e.g. agent heterogeneity, expectation formation, and land market transactions and help to explore potential feedbacks on land surface processes.

2.1.2 Entities, state variables and scales

Every full-time farm of the study area is represented by an individual model agent. The state of the agent is characterized by individual household composition, asset ownership, soil endowment and current expectations. The state of the household includes gender, age, the status of household members, and the expected remaining lifetime of the farm (until retirement of the household head, resp. their potential successor). The state of assets includes the age and time value of tangible assets and intangible assets (quotas and entitlements) as well as equity, cash and liabilities. Expectations are related to expected future values of prices, average yields and household composition. Agricultural land is characterized by different soil types and represented at a resolution of one hectare.

2.1.3 Process overview and scheduling

Figure 1.1 summarizes the sequence of actions, which is repeated for every agent in every simulation period. Agents start each season by forming expectations about future conditions (e.g. prices, yields). Expectations are specified by the modeler and are constant over time. In the next step, the agents decide on investments into assets (e.g. machinery, stables, etc.) by solving the decision problem for an expected average year of the future, and the chosen investments are then implemented, i.e. the state of the agents' asset is updated accordingly.

Agents then make a production plan for the current season, and the actual physical and economic outcomes of production are determined. The model calculates income, cash flow, debt service and rental payments, and increases the age of assets and household members. Assets, which have reached end of use life, are removed from the list of assets, and the agent decides on withdrawals for consumption. The agent shuts down the farm if bankruptcy cannot be avoided. Finally, the model determines, whether household members die, retire or give birth. If the household head is scheduled to retire, retirement and continuation of the farm depend on the willingness of a potential successor to continue the business.

2.2 Design concepts

2.2.1 Theoretical and empirical background

The model rests on the traditional agricultural economics approach of representing farm decisions as mathematical programming problems of choosing an optimal set of activities given technological and resource constraints specific to the farm [Hazell and Norton, 1986]. In a wider sense, if falls under the set of recursive-dynamic programming models representing economic decisions as described by Day [2008]. Besides the economic considerations of maximizing expected household income, while ensuring liquidity and long-run survival of the farm, agents have a preference for employing their own children and are assumed to comply with good farming practice and agri-environmental regulations.

Technical coefficients are based on standard references for farmers provided by extension services [e.g. KTBL, 2010; LfL, 2010, 2011], expert interviews and a farm survey. Farm census data and official demographical statistics are used to initialize the

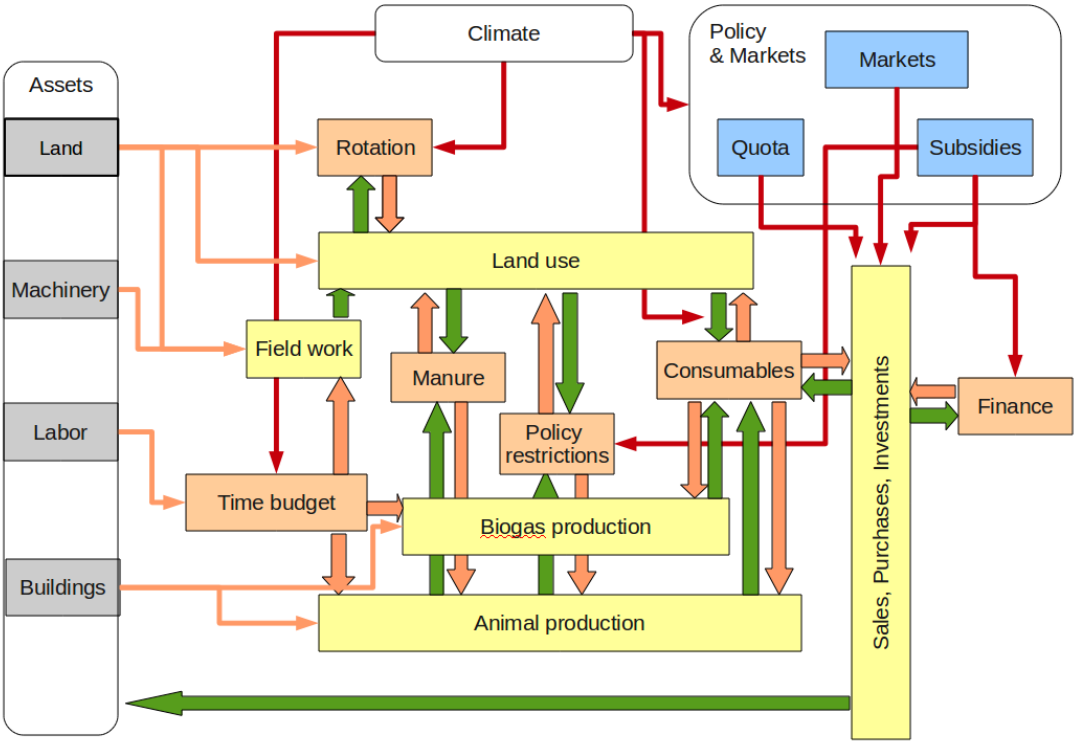


Figure 2.1: *Interdependencies of farm activities including the impact of climate change*

agent population. Price data is derived from various statistical databases LEL [2010, 2011a,b]; KTBL [2010]; destatis [2012d]. Crop yield information is taken from statistics and simulations with crop growth models, which were calibrated by multi-year field observations in the area (XXX).

2.2.2 Individual decision making

Farmers usually dedicate themselves to a number of mutually interdependent production activities. Figure 2.1 shows a rather aggregate representation of the conceptual model of the farmer's decision problem. The basic assumption of the decision model is that farmers maximize expected total income by choosing an optimal combination of production activities (shaded in yellow in the figure). Crop production, grassland use, animal production and biogas production constitute the major alternatives, but decisions also have to be taken on the selling and buying of products and inputs, field work, investments and application for premiums (Marktentlastungs- und Kulturlandschaftsausgleich (MEKA), European Union (EU)). The choice of activities is constrained by a number of restrictions and balances (shaded in light orange in the figure) including the manure balance, the time budget, crop rotation, the financial balance, the balance of products and inputs (yield, feedstock) and restriction imposed by policy regulations or subsidy conditions. This coarse conceptualization is mathematically represented as a mixed integer programming problem MIP in the model (see section 2.5), which serves as the basis for the two decision problems: investment and pre-season production decision. While the general structure of the problem remains the same in both decision stages, though investment activities are only included for the investment stage.

2.2.3 Learning, individual sensing and individual prediction

In the present study, we abstract from individual or collective learning processes. As we use long-term averages of yields, prices and environmental conditions, we assume they are known to the agents.

2.2.4 Interaction and collectives

Interactions and collective actions of agents are not considered in the present model version.

2.2.5 Heterogeneity

The structure of the agents' objective function and constraints is identical for all agents: it is a comprehensive representation of technology packages and local conditions for agricultural production. Heterogeneity is introduced into the decision module by different household compositions and resource availabilities of individual agents, e.g. the amount and type of available farm labor and land as well as the machinery and buildings owned at the start of the simulation.

These differences in starting conditions affect the profitability of production options, and in this way produce heterogeneous agent behavior. Soil types determine attainable crop yields and the tractor-power required for field work. Existing machinery and buildings are associated with sunk costs. Profitability of crops can differ given the possibility of selling or use for feeding. Household composition determines the amount of household labor available and affects the planning horizon for investment decisions. The household head's age and, respectively, his/her potential successor's age determines the expected remaining operating time of the farm, i.e. the maximum lifetime considered in agent investment calculations. Farm succession is an important topic in family business and requires some additional rules for implementation in dynamic simulation models. Here, we assume agents are glad to employ their potential heirs on the farm, and even willing to forgo own-income if a major investment or expansion of the farm is necessary to employ their successors. In the MIP decision problem, agent household heads have to remunerate their adult children's work on the farm, but they do not consider this a cost as long as their own minimum income expectation has been met.

Further, the model structure exhibits economies of size at farm level. In Southwest Germany with rather small farm sizes and indivisible tractors and other field work implements, the capacity-to-cost ratio usually declines with increasing capacity, which we considered accordingly. Building costs and livestock-related labor demands are implemented using fixed and size-dependent costs, leading to decreasing average cost functions. Certain policy schemes, however, include special regulations for smaller farms leading to dis-economies of size: under EEG regulations, for example, guaranteed biogas electricity prices decrease with volume. Again, we considered these farm-level effects accordingly.

2.2.6 Stochasticity

The model is deterministic.

2.2.7 Observation and emergence

Aggregate total land use and crop production in the study area emerge as the sum of individual agent decisions. In principle, the full state of all agents is accessible, and the individual courses of actions can be traced with log and debug tools. Analysis mostly focuses on agents' production activities, income, asset ownership and household composition.

2.3 Details: Initialization

The initial state of the model is defined by the initial agent population including their assets, household members, owned and rented land and expectations. Initialization differs between scenarios and repetitions for uncertainty analysis, and is described in more detail in chapter 3.

2.4 Details: Input data

There is no exogenous influence over the run time of the model. Exogenous influences, which change between scenarios – representing changes of exogenous conditions over time – include prices, yields, rotations options and available days for field-work. A second set of exogenous conditions, which differs between the simulation years of the validation runs, is defined by policies, e.g. EU regulations and support, MEKA agri-environmental measures, and biogas support.

2.5 Details: The farm model

In the following sections, the elements of the mixed-integer programming (MIP)(decision variables, objective function and constraint equations) used to represent production, investment, harvest and rental decisions are described in detail.

2.5.1 Objective function

The objective function that agents maximize results from subtracting the sum of all planning-dependent costs from the sum of all revenue. Revenue can be created by selling of goods (x^{sG}) , receiving interest on deposits (x^{dC}), receiving premiums awarded by different policy schemes (x^{Y}), and selling biogas electricity (x^{sYyYu}) including a potentially associated manure bonus (x^{osYyYu}). Costs result from the purchase of goods (x^{bG}), the use of machinery and buildings ($\mathrm{x}^{ZTK}, \mathrm{x}^{\alpha O}, \mathrm{x}^{\beta O}, \mathrm{x}^{\beta tM}, x^{\beta mUe}$), hiring permanent ($\mathrm{x}^{bpH}$) and temporary labor ($\mathrm{x}^{btHTK}$), payment of interest on short-term credit (x^{bC}) and the direct cost of land use (x^{L}) and animal production activities (x^{A}).

Debt payments on assets bought in the past are omitted from the objective function, which thus represents expected total farm gross margin ($\tilde{\pi}_{tgm}$) rather than expected farm income ($\tilde{\pi}$). However, as debt payments are considered planning independent fix cost, maximizing the total gross margin function is equivalent to maximizing total income.

The complete objective function is shown in equation 2.1. Explanation of the individual decision variables and the associated objective function coefficients are given throughout the subsequent sections. A comprehensive overview and explanation of symbols used in the MIP equations is also given in annex A.1.

$$
\begin{aligned}
&\max! \quad \tilde{\pi}_{tgm} = \\
&\sum_{g} \tilde{c}_{g}^{sG} \tilde{x}_{g}^{sG} + \sum_{y} c_{y}^{Y} x_{y}^{Y} + c^{dC} x^{dC} + \sum_{y_u, y_y} \left(c_{y_u, y_y}^{sYyYu} x_{y_u, y_y}^{sYyYu} \right) + \sum_{y_u, y_y} \left(c_{y_u, y_y}^{osYyYu} x_{y_u, y_y}^{osYyYu} \right) \\
&\quad + c^{sYx} x^{sYx} + \sum_{y_u, y_y, y_z} \left(c_{y_u, y_y, y_z+1}^{slYyYuYz} x_{y_u, y_y, y_z+1}^{slYyYuYz} + c_{y_u, y_y, y_z}^{slYyYuYz} x_{y_u, y_y, y_z}^{suYyYuYz} \right) \\
&- \sum_{g} c_{g}^{bG} x_{g}^{bG} - \sum_{l} \left(c_{l}^{L} x_{l}^{L} \right) - \sum_{a} \left(c_{a}^{A} x_{a}^{A} \right) - c^{pH} x^{bpH} - \sum_{t,k} c_{t,k}^{tH} x_{t,k}^{btHTK} - c^{bC} x^{bC} \\
&- \sum_{z,t,k} \left(c_{z}^{Z} x_{z,t,k}^{ZTK} \right) - \sum_{o} \left(c_{o}^{\alpha O} x_{o}^{\alpha O} \right) - \sum_{o} \left(c_{o}^{\beta O} x_{o}^{\beta O} \right) - \sum_{m} \left(c_{m}^{\beta M} x_{m}^{\beta tM} \right) - c^{\beta mUe} x^{\beta mUe}
\end{aligned}
\tag{2.1}
$$

When considering the employment of a potential successor, the labor cost for employing the family member is part of the total gross margin, but is counterbalanced by the utility of employing the successor as described in section 2.5.14. The objective function then differs from the total gross margin function by the wage paid to the potential successor.

2.5.2 Market interaction & goods balances

The agent interacts with goods markets by selling (x_{g}^{sG}) or buying goods (x_{g}^{bG}). The goods balances ensure that the farm agent cannot sell or use more of a good $g \in G$ than he/she bought or produced him/herself. Not all goods can be sold or bought, and only a limited range of goods can be produced by the farm agent itself, and of course, goods have different potential uses, e.g. as fodder, fertilizer or fuel. So for an individual good many terms in the following general balance equation are actually omitted or their corresponding coefficients are zero for an individual good.

$$x_g^{sG} - x_g^{bG} \pm \sum_l \left(a_{l,g}^{LG} x_l^L\right) \pm \sum_a \left(a_{a,g}^{AG} x_a^A\right) + \sum_{j_a,d} x_{g,j_a,d}^{fGJaD} + x_g^{uG} + \sum_a \left(a_{z,g}^{ZG} x_{z,t,k}^{ZTK}\right) \leq 0 \quad \forall g \tag{2.2}$$

Specifically, we distinguish

Pure products (**Gs**) Goods that are only sold by the farm agent, but not bought, e.g. malting barley, rapeseed, milk, meat.

Pure inputs (**Gi**) Goods that are only bought by the farm agent, but not sold, e.g. fuel, soybean meal and other industrial fodder.

Traded intermediates (**Gt**) Goods that can be both sold and bought by the farm agent, e.g. fodder barley, fodder wheat, young animals.

Non-traded intermediates (**Gn**) Goods that are produced by one process an used as an input for another process on the farm, but generally not traded in the study area, e.g. hay, grass silage.

Manure (**Go**) Non-traded intermediates with specific treatment due to their potential use in biogas plants (see section 2.5.6).

Fresh grass (**Gg**) Non-traded intermediates with specific treatment due to their only seasonal availability (see section 2.5.3).

For some cases, e.g. silage maize, the group a good falls into, is varied according to the assumptions embodied in a specific parameter combination selected during calibration.

2.5.3 Land use: crop production and grassland cultivation

Crop production and grassland cultivation are the major land uses considered in the model. Each element l of the vector of land use activities (x^L) represents a combination of a crop g_c, a soil type s, and a management plan. Grassland cultivation and grass/clover cultivation on arable land can have several products as the same area can be used up to four times a year for silage, hay, pasture and cutting of fresh grass.

We do not explicitly account for the perennial nature of grasslands at the moment, but rather distinguish between arable land and grassland. Conversion of grassland to arable land or vice versa is not considered for simplification. The land use statistics of the area also show no significant changes in overall grassland area in the study area between 1999 and 2007 and grassland conversion has effectively been forbidden in the state of Baden-Württemberg as of 2011, so this simplification seems justified.

Arable crops

The model includes winter wheat, winter wheat silage, winter rapeseed, winter fodder barley, summer fodder barley, summer malting barley, silage maize and fallow as potential production activities for arable land. Cultivating field grass is also possible on arable land, but due to the several potential harvests the associated management plan is structurally more similar to grassland, and thus described in the next subsection.

Management plans were derived from standard recommendations of German extension services [KTBL, 2010, 2008; LEL, 2012; LfL, 2012] and cross-checked and updated in expert interviews, survey results and observations on field measurement sites in the project XXX. In general, we distinguish three fertilization schemes (only mineral fertilizer, with pig manure and with cattle manure) and two tillage regimes (plough tillage and low tillage using rotary tillers). For the two summer crops, summer barley and silage maize, we include management plans with and without winter cover crops (field mustard). We do not distinguish different levels of pesticide use, but only assume a standard plant protection practice for each crop, as we are not able to simulate or estimate the yield effect of pesticide use. For some EU Agenda 2000 support schemes, it is necessary to ensure a certain use of the product (see 2.5.9). Whenever this applies, this commitment can also be considered part of the management plan.

Management plans determine the quantity of physical inputs required, and the necessary field work (tillage, sowing, fertilization, plant protection, harvest) and its timing. Except for animal manure, physical inputs are multiplied by prices and aggregated to direct cost, which enter the MIP as the objective function coefficient (c_l^L) of the corresponding land use activity (x_l^L).

Expected ($\tilde{a}_{l,g}^{LG}$) and actual yields ($a_{l,g}^{LG}$) are discussed in detail in section 4.1.

Grassland & field grass production

For grassland cultivation, we distinguish four intensity levels of production:

			Production
Level	Description	Use	[t dm /a]
0	Abandoned	Not even minimum requirements of cross compliance fulfilled	0
1	Very extensive use	conservation cuts, extensive pasture	25
2	Extensive grassland use	maximum two cuts per year	62
3	Intensive grassland use	maximum three cuts per year	83

Potential uses of grassland are grazing (G), cutting fresh grass for direct feeding (C), production of grass silage (S) or production of hay (H). Combinations of one type of fresh and one type of conserved fodder production on a single grassland plot are possible. In this case, the harvest of conserved fodder always precedes harvests of fresh fodder. The maximum number of uses is determined by the intensity level. Thus, SG, SS, SC, HH, HC, HG, G, C[1] are the potential combinations of uses for a plot managed at intensity level 2, and SSS, SSG, SSC, SG, SC, HHH, HHC, HC, HHG, HG, G, C are potential use combinations for intensity level 3. Additional grazing at the end of the season is always possible for both intensity levels. For extensive grassland (level 1), we consider three use options: year-round pasture, late cut (beginning of July, every two years) and very late cut (beginning of October, every two years). Abandoned grassland is not used at all.

The implementation of field grass production is structurally similar to grassland, the only difference being that it is not restricted to grassland plots, but rather to arable plots. Only one intensity level is considered for field grass, which allows up to four uses a year.

Grass yields of the individual uses are determined using simple regrowth parameters that relate daily regrowth during a specific half month to total expected dry matter production in the year. These parameters were calculated from data given in Berendonk [2011]. The yield obtained by a specific use is then the total regrowth between the date of the harvest and the date of the previous harvest, respectively the beginning of the growing season. Total annual dry matter production depends on the intensity level. The specific dates for the individual uses depends on both, the intensity level and the combination of uses (use profile) of a grassland plot.

Intensity level and use profile also determine the amount of fertilization and cultivation work. Similar to the implementation of arable crops, we consider three fertilization schemes: one with mineral fertilizer only, one with pig manure and one with cattle manure. Again, only manure is treated as an explicit input in the MIP (see 2.5.6) and other physical inputs are aggregated to direct costs and form the objective function coefficient (c_l^L) of the corresponding land use activity (x_l^L).

Due to the different potential uses, several different products ($g \in G$) can be obtained from one grassland plot. Additionally, for both, conserved and fresh grass, we distinguish fodder obtained from the first cut of the year, and fodder obtained from latter cuts, due to their different nutritional composition.

For fresh grass products ($g \in G_g$), separate balances for different feeding seasons ($d \in D$) are distinguished to take account of the fact that these cannot be stored and are only available at the certain point of time, when they are mature and harvested (see 2.5.4 for more detail).

The part of the MIP that links grassland and field grass activities to the balances of its products and the objective function, looks like this:

---

[1] For simplification, letters C and G are not repeated and always refer to all further potential uses till the end of the season

$$
\begin{array}{llllllll}
\sum\limits_g c_g^{sG} x_g^{sG} & -\sum\limits_g c_g^{bG} x_g^{bG} & -\sum\limits_l \left(c_l^L x_l^L\right) & +\dots & & \rightarrow & \text{obj} & \\
x_g^{sG} & -x_g^{bG} & -\sum\limits_l \left(a_{l,g}^{LG} x_l^L\right) & +\sum\limits_{j_a,d} x_{g,j_a,d}^{fGJaD} & +x_g^{uG} & \leq & 0 & \forall g \notin Gg \\
 & & -\sum\limits_l \left(a_{l,g,d}^{LG} x_l^L\right) & +\sum\limits_{j_a} x_{g,j_a,d}^{fGJaD} & & \leq & 0 & \forall (g \in Gg, d)
\end{array}
\tag{2.3}
$$

While all of the grassland areas become available for grazing at a certain point of the season, most farmers in the study area often do not make use of this option as pasturing can be quite labor intensive. For this reason, work for cutting grass is accounted for with the land use activity, but work for pasturing is accounted for at the respective feeding activities and thus only required if the area is actually used for pasturing. This implementation, however, would allow a farm agent to declare a grassland plot to be pasture only to fulfill cross compliance requirements, but then not use it at all avoiding both the work for pasturing and conservation cuts. To avoid this, an additional constraint is introduced requiring that at least 50% of the grass available in the second feeding period is actually pastured, whenever a plot is declared pasture only ($l \in Lgp$).

$$
0.5 \sum_{l \in Lp} \left(a_{l,g,d}^{LG} x_l^L\right) \quad - \sum_{j_a} x_{g,j_a,d}^{fGJaD} \quad \leq \quad 0 \quad \forall (g_g \in Ggp, d = 2) \tag{2.4}
$$

Crop rotation

Land use activities are obviously constrained by the area of a certain soil type available to each individual agent (b_s^S). More precisely, crops do either require a corresponding part of the soil to be incorporated into crop rotation (x_s^{rS}), or kept out of the crop rotation (x_s^{nS}), depending on whether they form part of the crop rotation or not.

$$
\begin{array}{lllrrl}
 & x_s^{rS} & +x_s^{nS} & = & b_s^S & \forall s \\
\sum\limits_l \left(1_{s,l}^{LS} 1_{s,l}^{LrS} x_l^L\right) & -x_s^{rS} & & = & 0 & \forall s \\
\sum\limits_l \left(1_{s,l}^{LS} 1_{s,l}^{LnS} x_l^L\right) & & -x_s^{nS} & = & 0 & \forall s
\end{array}
\tag{2.5}
$$

Cultivation of arable crops has to respect crop rotation rules. Following good agricultural practice and as observed in the study area, we assume farmers to implement production plans that can – at least in theory – be upheld for several years without violating crop rotational rules. We distinguish two types of rotation rules in the model:

First, there are maximum limits on the share of a crop or crop group in the rotation, e.g. if a crop should be grown maximum once every three years a maximum of 33% of the arable area should be cultivated with this crop. We classified the crops under consideration into rotation groups (Jr), for each of which a specific limit a_{j_r} applies. This limit is multiplied with the part of the soil that is included into the crop rotation x_s^{rS} to give the maximum area of crops of this group that can be grown by the agent.

$$\sum_{l}\left(1^{LS}_{s,l}1^{LJr}_{l,j_r}x^{L}_{l}\right) \qquad -a^{Jr}_{j_r}x^{rS}_{s} \qquad \leq \quad 0 \quad \forall(s,j_r) \tag{2.6}$$

Second, in the study area it may not be recommendable to grow crop A after crop B for plant health reasons or it may even be impossible due to incompatible timing of sowing A and harvesting B. We created two classifications of land use activities, which group together land use activities with similar characteristics as a preceding land use (Jp), respectively as a following land use (Jf). This classification is not only crop, but also management-specific, as different management plans for the same crop may entail different timings and thus affect compatibility with other crops. We then established compatibility coefficients $1^{Jpf}_{j_p,j_f}$ for each combination of $(j_p \in Jp, j_f \in Jf)$. These coefficients may take a value of 1 indicating compatibility or 0 meaning non-compatibility.

$$\begin{array}{rrcl} -\sum_{l}\left(1^{LS}_{s,l}1^{LJp}_{l,j_p}x^{L}_{l}\right) & +\sum_{j_f}x^{SJpJf}_{j_p,j_f,s} & \leq & 0 \quad \forall(s,j_p) \\ \sum_{l}\left(1^{LS}_{s,l}1^{LJf}_{l,j_f}x^{L}_{l}\right) & -\sum_{j_p}\left(a^{Jpf}_{j_p,j_f}x^{SJpJf}_{j_p,j_f,s}\right) & \leq & 0 \quad \forall(s,j_f) \end{array} \tag{2.7}$$

Some crops are compatible with themselves, but should not be repeated on the same plot more than a certain number of times (n) in a row. In this case, all sequences containing only crop Js require the inclusion of sufficient other preceding-following crop relations. E.g. if a crop is to follow itself maximum once, each hectare, where it is grown after itself has to be complemented by another hectare, where it is grown after another crop.

The corresponding coefficient $a^{JsJpJf}_{j_p,j_f,j_s}$ is -1 for all relations including crop j_s only as a following crop, and $\frac{n+1}{n}-1$ for all relations, where crop j_s follows itself.

$$\sum_{j_p,j_f,j_s}\left(a^{JsJpJf}_{j_s,j_p,j_f}x^{SJpJf}_{j_p,j_f,s}\right) \qquad \leq \quad 0 \quad \forall(s,j_s) \tag{2.8}$$

Field work & weather dependency

Every land use activity x^{L}_{l} requires certain types of field work ($w \in W$) to be executed at certain points of time. We defined nine work seasons ($t \in T$) comprising between one and seven half-months (with fine resolution in summer and coarse resolution in winter). Each type of field work requires a different amount of time and tractor power depending on the equipment used and the resistance of the soil. Field work activities ($x^{WTSqE}_{w,t,s_q,e}$[h]) are therefore combinations of a type of work w, the equipment e used, the work season t and the soil resistance class s_q .

Farm agents can do field work using own machinery or by contracting external machinery and workers (x^{bWTSq}_{w,t,s_q}). The amount of field work has to be balanced for each type of work, work season and soil resistance class as shown in equation 2.9,

$$\sum_{l} \left(a^{LWT}_{l,w,t} 1^{LSq}_{l,s_q} x^{L}_{l}\right) \quad - \sum_{w,e} \left(a^{WTSqE}_{w,e} x^{WTSqE}_{w,t,s_q,e}\right) \quad - x^{bWTSq}_{w,t,s_q} \quad \leq \quad 0 \quad \forall (w,t,s_q) \tag{2.9}$$

with $a^{LWT}_{l,w,t}$ being the amount of field work w required in work season t for land use activity x^{L}_{l}, $a^{WTSqE}_{w,e}[\mathrm{ha} \cdot \mathrm{h}^{-1}]$ the area that can be worked when using equipment e for work type w for an hour, and $1^{LSq}_{l,s_q}$ being an indicator function that links land use activities to soil resistance classes.

The amount of work with own machinery that can be done in a work season is limited by the number of equipments and tractors owned by the agent, the amount of labor available and the number of days with suitable weather for the type of work to be done. KTBL [2010] provides a division of Germany into climatic regions and an estimate of expected days for field work of different weather sensitivity levels (k) for each region and half-month of the growing season. Following this approach, we distinguish five levels of weather sensitivity: (i) cereal harvest; (ii) hay harvest (soil dried); [(iii) hay harvest (shed dried)] [2]; (iv) harvest of grass silage; (v) medium sensitive activities such as harvest of silage maize, mineral fertilization, and sowing; and finally (vi) less sensitive activities such as organic fertilization and incorporation of crop residues into the soil. Based on this we calculated $a^{TK}_{t,k}$, the available hours for field work of level k in work season t by assuming a certain amount ($\zeta_{\text{hoursaday}}$) of work hours per day.

The indicator function $1^{WK}_{w,k}$ links every type of field work to the corresponding weather sensitivity level, but also to all levels representing lower weather sensitivity: A day that is suitable for field work of level (v) is also suitable for work of level (vi), and if a person performs level (v) work it cannot simultaneously perform level (vi) work.

Equipment capacity The capacity constraints for equipment e are then represented by the following linear equations (2.10), with b^{E}_{e} being the number of equipments of type e owned.

$$\begin{aligned} \sum_{w,s_q} \left(1^{WK}_{w,k} x^{WTSqE}_{w,t,s_q,e}\right) \quad - a^{TK}_{t,k} x^{tE}_{e} \quad &\leq \quad 0 \quad \forall (k,t,e) \\ x^{tE}_{e} \quad &\leq \quad b^{E}_{e} \quad \forall e \end{aligned} \tag{2.10}$$

Tractor capacity In a similar fashion, field work is constrained by available tractor capacity. The major difference is that a 83 kW tractor can, of course, also be used for work, which requires only 45 kW of tractor power, although fuel consumption and variable cost will be higher than for a less powerful tractor. To take account of this fact we introduced tractor capacity balances and tractor power balances. The capacity of a tractor type is calculated in a similar way as equipment capacities,

$$\begin{aligned} x^{ZTK}_{z,t,k} \quad - a^{TK}_{t,k} x^{tZ}_{z} \quad &\leq \quad 0 \quad \forall (k,t,z) \\ x^{tZ}_{z} \quad &\leq \quad b^{Z}_{z} \quad \forall z \end{aligned} \tag{2.11}$$

---

[2] Currently not considered

and feeds into the corresponding tractor power balance for each combination of work season and weather sensitivity. The transfer activities x^{PPTK} allow the use of higher tractor power for less power-demanding work, too.

$$\sum_{w,s_q,e}\left(1^{WK}_{w,k}\, a^{WSqEP}_{w,s_q,e,p}\, x^{WTSqE}_{w,t,s_q,e}\right) \quad -\sum_{z}\left(1^{ZP}_{z,p}\, x^{ZTK}_{z,t,k}\right) \qquad -x^{PPTK}_{p+1,t,k} \quad +x^{PPTK}_{p,t,k} \quad \leq \quad 0 \quad \forall(k,t,p) \tag{2.12}$$

In any case, fuel consumption ($a^{ZG}_{z,g}$) and maintenance cost (c^{Z}_{z}) per hour of tractor use ($x^{ZTK}_{z,t,k}$) depend only on the type of tractor used, not on the type of work done with it. So, whenever available, a smaller tractor will be preferred over a heavier one if both can do the work in question.

Labor capacity As a third restriction, the farm needs to be able to muster the necessary amount of labor during the suitable days, which is calculated similarly to equipment and tractor capacity (see section 2.5.7).

Restrictions on hiring labor and machinery It is debatable and therefore left to sensitivity analysis and open discussion at this point, in how far contracting of field work is restricted by days with suitable weather. Of course, also contracted work can only be done with suitable weather, however, how much work can actually be done depends on whether the farmer is able to find a provider with open capacity and available capacity of the provider.

$$\begin{aligned} x^{btHTK}_{t,k} \quad & \leq \quad b^{btHTK}_{t,k} \quad \forall(t,k) \\ x^{bWTSq}_{w,t,s_q} \quad & \leq \quad b^{bWTSq}_{w,t,s_q} \quad \forall(w,t,s_q) \end{aligned} \tag{2.13}$$

The capacity b^{bWTSq}_{w,t,s_q} is calculated as follows:

$$b^{bWTSq}_{w,t,s_q} = \zeta_{\text{proptohire}} * \zeta_{\text{hoursaday}} * a^{TK}_{t,k} * 1^{WK}_{w,k} \tag{2.14}$$

The parameter $\zeta_{\text{proptohire}}$ is subject to calibration.

2.5.4 Animal husbandry

We considered cattle and pig related animal production activities. For cattle, we distinguished dairy production, calf raising (0-3 months, male and female), heifer raising (3 - 30 months), bull fattening (3 - 18 months) and suckler cows. For pigs, we considered piglet production (< 8 kg), piglet raising (8-28 kg) and pig fattening (28-118 kg).

In the model, each animal production activity ($a \in A$) is associated with a decision variable (x_a^A) in the MIP, which indicates the number of stable places $[\mathrm{sp}]$ used. For dairy cows, we included two different production levels (3000 kg/a and 8000 kg/a). Due to the linear nature of the MIP, any production level between the two levels included is (theoretically) achievable by a linear combination of the two. For all other animal production activities, we just included one standard specification. The standard specification defines the duration of a turn-over and the quantity produced, and correspondingly nutrition requirements, input use, manure production and work and infrastructure requirements.

Animal production activities usually produce several products including live animals, which form an input for other animal production activities. E.g. dairy production apart from milk supplies male and female calves. Live animals and nutritional inputs, and heating are treated in explicit balances in the MIP, while all other inputs (e.g. water, straw, veterinary costs, insurance) are multiplied by prices and aggregated to direct cost, which enter the MIP as the objective function coefficient (c_a^A) of the corresponding animal production activity (x_a^A).

The part of the MIP which links animal production activities to the balances of its products and the objective function, looks like this:

$$\begin{array}{cccccl}
\sum\limits_g c_g^{sG} x_g^{sG} & -\sum\limits_g c_g^{bG} x_g^{bG} & -\sum\limits_a \left(c_a^A x_a^A\right) & +\ldots & \rightarrow & \text{obj} \\
x_g^{sG} & -x_g^{bG} & -\sum_l \left(a_{a,g}^{AG} x_a^A\right) & & \leq & 0 \quad \forall g \in Ga
\end{array} \tag{2.15}$$

Nutrition

Each animal production activity ($a \in A$) requires the provision of certain quantities ($a_{a,n,d}^{AND}$) of selected basic nutrients ($n_m \in N_m$). Nutrients considered are metabolizable energy (ME), raw protein (XP) and lysine for pigs, net energy lactation (NEL) and usable raw protein usable raw protein (nXP) for dairy cows, and metabolizable energy (ME) and usable raw protein (nXP) for other cattle. For cattle, it has also to be made sure that the raw fibre content of the fodder ration is high enough, and for dairy cows standard limits on structure value (SV, de Brabander et al. 1999), sugar and starch content, raw fat content and ruminal nitrogen balance (RNB) are applied. Nutrition demand of animals was taken from KTBL [2010] and LfL [2010, 2011].

Balances for these nutrients are distinguished for feeding groups ($j_a \in Ja$), which comprise several animal production activities, and for six feeding seasons ($d \in D$).

The agent is free to choose any suitable combination of bought or self-produced fodder in order to satisfy the nutrient demand of its animals. The feeding decision is represented in the MIP by the vector of feeding activities ($x^{fGJaD}_{g,j_a,d}$), which indicate the quantity of a product g fed to a feeding group j_a in feeding season d. Feeding activities, obviously, form part of the product balance:

$$-x^{bG}_{g} \quad -\sum_{l}\left(a^{LG}_{l,g}x^{L}_{l}\right) \quad +\sum_{j_a,d}x^{fGJaD}_{g,j_a,d} \quad \leq \quad 0 \quad \forall g \tag{2.16}$$

Each feeding activity is associated with coefficients (a^{GJaNb}_{g,j_a,n_b}), which determine the quantity of the respective nutrients (n) in each unit of the good g fed to animal group j_a. Values were taken from KTBL [2010] and LfL [2010]. In order to ensure a healthy diet, the nutrient demand and supply should be balanced for each animal group in each feeding seasons, allowing for sufficient fodder with the right mixture of nutrients. To allow for some flexibility in solving the model, a minimum and a maximum constraint was included instead of an equality, allowing a 1% oversupply per nutrient:

$$\begin{array}{rrcl}
\sum_{a}\left(1^{AJa}_{a,j_a}a^{AND}_{a,n_b,d}x^{A}_{a}\right) & -\sum_{g}\left(a^{GJaNb}_{g,j_a,n_b}x^{fGJaD}_{g,j_a,d}\right) & \leq & 0 \quad \forall(n_b,j_a,d)\\
-1.01\sum_{a}\left(1^{AJa}_{a,j_a}a^{AND}_{a,n_b,d}x^{A}_{a}\right) & +\sum_{g}\left(a^{GJaNb}_{g,j_a,n_b}x^{fGJaD}_{g,j_a,d}\right) & \leq & 0 \quad \forall(n_b,j_a,d)
\end{array} \tag{2.17}$$

For cattle (with the exception of calves), a healthy diet requires a minimum raw fibre content of 18% of dry matter fed. For dairy cows, sugar and starch content should not surpass 28% and raw fat content should lie below 4%, while the structure value (SV, de Brabander et al. 1999) should at least reach an average of 1.2 per kg dry matter following recommendations of LfL [2010].

These restrictions have been implemented in the model using the following system of equations, distinguishing between those nutrients with an upper limit (N_u) on dry matter share and those with a lower limit (N_l). The coefficients a^{GJaNl}_{g,j_a,n_l}, resp. a^{GJaNu}_{g,j_a,n_u} indicate the nutrient content (% of dm) of feedstock g, while a^{GJaNd}_{g,j_a} indicates its dry matter content. Values were obtained from LfL [2010]. x^{tNd} are transfer activities used to close the equations, and the coefficients $a^{Nl}_{n_l}$ and $a^{Nu}_{n_u}$ represent the lower, respectively upper limits imposed on dry matter share for each nutrient.

$$\begin{array}{rrcl}
-\sum_{g}\left(a^{GJaNl}_{g,j_a,n_l}x^{fGJaD}_{g,j_a,d}\right) & +a^{Nl}_{n_l}x^{tNd}_{j_a,d} & \leq & 0 \quad \forall(j_a,d,n_l)\\
\sum_{g}\left(a^{GJaNu}_{g,j_a,n_u}x^{fGJaD}_{g,j_a,d}\right) & -a^{Nu}_{n_u}x^{tNd}_{j_a,d} & \leq & 0 \quad \forall(j_a,d,n_u)\\
\sum_{g}\left(a^{GJaNd}_{g,j_a}x^{fGJaD}_{g,j_a,d}\right) & -x^{tNd}_{j_a,d} & \leq & 0 \quad \forall(j_a,d)
\end{array} \tag{2.18}$$

The ruminal nitrogen balance for dairy cows is restricted to lie between 0 and 30 g per day, by the following constraints, in which $a^{NrDl}_{n_r,d}$ and $a^{NrDu}_{n_r,d}$ represent the lower, respectively upper limit to the ruminal balance in each feeding period d.

$$\begin{array}{rlll} a^{NrDl}_{n_r,d} \sum_a \left(1^{AJa}_{a,j_a} a^{AND}_{a,n,d} x^A_a\right) & -\sum_g \left(a^{GJaN}_{g,j_a,n} x^{fGJaD}_{g,j_a,d}\right) & \leq 0 & \forall (n_{\neg s}, j_a, d) \\ -a^{NrDu}_{n_r,d} \sum_a \left(1^{AJa}_{a,j_a} a^{AND}_{a,n,d} x^A_a\right) & +\sum_g \left(a^{GJaN}_{g,j_a,n} x^{fGJaD}_{g,j_a,d}\right) & \leq 0 & \forall (n_{\neg s}, j_a, d) \end{array} \tag{2.19}$$

Work, machinery and infrastructure: services for animal production

Animal production is usually subject to economies of scale, because work time required per stable place and cost of infrastructure per animal decline with an increasing number of stable places. In the model, labor, infrastructure and machinery use of animals are subsumed under the term services. Each animal production activity (a) requires certain quantities of different types of services. The required type of service ($v \in V$) is given as a scale independent coefficient $a^{AV}_{a,v}$ per stable place, e.g. for dairy cows it just indicates that every stable place used requires the capacity to milk one cow.

These service requirements can be satisfied by provisions of services ($o \in O$), which may require labor, machinery or infrastructure capacity ($m \in M$), cash or again other service types. E.g. milking with a herringbone milking parlor requires labor and the capacity of a corresponding parlor. Economies of scale are represented in the model by assuming that for (most) services, the quantity of labor, cash or other inputs can be represented by a linear equation with a binary fixed ($x^{\alpha O}$) and a proportional ($x^{\beta O}$) component, resulting in a decreasing specific input demand function with increasing production volume. A similar assumption holds for investments into infrastructure ($x^{\alpha iM}, x^{\beta iM}$), which also necessitates the inclusion of fixed and proportional capacities ($b^{\alpha iM}, b^{\beta iM}$). Services related to feeding are represented by separate balances for each feeding season $d \in D$.

The corresponding system of equations is shown in block 2.20.[3]

$$\begin{array}{rrrrll} \sum_a \left(a^{AV}_{a,v} x^A_a\right) & & -\sum_o \left(a^{OV}_{o,v} x^{\beta O}_o\right) & -\sum_m \left(a^{\beta MV}_{m,v} x^{\beta tM}_m\right) & \leq & 0 \quad \forall v \\ & -M x^{\alpha O}_o & +x^{\beta O}_o & & \leq & 0 \quad \forall o \\ & & & x^{\beta tM}_m & \leq & b^{\beta M}_m \quad \forall m \end{array} \tag{2.20}$$

Some services are specific to the type of fodder fed to the animals – e.g. feeding of silage requires a totally different type of work than pasturing – and are thus associated with the feeding rather than the animal production activities. This also makes it necessary to disaggregate related services and service types by feeding season (d), and requires the distinction between time-specific (Od, Vd) and non-time-specific (On, Vn) services and service types. Cutting of fresh grass requires field work, and pastur-

---

[3] Note: Some services/infrastructure items have no independent part, while others have a fixed size. Equations 2.20 include terms for service provision of the fixed part, too, which have been omitted for clearer exposition here.

ing requires labor in certain field work seasons. (These details have been omitted from equation block 2.20.)

2.5.5 Biogas production

Maize, wheat and grass silage as well as manure can be used in fermenters to produce biogas, which is then transformed to heat and electricity in generators. The production of biogas electricity from specific goods is represented by the decision variables x_g^{uG}. Electricity yields a_g^{uG} are specific to the feedstock used.

Production of biogas from a certain feedstock is obviously constrained by production or purchase of this feedstock,

$$x_g^{uG} - \sum_l \left(a_{l,g}^{LG} x_l^L \right) - \sum_a \left(a_a^{AG} x_a^A \right) - x^{bG} \leq 0 \quad \forall g, \tag{2.21}$$

as well as the total electric (b^{Ue}) capacities installed. Similar to other infrastructure, we also split biogas plants into a fixed ($b^{\alpha Ue}$) and a variable ($b^{\beta Ue}$) part to reflect economies of scale in investment size in the model. Use of capacity requires maintenance (x^{mUe}), with size-dependent monetary maintenance cost ($c^{\beta mUe}$) and size-independent daily maintenance work ($a^{H\alpha mUe}$).

$$\begin{array}{llllr} \sum_g \left(a_g^{GUe} x_g^{uG} \right) & -x^{\beta mUe} & & \leq & 0 \\ & x^{\beta mUe} & & = & b^{\beta Ue} \\ & x^{\beta mUe} & -Mx^{\alpha mUe} & \leq & 0 \\ & & x^{\alpha mUe} & = & b^{\alpha Ue} \end{array} \tag{2.22}$$

Biogas production (x_g^{uG}) further requires constant daily labor and process electricity, which is reflected by including corresponding coefficients (a_g^{uGH}, a_g^{GbG}) for x_g^{uG} in the daily labor, respectively the product balance of conventional electricity.

The German renewable energy act (*Erneuerbare Energien Gesetz* (EEG)) obliges electricity companies to purchase electricity from renewable sources like biogas plants at a fixed price. Prices are guaranteed to the electricity producer for twenty years from the start of electricity production. The individual price paid for a kWh of a certain biogas plant depends on the year the plant first entered production ($y_y \in Yy$) and is tiered by volume. The EEG was first established in 2000 and has been subject to revisions in 2004, 2009 and 2012. Farmers and also agents in the model, who are not not willing to comply with the EEG requirements for receiving the guaranteed prices, can still sell the electricity at market prices (x^{sGe});

EEG 2000-2009

The general mode of tiered payments has not changed between the 2000, 2004 and 2009 versions of the EEG. The EEG 2004 introduced additional boni for the use of energy plants (NaWaRo) and farm manure, as well as combined heat and power generation (KWK), while the EEG 2009 mainly changed the amount of the guaranteed prices and added a specific manure bonus on top of the NaWaRo bonus. In the model, an individual decision variable $x^{sYyYu}_{y_u,y_y}$ reflects the sale of a quantity of electrical energy for the price c^{sYyYu} valid under tier $y_u \in Yu$ for plants established in year $y_y \in Yy$.

$$x^{sGe} + \sum_{y_u,y_y} x^{sYyYu}_{y_u,y_y} - \sum_{g \in Gb} \left(a^{uG}_g x^{uG}_g \right) \leq 0. \tag{2.23}$$

Further, it requires an EEG application corresponding to biogas plant capacity established in the given year ($b^{Yy}_{y_y}$),

$$\sum_{y_u,y_y} x^{sYyYu}_{y_u,y_y} \leq b^{Yy}_{y_y} \quad \forall y_y \tag{2.24}$$

and that the volume allowed under the corresponding tier ($b^{YyYu}_{y_u}$) has not yet been exhausted.

$$x^{sYyYu}_{y_u,y_y} \leq b^{Yu}_{y_u} \quad \forall (y_u, y_y) \tag{2.25}$$

Since all feedstock categories considered in our model fulfill the requirements for the NaWaRo bonus, it is automatically added to the biogas sales price. The manure bonus of EEG 2009 requires a minimum of 30% manure (($g \in Go$)) in the total mass of the feedstock. This condition is implemented using a binary decision of either accepting the condition and receive the bonus (x^{yuo}), or relaxing the condition on minimum manure use (x^{nuo}) and forgo the bonus (eq. 2.26).

$$\begin{array}{llllllcll}
\sum\limits_{y_u,y_y} x^{osYyYu}_{y_u,y_y} & -\sum\limits_{g \in Go} \left(a^{uG}_g x^{uG}_g \right) & & & & & \leq & 0 & \\
 & & \sum\limits_{g} x^{uG}_g & -x^{tuGo} & & -Mx^{nuo} & \leq & 0 & \\
 & & -\sum\limits_{g \in Go} \left(x^{uG}_g \right) & 0.3x^{tuGo} & & & \leq & 0 & \\
\sum\limits_{y_u,y_y} x^{osYyYu}_{y_u,y_y} & & & & -Mx^{yuo} & & \leq & 0 & \\
 & & & & x^{yuo} & +x^{nuo} & \leq & 1 & \\
x^{osYyYu}_{y_u,y_y} & & & & & & \leq & b^{oYyYu}_{y_u,y_y} & \forall (y_u, y_y)
\end{array} \tag{2.26}$$

Apart from the electricity also the heat produced during the burning of biogas, can potentially be sold or used as input for animal production on the farm. The combined use of heat and electricity is rewarded with an additional KWK bonus under EEG 2004 and 2009 (^{sYuh}).

$$\begin{array}{lllllllcll}
 & & & -\sum\limits_{g \in Gb} \left(a^{uGh}_g x^{uG}_g \right) & & +x^{tuh} & & \leq & 0 & \\
 & x^{suh} & & & & -x^{tuh} & +x^{tuh2} & \leq & 0 & \\
 & & -x^{bG}_g & & +\sum\limits_{a} \left(a^{AG}_g x^{A} \right) & & -x^{tuh2} & \leq & 0 & , \quad g = heat \\
x^{yuh} & & & & & -a^{tueh} x^{tuh} & & \leq & 0 &
\end{array} \tag{2.27}$$

Both, manure and KWK bonus can only be rewarded for electricity for which also the base rate is awarded:

$$\begin{aligned} \sum_{y_u,y_y} x^{osYyYu}_{y_u,y_y} \quad - \sum_{y_u,y_y} x^{sYyYu}_{y_u,y_y} &\leq 0 \\ x^{suh} - \sum_{y_u,y_y} x^{sYyYu}_{y_u,y_y} &\leq 0 \end{aligned} \tag{2.28}$$

EEG 2012

The newest revision of the EEG replaces the old system of a base price and boni by introducing two remuneration classes, into which biogas feedstock is classified. The remuneration is granted according to the share of the feedstock classes in the total methane produced. As the remuneration remains tiered (Yu), this introduces a quadratic relationship into the constraints, which has to be resolved using discretization in our mixed integer linear model. We defined remuneration activities (Yz) with fixed relationships between the two remuneration classes ranging from 100% remuneration class I to 100% remuneration class II in steps of 10%. Except for the extremes, we introduced two activities at each step, one ($x^{slYyYuYz}$) serving as the lower bound of a 10% interval and the other as the upper bound ($x^{suYyYuYz}$).

We further defined mutually exclusive binary activities ($x^{ynoYyYz}$), which make sure the boundary activities of only one interval within a tier can be used. In this way, we make sure that the relationship between the remuneration classes is (at least approximately) equal in all tiers. Otherwise the optimization might lead to the remuneration of electricity of one class in the lower tier and of the other one in a higher tier (the relationship between rewards granted for each remuneration class is not the same between the tiers).

$$\begin{aligned} x^{slYyYuYz}_{y_u,y_y,y_z+1} + x^{suYyYuYz}_{y_u,y_y,y_z} - a^{Yu} x^{ynoYyYz}_{y_y,y_z} &\leq 0 \quad \forall y_z, y_u, y_y \\ \sum_{y_z} \left(x^{ynoYyYz}_{y_y,y_z} \right) &\leq 1 \quad \forall y_y \end{aligned} \tag{2.29}$$

The EEG 2012 further restricts the share of maize in the total feedstock mass to 60%,

$$\begin{array}{lllllll} & -\sum_{g} x^{uG}_{g} & +x^{tuGm} & & & \leq & 0 \\ & \sum_{g \in Gm} \left(x^{uG}_{g} \right) & -0.6x^{tuGm} & & -Mx^{n12m} & \leq & 0 \\ \sum_{y_u,y_y,y_z} \left(x^{slYyYuYz}_{y_u,y_y,y_z+1} + x^{suYyYuYz}_{y_u,y_y,y_z} \right) & & & -Mx^{y12m} & & \leq & 0 \\ & & & x^{y12m} & +x^{n12m} & \leq & 1 \end{array} \tag{2.30}$$

and requires the combined use of at least 60% of the heat for plants, whose feedstock consists of less than 60% manure. A special unitary premium (x^{sYx}) is granted for small plants up to 75 kW, which use manure for more than 80% of the electricity production. Together with the manure bonus of EEG 2009, these are combined into a mutually exclusive set of reward options in the model using binary activities,

$$x^{yYx} + x^{y60o} + x^{yuo} + x^{nuo} \leq 1 \tag{2.31}$$

which are used to apply different manure share requirements,

$$\begin{array}{rrrrrcl} \sum_{g} x_g^{uG} & -x^{tYx} & -x^{tu60o} & -x^{tuGo} & -Mx^{nuo} & \leq & 0 \\ -\sum_{g \in Go} \left(x_g^{uG}\right) & +0.8x^{tYx} & +0.6x^{tu60o} & +0.3x^{tuGo} & & \leq & 0 \end{array} \tag{2.32}$$

with

$$\begin{array}{rrrrcl} x^{tu60o} & & -Mx^{y60o} & & \leq & 0 \\ & x^{tuGo} & & -Mx^{yuo} & \leq & 0 \end{array} \tag{2.33}$$

and then allow the use of the respective schemes:

$$x^{sYx} - Mx^{yYx} \leq 0 \tag{2.34}$$

$$a^{rKWK} \sum_{y_u, y_y, y_z} \left(x_{y_u, y_y, y_z+1}^{slYyYuYz} + x_{y_u, y_y, y_z}^{suYyYuYz} \right) - Mx^{y60o} \leq 0 \tag{2.35}$$

(For lack of data on potential heat uses, we only consider the two extreme scenarios that either all or no agents can use all of the available heat. The requirement to use the heat is therefore not explicitly implemented in the model. Under the assumption that all agents have the potential of external heat use, the coefficient a^{rKWK} is set to zero as the condition will be fulfilled per se, while in the other case no biogas plant with less than 60% of manure can be rewarded according to EEG 2012.)

Further the remuneration activities are subject to the same constraints regarding biogas production and the establishment of an EEG contract as the EEG 2000-2009 activities.

$$\begin{array}{llllcll} x^{sYx} & +x^{sGe} & + \sum_{y_u, y_y, y_z} \left(x_{y_u, y_y, y_z+1}^{slYyYuYz} + x_{y_u, y_y, y_z}^{suYyYuYz} \right) & - \sum_{g \in Gb} \left(a_g^{uG} x_g^{uG} \right) & \leq & 0 & \\ x^{sYx} & & + \sum_{y_u, y_y, y_z} \left(x_{y_u, y_y, y_z+1}^{slYyYuYz} + x_{y_u, y_y, y_z}^{suYyYuYz} \right) & & \leq & b_{y_y}^{Yy} & \forall y_y \end{array} \tag{2.36}$$

2.5.6 Manure

The manure balance links land use, animal production and biogas production. Manure produced by animals can be either used in a biogas plant ($x_{g_o}^{uG}$) or directly spread on the field or grassland ($x_{g_o}^{tGo}$).

$$x_{g_o}^{uG} + x_{g_o}^{tGo} - \sum_{a} \left(a_{a, g_o}^{AGo} x_a^A \right) \leq 0 \quad \forall g_o \tag{2.37}$$

Currently, we distinguish only two types of manure ($g_o \in Go$), cattle and pig manure. For simplification, we assume that the residue from biogas production from manure is equivalent to the manure input with respect to fertilization (which seems justified at least with respect to total nitrogen amounts). Residue from biogas production with silage feedstock is transformed into pig and cattle manure equivalents based on nitrogen content, so that the balance for organic fertilization can be formulated as:

$$\sum_{l} \left(a^{LGo}_{l,g_o} x^{L}_{l} \right) - x^{uG}_{g_o} - \sum_{g_b \notin Go} \left(a^{uGGo}_{g_o,g_b} x^{uG}_{g_b} \right) - x^{tGo}_{g_o} \leq 0 \quad \forall g_o \tag{2.38}$$

At the same time, all the manure produced also has to be spread on the field. For computational reasons, we allow a certain slack here in order to give some flexibility to the MIP solver; the corresponding coefficient ζ_{manure} is subject to calibration.

$$- \zeta_{\text{manure}} \sum_{l} \left(a^{LGo}_{l,g_o} x^{L}_{l} \right) + \sum_{a} \left(a^{AGo}_{a,g_o} x^{A}_{a} \right) + \sum_{g_b \notin Go} \left(a^{uGGo}_{g_o,g_b} x^{uG}_{g_b} \right) \leq 0 \quad \forall g_o \tag{2.39}$$

Second, the farm needs to have storage capacity for all manure produced, which is implemented as an infrastructure service (v_{go}) as described in section 2.5.4.

$$\sum_{a,g_o} \left(a^{AGo}_{a,g_o} x^{A}_{a} \right) + \sum_{g_o,g_b \notin Go} \left(a^{uGGo}_{g_o,g_b} x^{uG}_{g_b} \right) - \sum_{m} \left(a^{\beta MV}_{m,v} x^{\beta tM}_{m} \right) \leq 0 \quad , v = v_{go} \tag{2.40}$$

2.5.7 Labor

Labor capacity depends on the number of household members working on the farm ($x^{H1} + x^{H2}$, see section 2.5.14) and hired permanent employees (x^{bpH}). This labor can be either used for the seasonal field work (x^{tHw}), or for constant daily tasks (x^{tHd}), as they typically are required for animal and biogas production.

$$x^{tHw} + x^{tHd} - x^{bpH} - x^{H1} - 0.25 x^{H2} \leq 0 \tag{2.41}$$

Labor capacity in each field work season is calculated like equipment and tractor power capacity, using the available number of field working days expected in each work season ($a^{TK}_{t,k}$). Additionally, temporary labor can be hired (x^{btHTK}_{t,k_i}) on an hourly basis for each work season.

$$\sum_{w,s_q,e} \left(1^{WK}_{w,k} x^{WTSqE}_{w,t,s_q,e} \right) - a^{TK}_{t,k} x^{tHw} - \sum_{k_i \leq k} \left(x^{btHTK}_{t,k_i} \right) \leq 0 \quad \forall (k,t) \tag{2.42}$$

Labor reserved for constant daily labor is multiplied by the assumed amount of daily working hours (a^{Hd}), and is available for animal production and related services as well as biogas production, where we assume that the same tasks have to be realized everyday.

$$\begin{aligned} & - a^{Hd} x^{tHd} \\ & \quad + \sum_a \left(a_a^{AH} x_a^A\right) + \sum_o \left(a_o^{\alpha OH} x_o^{\alpha O}\right) + \sum_o \left(a_o^{\beta OH} x_o^{\beta O}\right) + \sum_m \left(a_m^{\beta MH} x_m^{\beta tM}\right) \\ & \qquad + \sum_g \left(a_g^{uGH} x_g^{uG}\right) + a^{H\alpha mUe} x^{\alpha mUe} \leq 0 \end{aligned} \tag{2.43}$$

2.5.8 Financial activities & liquidity

While the objective function reflects the expected total farm gross margin , which could be negative, the financial balances ensure that the agent cannot use more liquid means than is actually available. Cash available to the agent at the start of the season (b^C) can either be deposited on the bank to earn interest (x^{dC}) or used in the production process (x^{tC}).

$$x^{dC} + x^{tC} \leq b^C \tag{2.44}$$

This mainly concerns expenses for inputs of crop production x^L , which have to be pre-financed. If cash reserves are insufficient for the later, they can be extended by short-term credit x^{bC}.

$$\sum_l a_{l,c}^{LC} x_l^L - x^{tC} - x^{bC} \leq 0 \tag{2.45}$$

Usually, the standing crop can be used as a collateral and extends the credit limit of the farm.

$$x^{bC} - \sum_l a_{l,c}^{LC} x_l^L \leq 0 \tag{2.46}$$

2.5.9 EU CAP premiums

During the time covered in our hindcast simulations, the EU Common Agricultural Policy (CAP) regulations changed several times. The regulations applicable to the first season simulated (1998/99) still date from the 1992 MacSharry reforms. From 1999/2000 on, the changes under the Agenda 2000 applied. Regulations under MacSharry and Agenda 2000 are structurally similar and differ mainly in parameters, which is why their implementation is described in a common subsection. The CAP Mid-Term Review (MTR) of 2003 enacted regulations applying from seasons 2004/05 on, which were only slightly adapted under the CAP Health Check in 2008.

Milk quota

Throughout the whole period considered, milk sales are restricted by the milk quota b^{Yd}.

$$x^s G_{g=\text{milk}} \leq b^{Yd} \tag{2.47}$$

MacSharry reforms & Agenda 2000

Cereal, oilseed and protein crop premium Farmers could apply for area premiums ($y_c \in Yc$) for cereals (wheat, barley, maize), oilseed and protein crops, which were paid per area of crops grown.

$$x_{y_c}^{Yc} - \sum_l \left(1_{l,y_c}^{\text{LYc}} x_l^L\right) \leq 0 \quad \forall y_c \tag{2.48}$$

A certain percentage (10%) of the area to be subsidized had to be set aside (x^{Ys}), with the exemption of farmers applying for premiums on an area equivalent to less than 92 t reference yield ($a_{y_c}^{Yc}$). (The binary variables x^{yYs} and x^{nYs} represent the decision to set-aside land at all.)

$$\begin{array}{cccccrl}
 & x^{Ys} & -\sum_l \left(1_l^{\text{LYs}} x_l^L\right) & & & \leq & 0 \\
0.1\sum_{y_c} x_{y_c}^{Yc} & -0.9x^{Ys} & & -Mx^{nYs} & & \leq & 0 \\
\sum_{y_c}\left(a_{y_c}^{Yc} x_{y_c}^{Yc}\right) & & & & -Mx^{yYs} & \leq & 92 \\
 & & & x^{nYs} & x^{yYs} & \leq & 1
\end{array} \tag{2.49}$$

As a premium (c^{Ys}) was also paid for set-aside areas, farmers had an incentive to set-aside more land than required. This voluntary set-aside could be extended up to 33% of the total subsidized area.

$$\begin{array}{cccrl}
-0.33\sum_{y_c} x_{y_c}^{Yc} & -0.33x^{Ys} & x^{tYs} & \leq & 0 \\
 & x^{Ys} & -x^{tYs} & \leq & 0
\end{array} \tag{2.50}$$

Energy crops (NaWaRo) could be grown on set-aside land, if their use for non-food and non-feeding purpose was ensured. To capture this, we introduced separate NaWaRo product balances for relevant crops into the model. Biogas production is based on the NaWaRo product balances, while selling and feeding is based on the normal product balances. For all concerned production activities, a duplicate was introduced and marked as "production destined for NaWaRo". The yield of these activities is transferred to the NaWaRo product balance. Feedstock can be transferred from the normal product balance to the NaWaRo balance, but not vice versa.

Suckler cow premium and special premium for male cattle A special premium for male cattle (Yb) was granted by the EU for each bull once in its lifetime, and for each ox twice in its lifetime. As we assume a turnover time of 15 months for bull fattening, this results in 0.8 potential premium applications (a_a^{AYb}) per stable place and year.

Table 2.1: *EU area premiums and reference yields 1999-2004*

		Year (of harvest)			
Coefficient	Land use	1999	2000	2001	2002-2004
c^{Yc}[€]	cereals	279	303	324	324
	maize	396	429	459	459
	oilseeds	545	474	421	324
	protein crops	403	384	384	384
c^{Ys} [€]	set-aside	363	310	333	333
a^{Yc} [t/ha]	cereals	5.29	5.14	5.14	5.14
	maize	5.29	7.28	7.28	7.28
	oilseeds	2.97	5.70	5.79	5.79
	protein crops	5.29	5.29	5.29	5.29

$$x^{Yb} - \sum_a \left(a_a^{AYb} x_a^A\right) \leq 0 \tag{2.51}$$

Suckler cow premium (Yo) was granted for each suckler cow every year (a_a^{AYo} = 1).

$$x^{Yo} - \sum_a \left(a_a^{AYo} x_a^A\right) \leq 0 \tag{2.52}$$

The maximum amount of suckler cows to be subsidized was limited by the suckler cow quota (b^{Yo}) owned by the agent.

$$x^{Yo} \leq b^{Yo} \tag{2.53}$$

Further, a combined upper limit for mother cow and special cattle premium was given by the available area used for feeding these animals (x^{tYbo}) after subtracting the area used for feeding any dairy cows of the farmer, respectively model agent ($a \in Am$). Or, alternatively, under a small producer scheme (x^{yYbs}), subsidies for up to 15 livestock unit (LU) could be granted irrespective of feeding area. The amount of livestock unit which could be subsidized per hectare of feeding area (a^{tYbo}) was 2.0 in 1999, and later reduced to 1.9 in 2002 and to 1.8 in 2003. Suckler and dairy cows were counted as 1 LU, fattening bulls as 0.6 LU per stable place.

$$1.0x^{Yo} + 0.6x^{Yb} + 1.0 \sum_{a \in Am} \left(x_a^A\right) - a^{tYbo}x^{tYbo} - 15x^{yYbs} - Mx^{nYob} \leq 0 \tag{2.54}$$

Due to the involvement of dairy cows in equation 2.54, the constraint needs to be relaxed completely in case the agent chooses neither to apply for cattle nor suckler cow premiums (x^{nYob}). This is reflected in the following equations, which also incorporates the choice between the small and regular producer conditions (x^{nYob}, x^{yYob}, x^{nYbs}, x^{yYbs} are binary integer variables).

$$\begin{array}{llllllll} x^{Yo} & +x^{Yb} & & & -Mx^{yYob} & & & \leq 0 \\ & & x^{nYob} & +x^{yYob} & & & & \leq 1 \\ & & & & & x^{nYbs} & +x^{yYbs} & \leq 1 \\ & x^{tYbo} & & & & -Mx^{nYbs} & & \leq 0 \end{array} \tag{2.55}$$

Total forage area is calculated as the sum of all land uses suitable for feeding ruminants (1_l^{LYbo}).

$$x^{tYbo} - \sum_l \left(1_l^{LYbo} x_l^L \right) \leq 0 \tag{2.56}$$

Under the EU regulations of MacSharry and Agenda 2000, cereal area could be counted as forage area, but could then not be used to apply for the crop premium and its yield cannot be used for feeding other animals (e.g. horses or pigs). To account for this in our model, product balances were split up and growing activities duplicated in a similar fashion as for the NaWaRo rule of the set-aside scheme of the crop premium. Agents can use yields from feeding areas only to feed their cattle (or other ruminants), while yields of other areas can be used for selling, biogas production and feeding alike.

Extensification premium In addition to every suckler cow or special male cattle premium granted, an extensification bonus (Ye) could be awarded

$$x^{Ye} - x^{Yo} - x^{Yb} \leq 0 \tag{2.57}$$

if the ratio of livestock unit (a_a^{AYe}, counting only dairy cows, mother cows, heifers, male cattle and sheep) to forage area (x^{tYe}) is less or equal to 1.4,

$$\sum_a \left(a_a^{AYe} x_a^A \right) - 1.4x^{tYe} - Mx^{nYe} \leq 0 \tag{2.58}$$

and at least 50% of this area is pasture .

$$0.5x^{tYe} - \sum_{l \in Lgp} x_l^L \leq 0 \tag{2.59}$$

Again, equation 2.58 has to be relaxed in case the agent does not opt for the extensification premium (x^{nYe}), requiring the following additional restrictions (with x^{yYe} and x^{nYe} being binary integers):

$$\begin{array}{llll} x^{Ye} & -Mx^{yYe} & & \leq 0 \\ & x^{yYe} & +x^{nYe} & \leq 1 \end{array} \tag{2.60}$$

The potential feeding area is calculated from suitable land use activities (1_l^{LYe}), which does not include cereals and oilseeds in this case.

$$x^{tYe} - \sum_l \left(1_l^{LYe} x_l^L\right) \leq 0 \tag{2.61}$$

Slaughter premium for cattle For cattle over eight months of age to be slaughtered or exported outside the European Union a slaughter premium (Yk) was granted under MacSharry and Agenda 2000. This general premium could be topped-up by EU member states according to a fixed budget, which was distributed among all applicants. The corresponding model coefficient a_a^{AYk}, indicating the number of potential applications per stable place and year, is calculated based on the turnover time assumed for the respective animal production activity a.

$$x^{Yk} - \sum_a \left(a_a^{AYk} x_a^A\right) \leq 0 \tag{2.62}$$

A similar premium was granted for slaughtering calves (without top-up), though this was not considered in the present version of the model, as we cannot distinguish whether calves sold by farm agents are directly slaughtered or raised. For similar considerations, the cattle slaughter premium is only considered for fattening bulls, and replaced dairy cows and mother cows.

Table 2.2: *EU animal premiums 1999-2004*

		Year (of harvest)			
Coefficient	Type	1999	2000	2001	2002-2004
c^{Yo}[€]	suckler cow premium	145	163	182	200
c^{Yb}[€]	special premium male cattle	135	160	185	210
c^{Ye}[€]	extensification premium	51.65	100	100	100
c^{Yk}[€]	slaughter premium incl. top-up	0	34	66	100

EU Transition from Agenda 2000 to MTR

In the course of the MTR reforms, the crop-specific area premiums were transformed into payment entitlements, which now allow receiving the new farm premium, irrespective of what is grown on the plots as long as the area is kept in 'good' conditions according to cross compliance regulations [BMELV, 2006].

In 2005, farmers in the EU received four categories of payment entitlements: for grassland (Yg), arable areas (Ya), set-aside (Yf) and special entitlements. Special entitlements could be awarded to producers without any land (e.g. sheep herders), but are currently not considered in the model.

Until 2013, the amount payable for each entitlement was specific for each farmer as it was partly determined based on the crop area premiums, cattle, extensification and suckler cow premiums received between 2000 and 2002. Representing this in a linear model, requires discretization, i.e. including a separate entitlement for each potential value a grassland or arable entitlement could take on (set-aside entitlements have fixed values). We chose to use discretization steps of 25 €, leading to the inclusion of 198 levels of grassland entitlements ($y_g \in [72, 107, \dots, 5000]$) and 188 levels of arable entitlements ($y_g \in [303, 328, \dots, 5000]$).

The transition itself has not been implemented in the model, as we have not found a feasible MIP implementation for this problem yet. (This currently bars us from running dynamic simulations from 2003 to 2007).

EU MTR and Health Check

Beginning with season 2004/2005, agents can receive the single farm payment ($x^{Ya}_{y_a}, x^{Yg}_{y_g}, x^{Yf}_{y_f}$) according to the respective entitlements owned (b^{Ya}, b^{Yg}, b^{Yf}),

$$x^{Yg}_{y_g} \leq b^{Yg}_{y_g} \quad \forall y_g \tag{2.63}$$

$$x^{Ya}_{y_a} \leq b^{Ya}_{y_a} \quad \forall y_a \tag{2.64}$$

Until the EU Health Check in 2007/2008, receiving premiums was bound to the condition that all set-aside entitlements were activated by setting a corresponding amount of area to set-aside.

This condition is enforced in the model by the following equality constraint,

$$x^{Yf}_{y_f} + x^{tnYf} = b^{Yf} \tag{2.65}$$

which is only relaxed (x^{tnYf}) if the agent chooses not to receive single farm payments. x^{yYf} and x^{nYf} are binary integer variables reflecting the decision for, respectively against receiving payments.

$$\begin{array}{cccccll} \sum_{y_g} x^{Yg}_{y_g} & + \sum_{y_a} x^{Ya}_{y_a} & & -Mx^{yYf} & & \leq & 0 \\ & & x^{tnYf} & & -Mx^{tnYf} & \leq & 0 \\ & & & x^{yYf} & +x^{nYf} & \leq & 1 \end{array} \tag{2.66}$$

Set-aside entitlements can only be activated using fallow or NaWaRo land use activities (1^{LYf}_{l}).

$$x^{Yf}_{y_f} - \sum_{l} \left(1^{LYf}_{l} x^{L}_{l} \right) + x^{tYf} \leq 0 \tag{2.67}$$

Grassland and arable entitlements can be activated using any land use that fulfills cross compliance requirements, including set-aside land (x^{tYf}):

$$\sum_{y_g} x^{Yg}_{y_g} + \sum_{y_a} x^{Ya}_{y_a} - \sum_{l} \left(1^{LY\neg f}_{l} x^{L}_{l} \right) - x^{tYf} \leq 0 \tag{2.68}$$

Following the EU regulations, from 2005 till 2009 the payment per entitlement in the model corresponds to the nominal amount of the agent entitlement (e.g. $c^{Ya}_{y_a} = y_a$). After that the payments are gradually adapted such that in 2013 all entitlements of every agent in the region have the same value (*regional target value, rtv*). Specifically, the difference between the individual value of an entitlement in 2009 and the regional target value is calculated, and in the next years the difference is gradually reduced such that agent entitlements worth less than the regional target value gradually increase in value, and higher valued entitlements decrease in value. The value in a given year between 2010 and 2013 is calculated according to the following formula

$$c^{Y\cdot}_{y.,year} = rtv + \psi_{year}(c^{Y\cdot}_{y.,2009} - rtv) \tag{2.69}$$

with ψ_{year} according to the following table [BMELV, 2006]:

Year	2010	2011	2012	2013
ψ_{year}	0.9	0.7	0.4	0

2.5.10 MEKA

The MEKA program rewards farmers with payments for agricultural practices, which contribute to extensification, landscape conservation and environmentally friendly production [MLREV, 2011]. The MEKA catalog – as of 2011 – contains about 30 different measures ($y_m \in Ym$). In our model, we consider only a few selected ones, which can reasonably be modeled with our current setup, and which are related to grassland extensification and crop rotation diversification.

So far, there have been three phases of the MEKA program: MEKA I from 1994 to 1999, MEKA II from 2000 till 2006, and MEKA III from 2007-2013. In each phase, the measures and associated conditions and rewards were revised substantially, and require a separate implementation in our model.

A constant feature throughout all phases has been the general principal of awarding a measure-specific number of points ($a^{Ym}_{y_m}$) per unit ($x^{Ym}_{y_m}$, e.g. ha, animal, farm) included under a certain measure y_m. For each point received (x^{sYm}), the agent in our model is rewarded with c^{sYm} Euro.

$$x^{sYm} - \sum_{y_m} \left(a_{y_m}^{Ym} x_{y_m}^{Ym} \right) \leq 0 \tag{2.70}$$

A minimum amount rewarded (a^{llYm}, in Euro) is required for participation (x^{yYm}, integer) and a maximum of b^{ulYm} Euro can be awarded per agent.

$$\begin{array}{rrcr} c^{sYm} x^{sYm} & & \leq & b^{ulYm} \\ x^{sYm} & -M x^{ylYm} & \leq & 0 \\ -c^{sYm} x^{sYm} & +a^{llYm} x^{ylYm} & \leq & 0 \end{array} \tag{2.71}$$

A second constant feature of the program has been the requirement to commit to the application of a measure for five years.[4] Agent participation in a measure is therefore bound to a corresponding commitment ($b_{y_m}^{Ym}$).

$$x_{y_m}^{yYm} = b_{y_m}^{Ym} \quad \forall y_m \tag{2.72}$$

Or, where the commitment is not a yes or no decision, but covers a specified area:

$$x_{y_m}^{Ym} = b_{y_m}^{Ym} \quad \forall y_m \tag{2.73}$$

MEKA I

During the first MEKA phase, 20 DEM (i.e. $c^{sYm} \approx 10.22€$) were awarded per point, the upper limit was 40,000 DEM ($b^{ulYm} \approx 20,452€$) and the lower limit 100 DEM ($a^{llYm} \approx 51.13€$).

For the first MEKA phase, we considered in our model only the extensive grassland measures listed under chapters 3.1 (use of grassland) and 3.2.2 (limits on the number of grassland cuts) of the MEKA I catalog. For MEKA I, the state of Baden-Württemberg was subdivided into three grassland support areas (“Förderkulissen”), where support was focused on either (i) groundwater protection, (ii) erosion prevention, or (iii) landscape value. For measure 3.1, differences between groundwater protection areas (3.1A) and the other two areas (3.1) applied. This differentiation is due to the fact that grassland conversion was (and is) not allowed in groundwater protection areas .

Extensive grassland (3.1) In groundwater protection areas, participation in measure 3.1A required maintaining an animal-to-land ratio between 0.3 and 1.4 roughage-consuming livestock unit (RLU) per ha of main forage area (MF), and rewards 8 points per ha of grassland of the agent. In other areas, merely maintaining grassland was rewarded with 2 points per ha, respectively 3 points if an animal-to-land ration of less than 1.8 RLU per ha MF was maintained, or 5 points if it was below 1.2 RLU per ha MF. In the model,

---

[4] on completion of the five years, usually a one or two year extension until the end of the phase was offered, if applicable

we implement these regulations accordingly and represent the agent decision to participate in MEKA 3.1 by the binary integers x^{yYm}_{131A}, x^{yYm}_{131h}, x^{yYm}_{131m} and x^{yYm}_{131l}, whereas x^{nYm}_{131} denotes non-participation. These decision alternatives are mutually exclusive,

$$x^{yYm}_{131A} + x^{yYm}_{131h} + x^{yYm}_{131m} + x^{yYm}_{131l} + x^{nYm}_{131} \leq 1 \tag{2.74}$$

and only possible if the agent is part of the corresponding support focus area:

$$x^{yYm}_{131A} \leq Mb^{y31A} \tag{2.75}$$

$$x^{yYm}_{131h} + x^{yYm}_{131m} + x^{yYm}_{131l} \leq Mb^{y31A} \tag{2.76}$$

Depending on the choice, different restrictions on the total farm RLU apply:

$$\begin{array}{llllll} \sum_a \left(a^{Arlu}_a x^A_a\right) & -1.4x^{clru}_{131A} & -1.2x^{clru}_{131l} & -1.8x^{clru}_{131m} & -Mx^{clru}_{131hn} & \leq 0 \\ -\sum_a \left(a^{Arlu}_a x^A_a\right) & +0.3x^{clru}_{131A} & & & & \leq 0 \end{array} \tag{2.77}$$

with

$$x^{clru}_{y_a} - Mx^{yYm}_{y_a} \leq 0 \quad \forall y_a \in \{131A, 131l, 131m\} \tag{2.78}$$

$$x^{clru}_{131hn} - Mx^{yYm}_{131h} - Mx^{nYm}_{131} \leq 0. \tag{2.79}$$

and the transfer variables (x^{clru}) required to equalize the main forage area:

$$\sum_{y_a \in Ym131} \left(x^{clru}_{y_a}\right) - \sum_{l \in Lmf} x^L_l = 0 \tag{2.80}$$

Fulfillment of these conditions allows agents to retrieve the associated bonuses :

$$x^{Ym}_{y_a} - Mx^{yYm}_{y_a} \leq 0 \quad \forall y_a \in \{131A, 131l, 131m, 131h\} \tag{2.81}$$

based on total grassland area of the agent:

$$\sum_{y_a \in Ym131} \left(x^{Ym}_{y_a}\right) - \sum_{l \in Lgg} \left(x^L_l\right) \leq 0 \tag{2.82}$$

Limitation of grassland cutting (3.2.2) Under measure 3.2.2, 1 point per ha was awarded per ha of grassland, whose use had to be restricted to two cuts per year (x^{Ym}_{13222}) ; 2 points for maximum one cut (x^{Ym}_{13221}).

The model implementation of these measures is straightforward:

$$\begin{array}{lllll} x^{Ym}_{13221} & +x^{tYm1322} & -\sum_{l \in Lgg1} x^L_l & \leq & 0 \\ x^{Ym}_{13222} & -x^{tYm1322} & -\sum_{l \in Lgg2} x^L_l & \leq & 0 \end{array} \tag{2.83}$$

MEKA II

During the second MEKA phase, $c^{sYm} = 10€$ were awarded per point, the upper limit (b^{ulYm}) was 40,000 € and the lower limit (a^{llYm}) 100 €.

Participation in any MEKA II measure required maintaining an animal-to-land ratio of 2.5 LU per ha of agricultural area (AA) as an overall condition. We implemented this rule in the model by using two binary integer variables x^{yYm2} and x^{nYm2}.

$$\begin{array}{cccccccc}
x^{yYm2} & +x^{nYm2} & & & & & \leq & 1 \\
 & -Mx^{nYm2} & +\sum_a \left(a_a^{Alu} x_a^A\right) & -2.5x_2^{clu} & & & \leq & 0 \\
-Mx^{yYm2} & & & +x_2^{clu} & & & \leq & 0 \\
 & & & x_2^{clu} & -\sum_l x_l^L & & \leq & 0
\end{array} \tag{2.84}$$

Diversification of crop rotation (A7) Under measure A7 of the MEKA II catalog, diversity in crop production was awarded, requiring the cultivation of at least 4 different crops, each with at least 15% of the total arable area of the farm and a restriction of maize area to 40% of the total arable area. Oilseeds could be counted as crops to fulfill diversification requirements, but no points were awarded for oilseed areas.

The corresponding model implementation therefore requires the inclusion of several binary integer variables: two variables to represent the decision whether to participate (x_{2A7}^{yYm}) or not (x_{2A7}^{nYm}), which are of course mutually exclusive.

$$x_{2A7}^{yYm} + x_{2A7}^{nYm} \leq 1 \tag{2.85}$$

x_{2A7}^{yYm} requires participation in MEKA II in general:

$$x_{2A7}^{yYm} - Mx^{yYm2} \leq 0 \tag{2.86}$$

Then for each group crop (Jym) potentially included in the agent crop rotation and counted for diversification, two binary integer variables indicate whether it has been included ($x_{j_{ym}}^{yJym}$) or not ($x_{j_{ym}}^{nJym}$). The condition of requiring at least four crops with a minimum share of 15% is enforced in the model by the following system of equations (using the soil in rotation variables x_s^{Sr} – see section 2.5.3 – to sum up all arable land):

$$\begin{array}{ccccccccc}
4x_{2A7}^{yYm} & -\sum\limits_{j_{ym}} x_{j_{ym}}^{yJym} & & & & & & \leq & 0 \\
 & x_{j_{ym}}^{yJym} & +x_{j_{ym}}^{nJym} & & & & & \leq & 1 \quad \forall j_{ym} \\
 & & & \sum\limits_s x_s^{Sr} & -x_{2A7}^{t1Ym} & -x_{2A7}^{t2Ym} & & \leq & 0 \\
 & & -Mx_{j_{ym}}^{nJym} & & +0.15x_{2A7}^{t1Ym} & +0.15x_{2A7}^{t2Ym} & -\sum\limits_l \left(1_{l,j_{ym}}^{LJym} x_l^L\right) & \leq & 0 \quad \forall j_{ym}
\end{array} \tag{2.87}$$

Two more variables are needed in the model to distinguish between oilseed area (x_{2A7}^{t1Ym}) and non-oilseed area (x_{2A7}^{t2Ym}). According to MEKA regulations, only the later can be counted to achieve the point:

$$\begin{array}{lllllll}
x^{sYm} & & & -1x^{t2Ym}_{2A7} & & \leq & 0 \\
& -Mx^{yYm}_{2A7} & +x^{t1Ym}_{2A7} & +x^{t2Ym}_{2A7} & & \leq & 0 \\
& & x^{t1Ym}_{2A7} & & -\sum\limits_{l \in Loil} x^{L}_{l} & \leq & 0 \\
& & & x^{t2Ym}_{2A7} & -\sum\limits_{l \notin Loil, Lgg} x^{L}_{l} & \leq & 0
\end{array} \tag{2.88}$$

Further, the MEKA restriction on maize cultivation is implemented as follows:

$$\sum_{l \in Lmai} -0.4x^{t1Ym}_{2A7} - 0.4x^{t2Ym}_{2A7} - Mx^{nYm}_{2A7} \leq 0 \tag{2.89}$$

Extensive grassland (B1, B2, B4) For MEKA II, the distinction of different support focus areas of MEKA I was dropped and support for extensive grassland use was unified in the whole area of Baden-Württemberg. Measure B1 awarded nine points for maintaining grassland use, i.e. abstaining from grassland conversion and maintaining a minimum level of use on all grassland plots, while restricting the animal-to-land ratio below 2 RLU / ha MF. Measure B2 awarded an additional 4 points for maintaining an animal-to-land ratio between 0.5 and 1.4 RLU / ha MF.

These MEKA restrictions are represented in the model by different limit calculation activities (x^{crlu}) in the following two equations,

$$\begin{array}{llllllll}
\sum_a \left(a^{Arlu}_a x^A_a\right) & -2.5x^{crlu}_{2B1} & -1.4x^{crlu}_{2B2} & -Mx^{crlu}_{n2B} & & \leq & 0 \\
-\sum_a \left(a^{Arlu}_a x^A_a\right) & & +0.5x^{crlu}_{2B2} & & & \leq & 0 \\
& x^{crlu}_{2B1} & +x^{crlu}_{2B2} & +x^{crlu}_{n2B} & -\sum\limits_{l \in Lmf} x^L_l & = & 0
\end{array} \tag{2.90}$$

while the choice of the correct calculation activity is a function of the choice of participating or not participating in B1 and B2 (x^{yYm}_{2B1}, x^{nYm}_{2B1}, x^{yYm}_{2B2}, x^{nYm}_{2B2}, all binary integer variables),

$$\begin{array}{llllllll}
x^{crlu}_{2B1} & +x^{crlu}_{n2B} & -Mx^{nYm}_{2B2} & & & & \leq & 0 \\
& x^{crlu}_{n2B} & & -Mx^{nYm}_{2B1} & & & \leq & 0 \\
& & x^{nYm}_{2B2} & & +x^{yYm}_{2B2} & & \leq & 1 \\
& & & x^{nYm}_{2B1} & & +x^{yYm}_{2B1} & \leq & 1
\end{array} \tag{2.91}$$

which then also allows receiving corresponding points depending on the total grassland area owned by the agent:

$$\begin{array}{rrrrrl} -Mx^{yYm}_{2B1} & & +x^{Ym}_{2B1} & & & \leq 0 \\ & -Mx^{yYm}_{2B2} & & +x^{Ym}_{2B2} & & \leq 0 \\ & & x^{Ym}_{2B1} & & -\sum\limits_{l\in Lgg}\left(x^L_l\right) & \leq 0 \\ & & & x^{Ym}_{2B2} & -\sum\limits_{l\in Lgg}\left(x^L_l\right) & \leq 0 \end{array} \tag{2.92}$$

Measure B4 rewarded very extensive grassland use with five points per ha (replacing MEKA I 3.2.2). The existence of at least four out of a catalog of 28 characteristic species was used as an indicator for low intensity grassland use. For simplification, we assumed in the model that grassland activities of intensity level 1 with conservation cuts ($Lb4$) fulfill these MEKA requirements.

$$x^{Ym}_{2B4} - \sum_{l\in Lb4}\left(x^L_l\right) \leq 0 \tag{2.93}$$

Participation in B1 actually was a prerequisite for participation in measures B2 and B4, and itself required participation in MEKA II in general. Further, farmers could not participate in MEKA B2 and at the same time receive the EU Agenda 2000 extensification premium (section 2.5.9).

$$\begin{array}{rrrrrl} x^{yYm}_{2B1} & & & -Mx^{yYm2} & & \leq 0 \\ -Mx^{yYm}_{2B1} & +x^{yYm}_{2B2} & +x^{Ym}_{2B4} & & & \leq 0 \\ & x^{yYm}_{2B2} & & & +x^{yYe} & \leq 1 \end{array} \tag{2.94}$$

MEKA III

As in the second MEKA phase, $c^{sYm} = 10$ € per point were awarded in the third MEKA phase. The upper limit (b^{ulYm}) remained at 40,000 €, while the lower limit (a^{llYm}) was raised to 250 €. The general requirement of restricting the animal-to-land ratio to 2.5 LU /ha AA to participate in the MEKA program was dropped.

Diversification of crop rotation (A2) With measure A2, the MEKA III catalog contained a diversification support similar to measure A7 of MEKA II. In contrast to phase II, points were also awarded for oilseed areas, while fallow/set-aside areas counted as element of the rotation, but no points were awarded for these areas. Compensation was increased to two points for each hectare of arable land of the farmer. The measure could thus be implemented analogous to the implementation of measure A7 of MEKA II (see section 2.5.10) and is not repeated here.

The MEKA III catalog included additional support for a five-part crop rotation (A3), which has not been used in the model so far as it requires at least 5% legumes in the rotation and we currently do not include any legume among the crops eligible to agents. Measure A3 has been implemented in the model for future use, though. and can be activated once legumes are included in the model .

Extensive grassland (B1, B2, B4) The measures supporting extensive use of grassland were revised again (and implemented in the model accordingly): Under measure B1, five points were awarded per ha grassland included under B1, if

- a maximum animal-to-land ratio of 2.0 LU per ha AA was not surpassed;
- at least 5% of the area was cut the first time after 15th July;
- no grassland conversion was performed;
- a number of other restrictions (e.g. no use of chemical plant protection on grassland at the farm, documentation of organic fertilization and use, pasture care) were respected (which cannot be represented in the model currently, however).

Under measure B2, ten points were awarded to farmers (and agents in the model) for each ha of grassland, if

- the animal-to-land ratio was 1.4 LU/ ha AA and between 0.3 and 1.4 RLU/ ha MF
- no grassland conversion was performed;
- and a number of other restrictions (e.g. no use of chemical plant protection on grassland at the farm, no sprinkler irrigation of grassland, no amelioration on grassland, pasture care) was respected (which again cannot be represented in the model currently).

As in reality, agent participation in B1 or B2 is mutually exclusive (x^{yYm}_{3B1}, x^{yYm}_{3B2}, and x^{nYm}_{3B} are binary integers).

$$x^{yYm}_{3B1} + x^{yYm}_{3B2} + x^{nYm}_{3B} \leq 1 \tag{2.95}$$

The restriction on the animal-to-land ratio is implemented in the model using different calculation activities (x^{clu}, x^{crlu}). For the LU to AA ratio the equation system is as follows:

$$\begin{array}{ccccccccc}
\sum\limits_{a}\left(a^{Alu}_{a} x^{A}_{a}\right) & -2.0x^{clu}_{3B1} & -1.4x^{clu}_{3B2} & & & -Mx^{nYm}_{3B} & \leq & 0 \\
 & x^{clu}_{3B1} & +x^{clu}_{3B2} & -\sum\limits_{l} x^{L}_{l} & & & \leq & 0 \\
 & x^{clu}_{3B1} & & & -Mx^{yYm}_{3B1} & & \leq & 0
\end{array} \tag{2.96}$$

, and for the RLU to MF ratio, the equation system is as follows:

$$
\begin{array}{rrrrrrl}
\sum_{a}\left(a_{a}^{Alu}x_{a}^{A}\right) & -1.4x_{3B2}^{crlu} & -Mx_{n3B2}^{crlu} & & & & \leq 0 \\
-\sum_{a}\left(a_{a}^{Alu}x_{a}^{A}\right) & +0.3x_{3B2}^{crlu} & & & & & \leq 0 \\
 & x_{3B2}^{crlu} & +x_{n3B2}^{crlu} & -\sum_{l\in Lmf} x_{l}^{L} & & & = 0 \\
 & & x_{n3B2}^{crlu} & & -Mx_{3B1}^{yYm} & -Mx_{3B}^{nYm} & \leq 0
\end{array}
\tag{2.97}
$$

Receiving points requires fulfillment of the conditions and is restricted by the available grassland area, and in the case of B1 on the additional condition of cutting 5% of the area the first time after the fifteenth of July.

$$
\begin{array}{rrrrrl}
x_{3B1}^{Ym} & +x_{3B2}^{Ym} & -\sum_{l\in Lgg} x_{l}^{L} & & & \leq 0 \\
0.05x_{3B1}^{Ym} & & -\sum_{l\in Lb4} x_{l}^{L} & & & \leq 0 \\
x_{3B1}^{Ym} & & & -Mx_{3B1}^{yYm} & & \leq 0 \\
 & x_{3B2}^{Ym} & & & -Mx_{3B2}^{yYm} & \leq 0 \\
-x_{3B1}^{Ym} & & & +x_{3B1}^{yYm} & & \leq 0 \\
 & -x_{3B2}^{Ym} & & & +x_{3B2}^{yYm} & \leq 0
\end{array}
\tag{2.98}
$$

In MEKA III, measure B4 was continued similar as in MEKA II, though six points were awarded per hectare of late-cut grassland and participation was not conditioned on participating in measure B1 anymore. The implementation in the model is analogous to the implementation of measure B4 of MEKA II (see section 2.5.10) and not repeated here.

2.5.11 Investments

For the agent investment decision, the production problem described in the previous subsections, is augmented by investment activities (x^{iB}). These include investments into tractors (x^{iZ}) and equipments (x^{iE}), the fixed and size-dependent part of biogas plants ($x^{i\alpha Ue}, x^{i\beta Ue}$) and infrastructure ($x^{i\alpha M}, x^{i\beta M}$), but also other decisions with effects that last longer than one season: for example, the 5-year commitment to a MEKA measure (x^{iYm}), and the right to sell biogas electricity at the guaranteed prices of the current year for the next twenty years (x^{iYy}).

In general, every investment relaxes the corresponding capacity constraint in the agent decision problem,

$$
\cdots - a_{b}^{iB}x_{b}^{iB} \leq b_{b}^{B}\forall b \tag{2.99}
$$

and after the decision is taken b_{b}^{B} will be increased by $a_{b}^{iB}x_{b}^{iB}$ before entering production decision stage for the current year.

Agent investments into assets which are split into fixed and size-dependent part usually are subject to the following condition:

$$x_b^{i\beta B} - Mx_b^{i\alpha B} \leq 0 \forall b \in M, Ue \tag{2.100}$$

In the investment decision, the production problem has been formulated for an average year in the near future. The objective function represents the annualized total farm gross margins of the next years, and consequently the objective function coefficients of the investment activities are the annualized investment cost, calculated as

$$c_b^{iB} = -\left(\eta_b \frac{a_b^B}{\lambda_b} + (1-\eta_b) a_b^B \iota_f \frac{(1+\iota_f)_b^\lambda}{(1+\iota_f)_b^\lambda - 1}\right) \tag{2.101}$$

with η being the share of the investment paid from equity, a^B the investment cost, λ the lifetime of the asset, and ι_f the interest rate on borrowed capital.

Investments are restricted by the liquid means available to the agent,

$$\sum_b \left(\eta_b a_b^B x_b^{iB}\right) - x^{tC} \leq 0 \tag{2.102}$$

further the continuous cash demand by equity fixed in the asset is considered in the general liquidity restriction (eq. 2.45):

$$\sum_l a_{l,c}^{LC} x_l^L - x^{tC} - x^{bC} - \sum_b \left(a_b^{BC} x_b^{iB}\right) \leq 0 \tag{2.103}$$

with the corresponding coefficient (a^{BC}) being calculated as

$$a_b^{BC} = \eta_b a_b^B \left(\frac{(1+\iota_e)_b^\lambda}{(1+\iota_e)_b^\lambda - 1} - \frac{1}{\lambda_b \iota_e}\right) \tag{2.104}$$

with ι_e being the discount factor applied to equity, which is assumed to be equal to the interest rate on short-term deposits ($\iota_e = c^{dC}$)

2.5.12 Post harvest decisions

After harvest, the production decision problem is solved again, with all land use activities fixed at the areas determined in the pre-season production decision, all expected yields replaced by the actual yields obtained and all expected prices replaced by the actual prices realized in the markets. This gives the agent the opportunity to adapt the production plan to the actual results from production: Buying less or more feedstock on the market, or increasing or reducing animal or biogas production.

In the case of perfect foresight of prices and yields, this step can be omitted as the agent harvest results correspond to their expected pre-season values.

2.5.13 Expectations & learning

The current version of the model does not include any updating of expectations or other form of learning.

2.5.14 The farm household and farm succession

All farms are modeled as family farms, as this remains the predominant form of farming enterprise in the study area.

Composition of the farm household

Each agent household in our model consists at least of one male or female household head, the farm manager. Further, it may comprise the farm manager's spouse, their children, a retired household head and his/her spouse (usually the household head's parents), and in some cases also siblings of the household head.

New members enter the agent household either by birth or by marriage. All female household members between 15 and 49 have a positive probability of giving birth. A newborn household member is randomly assigned a gender and a career path as young farmer or young non-farmer. Non-farmer members have no interest in farming as their profession. They may work on the agent farm between the age of 14 and 19, but leave the agent household with 20 years. Young farmer members, on the other hand, are eligible to succeed the current household head once they surpassed the age of 22. Whether they are employed on the agent farm or work somewhere else is part of the agent production decision. If they are over 23 and employed on the agent farm, they have to be paid and their remuneration is accounted as labor cost. A young farmer member, who did not become household head automatically retires at the age of 65 and becomes ineligible for employment on the farm.

All unmarried household members (except young non-farmers, children and seniors above 70 years) have a positive probability of marrying. The status of new household members marrying into the agent household is determined in the model by his/her spouse: The spouse of a young farmer household member will be a young farmer household member, the spouse of a retiree is a retiree, and the spouse of the household head is the spouse of the household head.

Like marriage and giving birth, also the death of agent household member is determined randomly based on their current probability of dying, which depends on their gender and current age. Fertility, mortality and marriage probabilities have been calculated using destatis [2012a,b,c]. The probability for a male newborn household member to be interested in farming (*potsuc_prob_male*) is assumed to range between 0.5 and 1, while the probability for a female newborn household member to become a young farmer is only 0.1, unless the household has a female household head, in which case it is 0.5. (This gender bias follows the patterns observed e.g. by Mann 2007)

Labor provided by other farm members is not accounted for as labor costs and is remunerated through the agent farm income. The minimum household consumption in the model is 26,000 Euro for the household head and 8,000 Euro for each retiree (former household head or spouse of household head). If the agent income is higher than the minimum consumption, a certain percentage determined by parameter *sconextra* of the additional income is consumed in addition. When agent income falls below minimum consumption, it is consumed entirely.

In our model, household heads and retirees until the age of 70 as well as young farmer household members over the age of 19 count as full workers. Household members between 14 and 18 years of age have a labor provision of 30% of a full worker. Spouses of household heads provide 60%, retirees between 70 and 75 years 50% and retirees between 76 and 80 years 20% of a full worker.

Household head succession

Mann [2007] groups the factors influencing the decision to take over a farm business into identity-related and environmental factors. Following this concept, we present the identity-related factors in our model by the distinction between young farmer and young non-farmer household members, which is modeled as a purely statistical relationship.

The environmental factors are mainly related to the economic situation of the farm, and determine whether a potential successor, who is generally interested in farming, finds it worthwhile to take over the farm once the decision has to be taken. In our model, farm succession, i.e. passing the responsibility of the household head to another household member, can be triggered by either death or retirement of the current household head. Succession requires the availability of a potential succeeding household member and the fulfillment of certain economic preconditions for the successor to accept the succession. These preconditions, the eligibility of household members for succession and the consequences of an unsuccessful succession depend on the event which triggered the attempt for succession.

In the case of the death of the current household head, potential candidates for succession in our model are the young farmer household members with at least 23 years of age and the spouse of the deceased household head, in case this member is not older than 65 years. Succession succeeds if an income of *suc_mincons* $*$ minimum household consumption is achieved. If succession fails in the case of death, the agent farm will be shut down.

Two different cases of retirement of the current household head are distinguished in the model: Between the age of 55 and 70, household heads may make a voluntary attempt to retire. Household heads between 55 and 64 attempt to retire in a given year with a probability of 10%, household heads above 65 will attempt to retire every year. Only young farmer household members with at least 23 years of age are eligible for succession and these will succeed only if they have been employed on the agent farm, and the agent income covers at least *suc_mincons* $*$ minimum household consumption of the farm household after succession. If succession fails, the current household head will remain farm manager and try to retire later. Household heads above 70 are obliged to retire, and if they do not find a successor or the later does not accept, the agent farm is shutdown.

Only young farmer household members of at least 23 years of age are eligible for succession and these will succeed if the farm income covers at least emphsuc_mincons $*$ the minimum consumption of the farm household after succession.

This tiered system of retirement implemented in our model is intended to make succession independent of a casual bad year, and let agents choose a suitable situation for succession.

If several potential succeeding household members are available, the one with the highest priority becomes the new household head. The priority ranking is as follows: the oldest male young farmer household member between 23 and 45 years has highest priority, followed by the youngest male young farmer between 46 and 65 years, the oldest female young farmer between 23 and 45 years, and the youngest female young farmer between 46 and 65 years. In the case of death of the current household head, the spouse of the deceased household head follows with lowest priority.

Influence on investment and production decisions

Apart from determining a potential closing of the agent farm business due to the death or retirement of the farm manager, labor provision and the household consumption, the household composition also affects the agent production and investment decisions in two other ways:

First, employment of a young farmer household member – though considered labor cost in the financial accounting of the agent farm – is not considered a cost by the farm manager during planning, as soon as the minimum consumption of the household is expected to be covered by the agent farm income. This model implementation reflects the empirical observation, that farm managers actually tend to enhance their business in order to be able to employ their potential successors, potentially even reducing their own income.

This condition is implemented in our model by distinguishing between ordinary household labor (b^{H1}) and young farmer labor (b^{H2}) . Employing young farmer labor x^{H2} is associated with a cost (c^{H2}), while employing other household labor x^{H1} is not. x^{H2} is an integer activity corresponding to 25% of a full workload, making sure, that the amount of hours worked by the young farmer household member is meaningful. If the total farm gross margin surpasses the sum of minimum consumption, depreciation, rental payments and other fix costs, the cost for employing young farmer household members can be (partially) offset or even overcompensated depending on the value of the parameter ζ_{H2ut}. (The either-or condition is implemented using the two auxiliary integer activities x^{yH2ut}, x^{nH2ut} .)

$$
\begin{array}{lllllllr}
\sum cx & & -0.25c^{H2}x^{H2} & +\zeta_{\mathrm{H2ut}}c^{H2}x^{H2ut} & & & \rightarrow & \mathrm{obj} \\
 & x^{H1} & & & & & \leq & b^{H1} \\
 & & 0.25x^{H2} & & & & \leq & b^{H2} \\
 & & -0.25x^{H2} & +x^{H2ut} & & & \leq & 0 \\
 & & & x^{H2ut} & -Mx^{yH2ut} & & \leq & 0 \\
 & & & & x^{yH2ut} & +x^{nH2ut} & \leq & 1 \\
\sum cx & & -0.25c^{H2}x^{H2} & & & +Mx^{nH2ut} & \geq & b^{Cmc}+b^{Cf}
\end{array}
\tag{2.105}
$$

Second, the age of the household head and the availability of a potential succeeding household member both influence the investment horizon of the farm. To avoid that farm managers close to retirement without successor make investments which pay out only over a long time, the expected remaining farm life is used in the agent investment calculus instead of the expected lifetime of an asset, whenever the later is greater than the former. The expected remaining farm life is the remaining time until the current household head turns 65. We also tested a second implementation, where in the presence of a potential successor, the time until the potential successor will turn 65 is used instead.

2.5.15 Land markets

In the current version of the model, the land market is inactive. Plots available to agents are either owned or rented in from an abstract land owner agent. For land rented in, agents pay a fixed rental payment at the end of the year. In this model version, neither renting in additional land, nor canceling a rental contract, nor renting out own land is considered (although implemented in the model).

Chapter 3

Initial agent populations

Agent populations are initialized based on data from FDZ 2010. The panel includes observations of land use and animal stocks at the farm level for the years 1999, 2003 and 2007. More specifically, the statistical surveys have been conducted always in May of the respective years and reflect production in the cropping seasons 1998/1999, 2002/2003 and 2006/2007

Table 3.1: *Full-time farm classification (FADN)*

Year	Class	Limits
1999	full-time	SBE $\geq$ 15000
	part-time	5000 $\leq$ SBE $\leq$ 15000
	not-represented	SBE $<$ 5000
2003, 2007	full-time	$\geq$ 16 ESU and $\geq$ 1 labour unit
	part-time	$\geq$ 8 ESU, but $\leq$ 16 ESU or $\leq$ 1 labour unit
	not-represented	$<$ 8 ESU

SBE: standard farm income
ESU: European size unit

Our agent population is constituted by the full-time farms for each observation year in the panel dataset. Panel farms were classified into full-time farms and others, using the classification rules used for the German Farm Accounting Data Network (FADN).[1] We restricted our simulations to the group of full-time farms due to constraints on data and resource availability and also because the basic assumption of income maximization in our farm decision model is less convincing for part-time and hobby farmers. Nevertheless, we included the non-full-time farms also in the distribution algorithm in order to achieve more realistic results in the spatial distribution of plots.

---

[1] There has been a change of classification rules between 1999 and 2003 (see tab. 3.1). We used the rules that were valid in each year in order to allow comparability with that year's FADN data.

The FDZ 2010 itself was used to derive marginal and joint distributions of farm areas, arable, grassland and forest shares and livestock numbers. This statistical information was extended by aggregate statistical information and theoretical rules and forms the basis for a step-wise distribution algorithm for agent asset endowments. Next, the farm area allocated to the agent was spatially distributed over the map defining the soil type distribution for each agent. Finally, the demographic composition of farm households was generated based on a random sampling from general demographic information for Germany.

3.1 Estimation of distributions

3.1.1 Estimation of marginal distribution

Marginal distributions for each variable v representing a household characteristic were estimated as empirical inverse cumulative distribution functions $icdf_v(p)$ for each population at a resolution of 0.01. Percentiles for 0% ($p = 0$) and 100% ($p = 1$) had to be excluded due to anonymization requirements. (The statistical office is not allowed to report the minimum and maximum of a variable.)

To arrive at complete continuous distribution functions, we used linear interpolation to infer values between the centiles and imputed values for $p = 0$ and $p = 1$ according to the following rules:

1. $icdf_v(0) = 0 \, \forall \, v$
2. For certain variables representing shares of aggregated land use groups (e.g. share of arable land), we could safely set $icdf_v(1) = 1$.
3. In cases where we had information available on the population total for the specific variable (e.g. for the total agricultural area or the total number of dairy cows), we assumed that due to the construction of the percentiles, summing over the values assigned to N agents (equally distributed over p) should equal the population total observed in the area. Given estimation and interpolation, we knew the values for roughly the first $0.99N$ agents. The values for the remaining $0.01N$ had to add up to the difference between observed total and total of $0.99N$ agents. $icdf_v(1)$ can then be calculated assuming linear progression among the last $0.01N$.
4. For variables where neither theoretical values nor population total was available, we extrapolated $icdf_v(1)$ using the slope of the linear interpolation between $icdf_v(.98)$ and $icdf_v(.99)$

3.1.2 Estimation of the joint distribution

The joint distribution of variables was determined as a frequency distribution of quintile combinations $f(c)$, which similar to a copula link the marginal distributions of the variable using the uniform distributions resulting from a probability integral transformation of the marginal distributions of the variables. In contrast to the canonical form of a copula, we used a frequency distribution instead of a cumulative distribution in this case.

The quintile combination is a vector with dimension V (the number of variables) associated with each household in the panel dataset. Each dimension $v \in \{1, \ldots, V\}$ can take on discrete values $c_v \in \{0, \ldots, 5\}$ indicating under which quintile of the variable v the household falls. E.g. $\mathbf{c} = (1, 3, 1, 5)$ denotes that the associated household falls into the first quintile of the first variable, into the third quintile of the second variable, into the first quintile of the third variable and into the fifth quintile of the fourth variable. A value of zero has been used to aggregate all lower quintiles, whose upper value is zero, in order to reduce the number of distinct vectors. The frequency distribution $f(\mathbf{c})$ reports the number of households h_c associated with a certain vector $\mathbf{c}$.

The privacy restrictions required the choice of quintiles instead of finer quantiles. Further, a full frequency distribution of quintile associations could only be estimated for the vector $\mathbf{d} = (c_1, c_2)$ containing the first two dimensions (total agricultural area of the farm and share of arable land), while the frequency distribution of the complete vector $\mathbf{c}$ could only be estimated from a 85% sample of the entire population.

Comparing the quintiles for the marginals of the 85% sample, with the marginals estimated from the full population, we found them to be an acceptable representation of the full population marginals.[2]

As a result of the sampling, the estimated number of households ($\hat{h}_c$) for a given quintile combination (c) was smaller or equal to the number of households (h_c) that were actually associated with c. This also led to quintile combinations not being reported at all, because their $\hat{h}_c$ is zero, although the true h_c was greater. (According to the FDZ statistician that roughly affected 10% of all c with a $h_c > 0$.) This had to be taken into account during the creation of the agent population by allowing all quintile combinations (including the ones reported as zero) to contain a higher number of agents than reported.

In a first attempt, we largely underestimated the forest area and thus overestimated arable areas and grassland areas. The uppermost quintiles of total area and forest share span relatively wide ranges (for example, 80 - 2393 ha, respectively 10 to 100% for the full-time farms in 1999). Given this result and the expert information that there is no farm with more than 500 ha of agricultural area in the region, we concluded that the largest enterprises are rather forestry than agricultural enterprises and changed quintile distributions such that the uppermost 2% of both the total area and the forest share distributions are now associated with each other. For the out-of-sample farms, we restricted the non-forest area to a maximum of 500 ha. This led to a satisfactory forest area (increase of about 30,000 ha compared to no constraint).

---

[2] This approach works due to the high sampling fraction. For a smaller fraction, one should probably re-estimate the marginals for the sample and later project the quintile association onto the original marginal.

3.2 The distribution algorithm for farm endowments

The estimated distributions were combined with theoretical constraints in order to ensure compatibility with the model. These were also necessary to avoid unrealistic combinations, given the fact that estimated quintile associations delivered only a relatively coarse representation of the joint distribution function and 15% of the agent population was not subject to the full joint distribution at all.

This required a stepwise sampling procedure using different techniques at different steps of the process, which were implemented using `mpmasdist` and are described in the following:

1. We created an agent population of size N and randomly distributed the different observed realizations of the vector d according to its frequency distribution among the agents.
2. We randomly distributed the different observed realizations of the vector c according to its observed frequency distribution among $0.85N$ agents, making sure the first two dimensions of the selected c fit the previously allocated vector d.
3. $0.15N$ agents remained without c imposing no statistical restriction on the joint distribution of characteristics for these agents (except for the farm size and share of arable land reflected in d).
4. Next, looping over farm size quintiles, random farm sizes were distributed among the $0.2N$ agents associated to each respective quintile according to the corresponding partial marginal distribution. A theoretical constraint ensured that the allocated arable land resulting from multiplying the allocated farm size with the minimum share of arable land of the agents defined by c did not surpass 500 ha. The simple, order-based distribution algorithm described in Ch. A.2 was used here and in the following steps unless otherwise noticed in order to ensure covering the full range of the distribution function.
5. Similarly, the arable, grassland and forest shares were allocated within each quintile, making sure that the sum of these was close to one and the resulting non-forest area not greater than 500 ha.
6. At this point, the resulting grassland and arable land ownership was used to spatially allocate plots in the study area to each agent as described in the next section and thus defined the soil composition of the land owned by the agents.
7. Again looping over quintiles, the observed animal numbers were randomly distributed to the agents. The basic restriction is the total animal-to-land ratio, which had to be lower than *gvpha* LU per ha, where *gvpha* was assumed to lie between 2.5 and 3 and subject to calibration. Further, we expected the number of calves and heifers to be characteristically related to dairy cows, and the number of farrows to be dependent on the number of sows, respectively fattening pigs. Specifically, the algorithm used the following steps (separately for the agents with and without associated c vector):

   (a) Dairy cows were randomly allocated, ensuring the animal-to-land ratio was respected, taking into account the expected number of young animals entailed by the number of dairy cows (0.35 calves and 0.35 heifers per dairy stable place), and the minimum numbers of other animals defined by the quintiles associated with the agent.
   (b) Medium-aged cattle and calves were allocated using the Hungarian Method with random component (see Ch. A.2), where the deterministic cost component was set to infinity if the animal-to-land ratio was violated, to zero if the ratio of young animals to dairy cows was greater or equal 0.35, and to $\ln\left(\frac{1}{0.35 dairy - young + 1}\right)$ otherwise.
   (c) Mother cows, horses, fattening pigs, sows and sheep were allocated subsequently ensuring the animal-to-land ratio was respected taking into account the already determined numbers of other animals, respectively the minima defined by the quintiles associated with the agent.
   (d) Other pigs (i.e. mostly farrows) were then distributed using several loops: First, it was attempted to distribute values only to agents, which had both sows and fattening pigs. Then, it was attempted to distribute the remaining values to agents which had sows or fattening pigs. Third, values were allocated to those agents, who neither had fattening pigs nor sows, but were supposed to have farrows. In the first two attempts, assigned values were accepted if they lay in a range of ± 15% of a third of the number of fattening pigs plus 6.21 times the number of breeding sows, reflecting the typical relation of stable places and turnover times of the production activities.

8. The statistical information on livestock randomly allocated to the agents was transformed into model assets:

   (a) The livestock numbers were transformed into corresponding types and quantities of stable capacities.
   (b) For dairy cows, stable places were assumed to be in stanchion stables up to a number of 40 cows, above this cubicle loose-housing stables were allocated. Up to 10 dairy cows, we allocated a bucket milking machine, up to 40 cows a milking pipeline, and above 40 cows usually a herringbone milking parlor. Alternatively, between 60 and 160 cows an automatic milking system (AMS) was allocated with 10% probability and above 160 dairy cows a rotary milking parlor was allocated with 50% probability.
   (c) Agents received milk and manure storage facilities, feeding equipment as well as milk quotas corresponding to the amounts required according to the model assumptions.

9. The number of biogas plants to be allocated in each of the years was inferred based on the results of the farm survey, which asked for the capacity and year of establishment of biogas plants currently installed, and scattered information found in Fachagentur nachwachsende Rohstoffe e.V., Dederer and Messner [2011] and Hartmann [2008]. We intended to allocate 17 biogas plants with capacities

ranging from 75 to 400 kW in 2007, nine biogas plants with capacities ranging between 40 and 420 in 2003 and four biogas plants ranging between 40 and 420 kW in 1999. These biogas plants were randomly distributed among those agents with the theoretical ability to produce feedstock for an electricity production using at least 80% of the plant capacity, taking into account the arable land, grassland and animals owned by the agent.

10. Tractors and other machinery were distributed according to rules developed based on the machinery endowments observed in the farm survey, expert information and model assumptions. The rules related the amount of arable land, grassland, expected manure to be spread and animals owned to certain combinations of tractors and was implemented as shown in table 3.2 in the appendix.

11. Lastly, relevant EU CAP entitlements had to be distributed. Milk quotas were handled as explained above. For the years 1999 and 2003, mother cow quotas were allocated by simply assuming agents own quotas corresponding to the mother cows they own. For 2007, single farm payment entitlements had to be allocated. This was done by allocating grassland, arable and set-aside entitlements according to the land endowments of each agents, and determining their values according to the regulations, assuming agents obtained all premiums they could have potentially received in 2003 given their current (i.e. 2007) asset and land ownership.

Table 3.2: *Rules used to distribute machinery among the agents*

Machinery	Conditions
Tractors	
157, 102, 67 kW	arable > 160 ha, or manure $\geq 3200\mathrm{m}^3$
120, 83, 45 kW	arable > 70 ha or grassland $\geq$ 180 ha
102, 67, 45 kW	arable > 50 ha, or manure $\geq 1200\mathrm{m}^3$
83, 45 kW	arable > 20 ha, or manure $\geq 200\mathrm{m}^3$
45 kW	grassland ≥ 15 ha, or dairy cows ≥ 10, or medium-aged cattle ≥ 20
Tillage and seeding implements	
seeder 2 m, plough 0.7 m	arable 20-50 ha
seeder 3 m, plough 1.05 m	arable 50-160 ha
seeder 4 m, plough 1.75 m	arable > 160 ha
Spraying and fertilizing equipment	
15 m	arable 20-110 ha
24 m	arable > 110 ha
Maize seeder	
	if biogas, or dairy cows > 30, or medium-aged cattle > 50, and …

Table 3.2: *Rules used to distribute machinery among the agents (cont.)*

Machinery	Conditions
3 m	… arable 20-160 ha
6 m	… arable > 160 ha
Manure trailer	
7 m^3	manure 200-1200 m^3
12 m^3	manure 1200-3200 m^3
20 m^3	manure $\geq 3200\mathrm{m}^3$
Manure drag hose	
12 m	manure 2000-3200 m^3
24 m	manure $\geq 3200\mathrm{m}^3$
Machinery combination for grass harvest (mowing, stirring, swathing)	
5 m	grassland ≥ 180 ha
3.2 m	grassland $\geq$ 90 ha, or dairy cows ≥ 104, or medium-aged cattle ≥ 150
2.4 m	grassland $\geq$ 20 ha, or dairy cows ≥ 10, or medium-aged cattle ≥ 20
Round baler	
1.2 m	grassland ≥ 20 ha
Self-loading trailer	
20 m^3	grassland ≥ 20 ha
Loader	
102 kW	if arable > 30 and (dairy cows > 60, or medium aged cattle > 120, or biogas)
Grassland cultivation	
roller 3 m, grass harrow 4 m	grassland 25-50 ha
roller 6 m, grass harrow 9 m	grassland ≥ 50 ha
Combine harvester	
125 kW, 4.5 m	arable 100-180 ha
175 kW, 6 m	arable ≥ 180 ha

3.3 Spatial distribution of farms

We used the CORINE land cover maps – more specifically CLC2000 [2004] for 1999 and 2003, and CLC2006 [2009] for 2007 – providing information on the basic spatial extent of urban, arable, grassland, forest and other natural areas. For our purpose, we aggregated the original 47 land use categories of the CORINE datasets into 13 categories shown in table 3.3.

Table 3.3: *Land use categories used for the spatial allocation of agents*

Code	Description	CLC Codes
0	Urban	111-112
1	Industrial & traffic	121-142
2	Arable	211-213
3	Permanent crops	221-223
4	Pasture	231
5	Mixed cultivation patterns	241, 242, 244
6	Agriculture & natural vegetation mixed	243
7	Forest	311-313, 323-324, 990
8	Heathland	322
9	Natural grasslands	321
10	Wetlands	411-423
11	Water	511-523, 995
12	Rocks	331-335

The spatial distribution of plots proceeded by first randomly distributing farmsteads over the plots classified as urban or arable (0 or 2). Then the forest, arable and grassland area previously determined for each agent was randomly distributed using the `mpmasdist` spatial allocation mechanism, which divides the area owned by an agent into random-sized plots and sequentially places these plots as close as possible to the farmstead or any other previously allocated plots of the agent. Forest area could be placed on plots of category 7 only. While categories 5 and 6 were considered suitable for both arable and grassland, plots were allocated to categories 2, respectively 4 first until all of these were used. Only after that, plots of category 5 and 6 were included into the distribution process.

The distribution mechanism in its current implementation took several days of run time to complete, such that only a limited number of different spatial distributions were generated.

The resulting agent property maps could then be overlaid with the soil maps described in section 4.1 to determine the composition of soil types on each agent's land.

3.4 Household composition

To generate realistic household compositions, we started by randomly determining the age of the household head (age_{hh}) and whether he is married or not based on the statistical distributions reported in destatis [2011] and destatis [2012a]. The age of the household head's wife was drawn from the normal distribution $N(age_{hh}, 2)$. Children were generated by randomly determining whether the household head's wife gave birth for each age between 15 and her current age. The probability of giving birth at each age was taken from destatis [2012b], but was proportionally increased by a factor *birth_factor_past* as the statistical data used cover only a relatively recent period and birth rates in the past have probably been higher. The career path and gender of children were determined using the the same coefficients, which are used in the model (see Sec. 2.5.14). The procedure of determining marriage status and potential descendants (i.e. grand children of the household head) is repeated for each child.

The presence of the household head's retired parents was determined by first individually drawing their potential age from the normal distribution $N(age_{hh} + 28, 2)$ and then using the mortality information from destatis [2012c] to determine whether they actually reached this age or died in the past.

Chapter 4

Exogenous, scenario specific variables

The different policy regulations valid in each of the selected years have already been described in the model description (see 2.5.9 and 2.5.10), and the respective setting was of course chosen for each year. The choice of crop yields, prices, available field work days and crop rotation options is explained in this chapter.

4.1 Crop yields

Aurbacher et al. [2013] used the Expert-N model package to simulate crop yields under current and future climate conditions in order to assess the influence of climate change on crop yields in the study area. The Expert-N model uses the CERES model for winter, wheat, barley, and silage maize, and the GECROS model for winter rapeseed. The study area specific parameterization was calibrated and validated against leaf area index (LAI) and phenological observations at the three field measurements sites measured between 2009 and 2011. As a consequence, the simulated yields reflect current technology, which may cause a bias when used in the calibration and validation process for the 1999, 2003 and 2007 observations. We therefore also considered alternative yield sets derived from public yield statistics in the calibration process in order to avoid overfitting of model parameters to a potentially biased, simulated yield set.

4.1.1 Simulated yields for current climate

Aurbacher et al. [2013] calibrated and validated the Expert-N model against LAI and phenological observations at three field measurements sites measured between 2009 and 2011. The calibrated model was then used to predict yields for each combination of reference soil profile and management for each season between 1951 and 2010 using the corresponding record from the meteorological time series of Stötten weather station. For the present study, the crop modeling team provided us with the results of extended simulations compared to Aurbacher et al. [2013]: The model was calibrated for barley and rapeseed and simulated for all relevant soil classes in the area.

Using LUBW [2007], ten soil mapping units were identified in the study area, which were then linked to eight reference soil profiles to obtain the relevant soil characteristics for modeling. Tables 4.1 and 4.2 give an overview of the importance of each soil mapping unit in the study area and the structure of reference profiles linked to each soil mapping unit.

Yield scenario xn3 uses the long-term average of simulated yields as expected yields for each crop production activity in each of the years ignoring any technology-induced yield difference between the years.

For wheat, we employed a yield reduction of 20% for wheat grown the second year on the same plot compared to wheat grown after other crops. As we assumed the observed or simulated yield to represent the area-weighted average wheat yield in the area, the first year wheat yield is increased using the scaling parameter *wheat_normal*, which is subject to calibration.

Table 4.1: *Overview of soil mapping units and their link to reference profiles*

Soil classification			Share of					
German	WRB	Area [ha]	total	arable	forest	grassland	mixed	Profile ID
Rendzina, Braunerde-Terra fusca, Terra fusca-Parabraunerde	Rendzic Leptosols, Chromic Cambisols, Chromic Luvisols	77,799	60.0%	48.0%	60.4%	68.9%	67.5%	RT
Pararendzina	Calcaric Regosols	5,509	4.2%	5.0%	2.8%	4.7%	5.0%	Z
Braunerde-Pelosol, Pseudogley-Pelosol, Pararendzina	Vertic Cambisols, Vertisols	1,007	0.8%	1.6%	0.5%	0.0%	1.1%	VB
Parabraunerde, Terra fusca, Terra fusca-Braunerde	(Chromic) Luvisols, Chromic Cambisols	5,557	4.3%	12.6%	2.6%	1.6%	0.9%	PT
Parabraunerde, Pararendzina-Braunerde	Luvisols, Cambisols	506	0.4%	0.7%	0.1%	0.3%	0.2%	PB
Parabraunerde	Luvisols	553	0.4%	1.4%	0.2%	0.0%	0.1%	P
Braunerde-Terra fusca	Cambisols, Chromic Cambisols	8,665	6.7%	12.8%	4.1%	3.9%	7.0%	BT
Brauner Auenboden, Auengley	Fluvisols	2,826	2.2%	1.2%	0.6%	5.1%	2.3%	A
Kolluvium	(Cumulic) Anthrosol	10,024	7.7%	14.1%	1.2%	9.7%	10.2%	K
Rendzina, Braunerde-Rendzina	Rendzic Leptosols, Mollic Leptosols	16,802	13.0%	2.5%	27.5%	5.5%	5.6%	RT
Niedermoor	Histosols	97	0.1%	0.1%	0.0%	0.4%	0.0%	-
Ortslage	Build-up areas	400	0.3%	0.2%	0.1%	0.0%	0.0%	-
Total		129,746	100.0%	100.0%	100.0%	100.0%	100.0%	

Table 4.2: *Reference soil profiles used in Expert-N simulations*

ID	Source	Horizon	Lower depth [cm]	Texture	Bulk density [g/cm^3]	Total C$_{org}$ %	C/N
RT	PAK-EC6	Ap1	12	Tu2	1.04	3.314	9.8
		Ap2	21	Tu2	1.29	2.540	10.3
Z	DE-7817-3	Ap	30	Tu2	1.14	2.140	10.7
		2lCv	60	Lts	1.65	0.290	9.7
		mCv	82	Lt2	1.36	0.230	11.5
VB	DE-7619-4	Ah	16	Tu2	1.22	2.950	10.2
		rApBv	35	Tu3	1.23	0.920	8.4
		BvP	65	Tu2	1.21	0.630	7.0
		lCvP	120	Tu2	1.21	0.340	4.9
PT	DE-6623-12	Ah	6	Tu4	0.98	3.300	15.0
		AhAl	26	Tu4	1.31	1.210	15.1
		Bt	41	Tu3	1.41	0.580	9.7
		2T1	58	Tu2	1.42	0.630	10.5
PB	DE-7926-204	Ah	6	Ls2	1.38	3.480	10.2
		Al	30	Ls2	1.41	1.970	12.3
		2Bt	50	Lt2	1.37	0.460	5.8
P	DE-IBS-265	Ah	5	Sl3	1.19	4.500	16.1
		Al	30	Sl4	1.45	1.200	13.3
		Btv	45	Ls4	1.56	0.500	10.0
		Bt	78	St3	1.39	0.100	10.0
		Cv	88	St2	1.39	0.010	10.0
		C	100	Sl3	1.40	0.010	10.0
BT	PAK-EC4	Ap1	21	Tu3	1.31	2.630	9.5
		Ap2	29	Tu3	1.34	1.293	9.8
		Tv	41	Tu2	1.32	0.972	9.1
A	DE-7518-1	rAp	30	Lu	1.27	2.490	8.6
		M1	82	Lu	1.40	1.040	8.0
		M2	140	Lu	1.37	0.010	-
		M3	167	Lu	1.47	0.010	-
K	PAK-EC5	Ap	20	Tu4	1.37	2.171	9.4
		eM1	60	Tu3	1.40	1.063	9.3
		eM2	90	Tu3	1.51	0.380	6.3

4.1.2 Simulated yields for future climate

The plant modeling team simulated yields for the years 2000-2030 using a statistically downscaled projection for the Stötten weather station taken from two realizations from the WETTREG [2010] project. We used the average yield over two WETTREG generalizations and all thirty simulated years as future climate scenario, while the baseline was based on the yields for the years 1981-2010 of the observed time series. Table 4.3 shows the relative yield changes resulting from the simulations compared to the baseline for each crop and soil type.

Table 4.3: *Relative change of yields in climate change scenario*

Soil	Silage maize	Summer barley	Winter barley	Winter rape	Winter wheat
0	-2.6	-6.9	-1.2	-3.5	8.7
1	4.8	-8.5	-1.7	5.3	7.5
2	-3.0	-0.8	-7.0	12.2	20.1
3	0.5	-5.4	-1.7	-1.6	11.7
4	-4.3	11.9	-14.6	6.6	24.6
5	-0.4	-3.2	-2.4	1.0	14.6
6	1.1	-19.7	-1.5	8.5	2.4
7	-2.1	7.2	-10.8	9.3	18.8
8	-4.5	12.1	-12.1	24.1	22.7

4.1.3 Alternative statistical yield sets for calibration-validation

For the calibration and validation of the short-term production decisions it was important to infer the yield farmers' calculated with during their production decision at the beginning of the year, which does not necessarily correspond to the real yield obtained by farmers later in the year.

In a farm survey, conducted between August and October 2010, farmers were asked to describe their expectation for wheat, barley and rapeseed yields as a triangle distribution. In table 4.4, we show mean, standard deviation, minimum and maximum of the modus of the triangle distribution over responding farmers practicing conventional farming on the Central Swabian Jura. We cannot rely on these figures to be statistically representative of all farmers in the study area due to the rather low number of respondents and nature of the survey sample, but these numbers are valuable as a first impression to derive calibration input.

As a second source of information, we recurred to the online database of the statistical office of the state of Baden-Württemberg [Statististisches Landesamt Baden-Württemberg, 2012], which provides yield averages for the two study area districts ranging back until 1983. As depicted in figures 4.1 and 4.2, we observed a long-term trend of increasing yields for most crops, maybe with the exception of summer barley in the

Table 4.4: *Medium yield expectations [dt/ha] in the farm survey (Sep/Oct 2010)*

Crop	N	Avg	Sd	Min	Max
Bread wheat	14	73.9	4.77	67.5	80
Fodder wheat	14	78.6	8.36	70.0	95
Organic wheat	4	41.3	8.10	35	53
Malting barley	5	61.6	11.63	50	80
Winter rapeseed	13	39.7	4.52	30	46

Reutlingen district and silage maize in the Alb-Donau district, combined with considerable interannual variability. Silage maize yields in Reutlingen experienced an abrupt upward shift around 1998 from stable levels below 300 dt/ha to stable levels above 400 dt/ha.

We assumed that farmers' yield expectations average out interannual variability, but do reflect long-term yield development. As an approximation, we calculated the average of the yield of the a years preceding the respective year of harvest for our simulation experiments. Table 4.5 shows the results for $a = 3$ and $a = 6$, reflecting two types of averages, which are rather more and rather less sensitive to short-term fluctuations in observed yields.

Table 4.5: *Yield average of the* a *years preceding the years of observation*

		Reutlingen				Alb-Donau			
	a	1999	2003	2007	2011	1999	2003	2007	2011
Silage maize	3	339.7	421.3	425.0	419.0	479.3	471.7	459.0	503.0
	6	299.8	420.2	418.0	434.8	486.5	478.7	456.7	470.3
Summer barley	3	53.3	46.7	43.3	49.0	51.3	52.3	53.0	57.7
	6	51.3	47.7	45.2	45.8	49.3	51.2	53.2	54.4
Winter barley	3	53.3	59.7	49.7	58.8	64.3	63.7	61.3	64.2
	6	48.5	56.3	54.0	55.4	62.0	63.2	61.8	63.3
Winter rapeseed	3	31.0	34.7	37.7	36.0	33.7	35.0	41.3	38.7
	6	31.0	34.0	34.8	37.5	33.8	34.3	36.8	40.0
Winter wheat	3	63.0	67.7	61.3	65.3	72.3	74.7	77.0	78.9
	6	58.2	65.2	63.3	63.8	69.2	72.7	74.8	77.8

Source: Own calculation based on Statististisches Landesamt Baden-Württemberg [2012]

Yields in the Reutlingen district were consistently lower than in the Alb-Donau district. Despite the fact, that we would have generally expected the Reutlingen district to be more representative of our study area, the farm survey results for wheat, rapeseed and barley seemed to be more consistent with the pre-2011 averages in Alb-Donau rather than Reutlingen. While a slight majority of survey respondents (8 out of 14, resp. 13) is located in the Alb-Donau district, there was no significant difference in yield expecta-

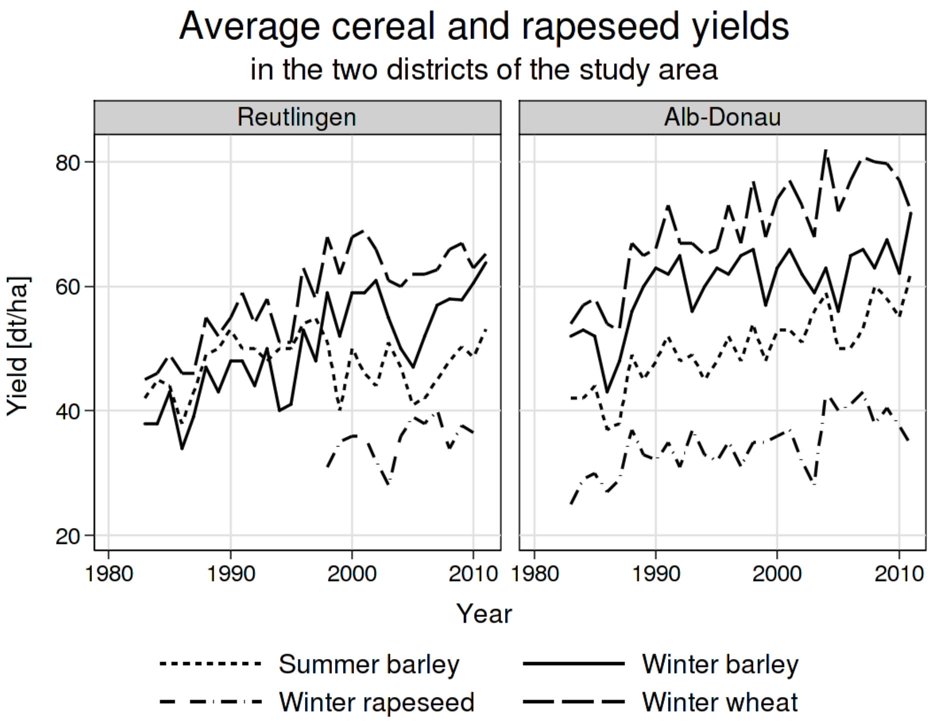


Figure 4.1: *Average wheat, barley and rapeseed yields in the two study area districts, 1983-2011 [Statististisches Landesamt Baden-Württemberg, 2012]*

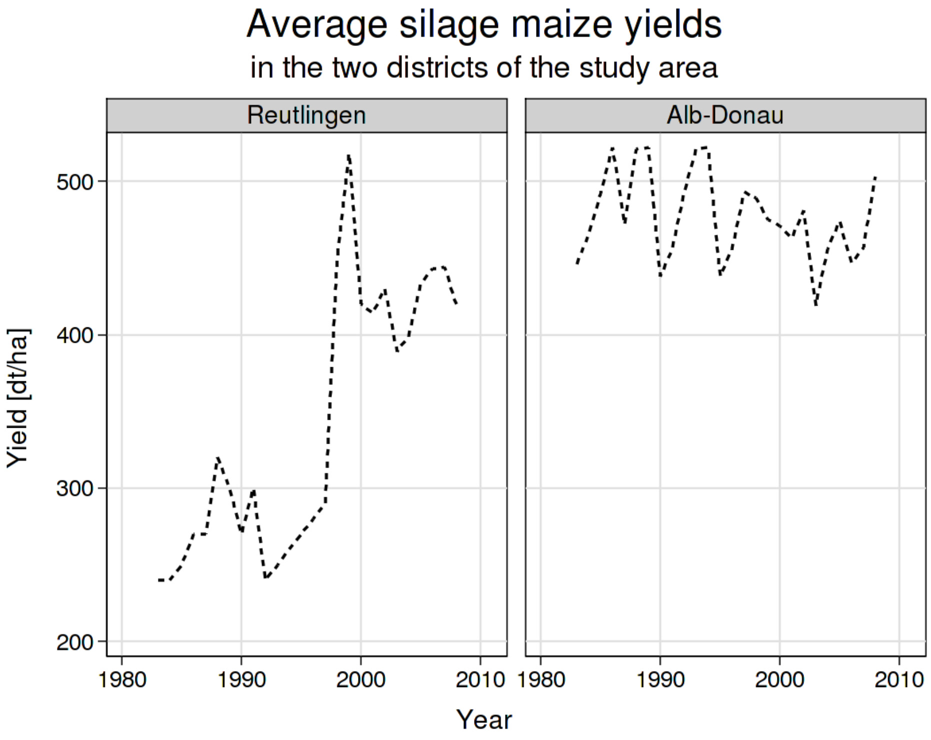


Figure 4.2: *Average silage maize yields in the two study area districts 1983-2011 [Statististisches Landesamt Baden-Württemberg, 2012]*

tions for wheat and rapeseed between the two districts in the survey, when using t-test for mean comparison. Even the lowest expectation for wheat mentioned in the survey was higher than the long and short-term averages recorded in the statistics for Reutlingen. For malting barley, only one respondent was from Reutlingen, but his answer also lay well within the range of answers provided by the four farmers from Alb-Donau.

Based on this information, we created a set of three uniform yield scenarios (table 4.6): one modeled on the Reutlingen yields (urt), one based on the Alb-Donau time series (uad), and one mixing summer crop yields from Reutlingen with winter crop yields from Alb-Donau (umx) always using the 6-year average. They are qualified as uniform, because they assume the same yield irrespective of soil type or type of fertilization (manure, mineral).

Table 4.6: *Alternative yield scenarios*

		Year			
Scenario	Crop	1999	2003	2007	2011
urt	Silage maize	340	420	420	420
	Summer barley	50	45	45	45
	Winter barley	55	55	55	55
	Winter rapeseed	31	34	34	37
	Winter wheat	63	63	63	63
uad	Silage maize	470	470	470	470
	Summer barley	50	51	53	54
	Winter barley	63	63	63	63
	Winter rapeseed	34	34	37	38
	Winter wheat	73	75	77	79
umx	Silage maize	340	420	420	420
	Summer barley	50	45	45	45
	Winter barley	63	63	63	63
	Winter rapeseed	34	34	37	38
	Winter wheat	73	75	77	79

4.2 Prices

The price information required for the model comprises producer prices for crops and animal products, purchase prices for consumable inputs, buying prices and maintenance cost of investment goods and wages for hired labor. Producer prices for major crop and animal products were taken from the regional statistical time series in LEL [2010, 2011a,b]. For other products and inputs, we constructed a time series combining prices reported for the year 2009 from KTBL [2010] and combined it with the corresponding price indices from destatis [2012d]. Figure 4.3 shows the development of the producer prices of the crops most relevant for the study area between 1995 and 2011.

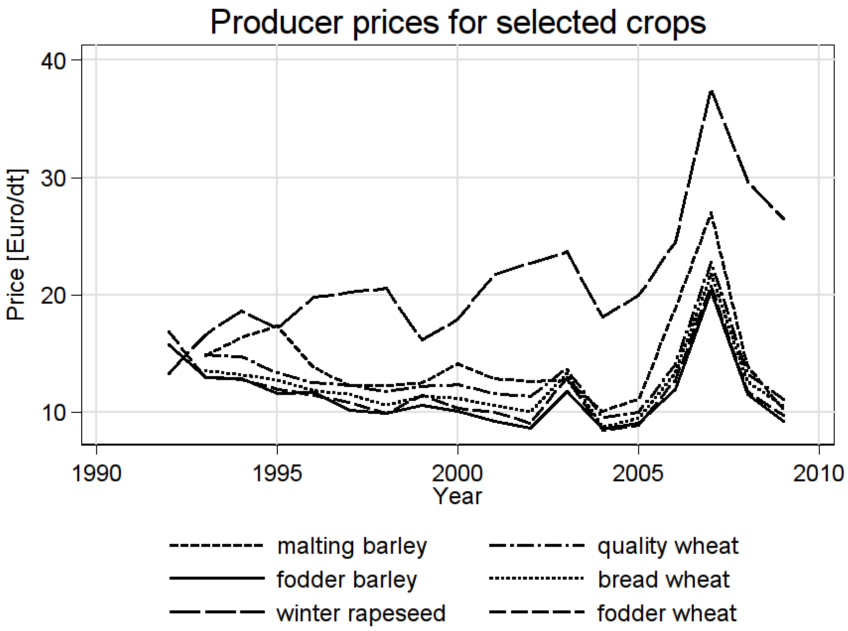


Figure 4.3: *Development of producer prices for the major crops of the area, 1993-2009. Illustration based on data from LEL 2010.*

The seasonal nature of agricultural production requires a distinction between products, whose prices are assumed to be known at the time of decision making (including investment goods, fuel and most other input prices), and those not known at the time of the production decision, including prices for crops and – to a lesser extent – also for animal products.

4.2.1 Prices used in calibration and validation experiments

Similar to crop yield expectations, assumptions on the formation of expectations are required to infer the product prices used during production planning. In the farm survey, farmers were also asked to describe their long-term expectation for producer prices as triangle distributions. Table 4.7 shows for each product the mean and the range of answers over all respondents, which were asked for the price they expected to see most frequently in the following years. We also asked for the lowest and the highest price they would expect to observe in the following years.

In a second question, farmers were also asked for the wheat price they would specifically expect for 2011, with answers shown in table 4.8. Except for bread wheat of quality A, the answers differ very little from the long-term expectations discussed above.

Like for crop yields, we calculated price averages over the three, respectively six years preceding each point of observation as a proxy for price expectations for our sim-

Table 4.7: *Long term price expectations of surveyed farmers (Lowest, most frequent and highest points of triangle distribution)*

			Triangle distribution of expected price [€]					
			Most frequent		Lowest		Highest	
Product	Unit	N	Mean	Range	Mean	Range	Mean	Range
Fodder wheat (C)	dt	9	13.98	[11; 16.5]	10.06	[7; 14]	20.56	[17; 25]
Bread wheat (B)	dt	4	16.75	[15; 20]	12.25	[9; 18]	23.75	[20; 25]
Bread wheat (A)	dt	6	15.67	[13; 18]	10.67	[9; 16]	24.33	[22; 26]
Bread wheat (E)	dt	4	19.50	[17; 25]	11.00	[8; 16]	43.00	[22; 100]
Malting barley	dt	5	18.00	[15; 23]	11.20	[6; 20]	37.20	[24; 80]
Winter rapeseed	dt	13	30.77	[25; 35]	23.58	[19; 29]	47.38	[35; 120]
Milk	100 l	14	31.14	[28; 35]	22.69	[18; 27]	38.77	[35; 45]
Beef	kg	5	3.06	[2.85; 3.3]	2.66	[2.5; 2.8]	4.02	[3.1; 6]
Pork	kg	10	1.38	[1.3; 1.5]	1.16	[1.1; 1.4]	1.71	[1.5; 1.9]
Piglet 25kg		2	52.75	[45.5; 60]	32.00	[30; 34]	70.00	[60; 80]

Table 4.8: *Wheat price expectations for 2011 for survey farmers*

Category	N	Mean	Range
Fodder wheat (C)	11	13.77	[10.5; 17]
Bread wheat (B)	4	16.50	[15; 18]
Bread wheat (A)	6	17.08	[15; 20]
Bread wheat (E)	4	18.13	[15.5; 20]
Organic bread wheat (E)	4	39.50	[36; 44]

ulation experiments (tab. 4.9). If we compare the results for 2010[1] with the expectations recorded in the farm survey, it seems that the 3-year average looks much more consistent with farmers' responses than the 6-year average. Based on this observation, we decided to use the calculated 3-year averages as proxies for expected prices in our calibration and validation simulations.

4.2.2 Price scenarios for climate change simulations

In the simulations used to test the effect of climate change on short-term production decisions, we used the average of the prices observed between 2000 and 2009 converted to 2009 real terms in all three scenarios (B, C1, C2). For specific assessments, individual prices were changed as described in the article.

---

[1] the most recent year we were able to calculate an average for

Table 4.9: *Price average of the x years preceding the years of observation. Source: Own calculations based on data from LEL 2010, 2011a,b*

		Price [€]			
Product	x	1999	2003	2007	2010
Malting barley	3	12.79	13.20	13.34	17.03
	6	14.48	12.78	13.04	15.19
Fodder barley	3	10.58	9.32	9.85	13.69
	6	11.52	9.77	9.87	11.77
Winter rapeseed	3	20.18	20.80	20.85	31.25
	6	18.82	19.88	21.78	26.05
Bread wheat	3	11.32	10.59	10.51	14.97
	6	12.24	10.88	10.91	12.74
Quality wheat	3	12.18	11.76	11.21	15.73
	6	13.26	11.92	11.71	13.47
Fodder wheat	3	10.71	9.80	10.02	14.08
	6	11.64	10.26	10.34	12.05
Piglets	3	45.75	46.97	42.67	42.45
	6	46.49	43.23	42.67	42.56
Pork	3	1.40	1.43	1.43	1.45
	6	1.44	1.35	1.40	1.44
Beef (young bulls)	3	2.70	2.43	2.92	3.12
	6	2.70	2.58	2.69	3.02
Milk	3	0.295	0.311	0.286	0.309
	6	0.294	0.304	0.294	0.297

4.3 Field work days

KTBL [2010] provided estimates of available field working days in each half month of the growing season. Estimates are specific to the weather sensitivity level of field work, probability of occurrence (60%, 70%, 80%, 90%) – respectively grain water content (14%, 16%, 18%) for cereal harvest activities –, and one of 12 agro-climatic subregions. Our study area falls into three of these subregions. Areas above 700 m fall into region 4, areas below 700 m into region 5 (The Hochalb, also above 700 m, falls into region 2). As our current model design does not allow to distinguish different climatic regions in our model area[2], we could only use the values of one of the regions, and we chose to test 4 and 5 in our validation experiments.

We also tested two levels of probability of occurrence, 60% and 80% (respectively 16% and 14% grain water content at 80% probability for cereal harvest). Further, we tested the parameter relating the potential to hire work of a certain type in a work season to the suitable field working days in the corresponding time span (see section 2.5.3).

For climate change scenario C2, we used the values for climate zone 7, which contains the lower, still hilly areas surrounding our study area. The number of suitable days for each field work season in the different climate zones and under the different probability levels is shown in table 4.10.

4.4 Rotation options

The compatibility of crops as direct neighbors in the crop rotation was obtained through expert interviews and recorded in the compatibility matrix shown in table 4.11.

For calibration and the baseline, we assumed the relationships as given in the matrix. For climate change scenario C2, we assumed that growing rapeseed directly after winter wheat becomes possible.

---

[2] This is theoretically possible by distinguishing soils not only by soil type, but also by climatic region. We chose not to do so for simplification

Table 4.10: *Suitable days for field work by work season in different climate zones for medium soil resistance (Source: KTBL [2010], own classification of seasons)*

			Weather sensitivity level[1]				
Field work season	Probability[2]	Zone	1	2	4	5	6
SPR	80%	4		1	3	9	27
Spring		5		0	3	11	31
early March		7		1	4	18	40
– early May	60%	4		3	5	19	40
		5		2	5	21	44
		7		3	6	28	53
ESU	80%	4		14	24	21	37
Early summer		5		12	23	28	43
late May		7		17	32	38	50
– early July	60%	4		21	31	33	46
		5		21	31	41	51
		7		27	37	49	56
HWB	80%	4	3.11	5	7	6	10
Harvest winter barley		5	1.89	4	8	7	11
late July		7	3	6	10	11	13
	60%	4	7.22	8	10	9	12
		5	6.22	7	10	11	14
		7	8.33	9	11	13	15
HWR	80%	4	1.89	5	8	6	9
Harvest winter rapeseed		5	1.89	5	8	8	11
early August		7	3.33	7	10	11	13
	60%	4	6	8	9	9	12
		5	6.33	7	10	11	13
		7	8.56	9	11	13	14
HWW	80%	4	1.33	4	7	6	10
Harvest winter wheat		5	1.67	4	7	7	11
late August		7	2.56	5	9	10	13
	60%	4	4.89	7	9	10	13
		5	5.56	6	9	11	14
		7	7.56	8	11	13	15
SP1	80%	4		3	6	6	11
Early September		5		2	6	7	11
		7		3	7	10	13

---

[1] 1 - cereal harvest; 2- hay harvest (soil dried); 4 - harvest of grass silage; 5 - medium sensitive activities such as harvest of silage maize, mineral fertilization, and sowing; 6 - less sensitive activities such as organic fertilization and incorporation of crop residues into the soil

[2] For sensitivity level 1, KTBL lists only the probability level for 80% for different levels of grain humidity, we use the values for 14% and for 16%, respectively.

Table 4.10: *Suitable days for field work by work season (cont.)*

			Weather sensitivity level				
Season	Probability	Zone	1	2	4	5	6
	60%	4		6	9	9	13
		5		4	8	11	13
		7		6	9	13	14
HSM	80%	4		0	0	7	11
Harvest silage maize		5		0	0	8	12
late September		7		0	0	11	13
	60%	4		0	0	11	13
		5		0	0	11	13
		7		0	0	13	14
AUT	80%	4		0	0	10	27
Autumn		5		0	0	7	26
early October		7		0	0	17	35
– early November	60%	4		0	0	22	37
		5		0	0	19	36
		7		0	0	31	42
WIN	80%	4		0	0	0	2
Winter		5		0	0	0	2
late November		7		0	0	0	7
– late February	60%	4		0	0	0	5
		5		0	0	0	6
		7		0	0	3	9

Table 4.11: *Compatibility of crops in rotation*

	Following crop								
Preceding crop	Fallow	Field grass	Silage Maize (+I)	Silage Maize	Summer barley (+I)	Summer barley	Winter wheat	Winter barley	Winter rape
Fallow	1	1	1	1	1	1	1	1	1
Field grass	1/2 *	2/3*	0	1/2 *	0	1/2 *	1/2 *	1/2*	0
Silage maize	1	1	0	X	1	1	1	0	0
Summer barley	1	1	1	0**	1/2	0**	0	1	1
Winter barley	1	1	1	0**	1	0**	0	0	1
Winter rape	1	1	1	0**	1	0**	1	1	0
Winter wheat	1	1	1	0**	1	0**	1/2*	0	(0)***

+I: with witer cover crop (e.g. field mustard)

0: incompatible

1: compatible, full area can be considered for following crop

1/2: maximum half of the area can be considered, e.g. wheat can directly follow wheat only once, then another crop has to be grown before wheat can be grown again

2/3: Field grass is a semi-permanent culture that is usually kept 2-3 years on the same field. So at maximum half the area can be considered preceding crop for other crops and at maximum 2/3 can be considered preceding crop for next year's fields grass.

X: uncertain, subject to calibration, parameter *maize_on_maize*

** Part of the study area is a water protection area, where a cover crop is required, whenever a summer crop follows a preceding crop harvested before September. We assume this restriction to hold for the whole area.

*** Currently, sowing dates of winter rapeseed and harvest dates of winter wheat often overlap and few farmers have the machinery to harvest, till and sow within 2 days. A change of climate might move wheat harvest to earlier dates, allowing wheat - rapeseed rotations regularly in the future.

Chapter 5

Validation and calibration of the short-term production decision

One of the basic assumptions of our model is that the short-term production decision ($\mathbf{a}$) of a farmer (i) for a given year (t) can be predicted with reasonable accuracy, if one knows his/her asset endowments at the beginning of the season (B_t) and his/her knowledge or expectations of major production parameters. These can be divided into parameters (θ) that are expected to remain constant over time or between scenarios, and those that constitute exogenous variables, which may potentially change over time/with scenarios. In our case, the exogenous variables considered are crop yields ($\mathbf{c}$), prices ($\mathbf{p}$), rotational constraints ($\mathbf{r}$), field working days ($\mathbf{c}$) and policy regulations ($\mathbf{z}$).

$$\hat{\mathbf{a}}_{\mathbf{t}} = f\left(\mathbf{B}_t, \mathbf{p}_t^*, \mathbf{y}_t^*, \mathbf{r}_t^*, \mathbf{c}_t^*, \mathbf{z}_t^*, \theta\right) \tag{5.1}$$

To empirically test our model, we needed simultaneous observations of production decisions, exogenous variables and asset endowments at the beginning of the period. Fortunately, we could construct several such consistent data points allowing some control against overcalibration to one specific situation. FDZ [2010] includes observations of land use and animal stocks at the farm level for the years 1999, 2003 and 2007. More specifically, the statistical surveys were conducted always in May of the respective years and reflect production in the cropping seasons 1998/1999, 2002/2003 and 2006/2007. If we consider total land endowment and its partition into grassland, arable and forest land, as well as livestock stalls induced from animal stocks as asset observation at the beginning of the season, and the particular crop choice and intensity of grassland use as well as actual stocking rates as major outcomes of the production decision, the panel dataset provides a good, but incomplete basis for a calibration and validation dataset.

Figure 5.1 summarizes the process of data gathering to arrive at three consistent datasets that was described in the previous section. In the following sections, we list the parameters reflecting uncertainty in model design, explain our goodness-of-fit criteria and our approach to reducing parameter uncertainty by stepwise calibration. We conclude with a final comparison of observed land use and farm type distributions with those simulated with the model using the reduced parameter set.

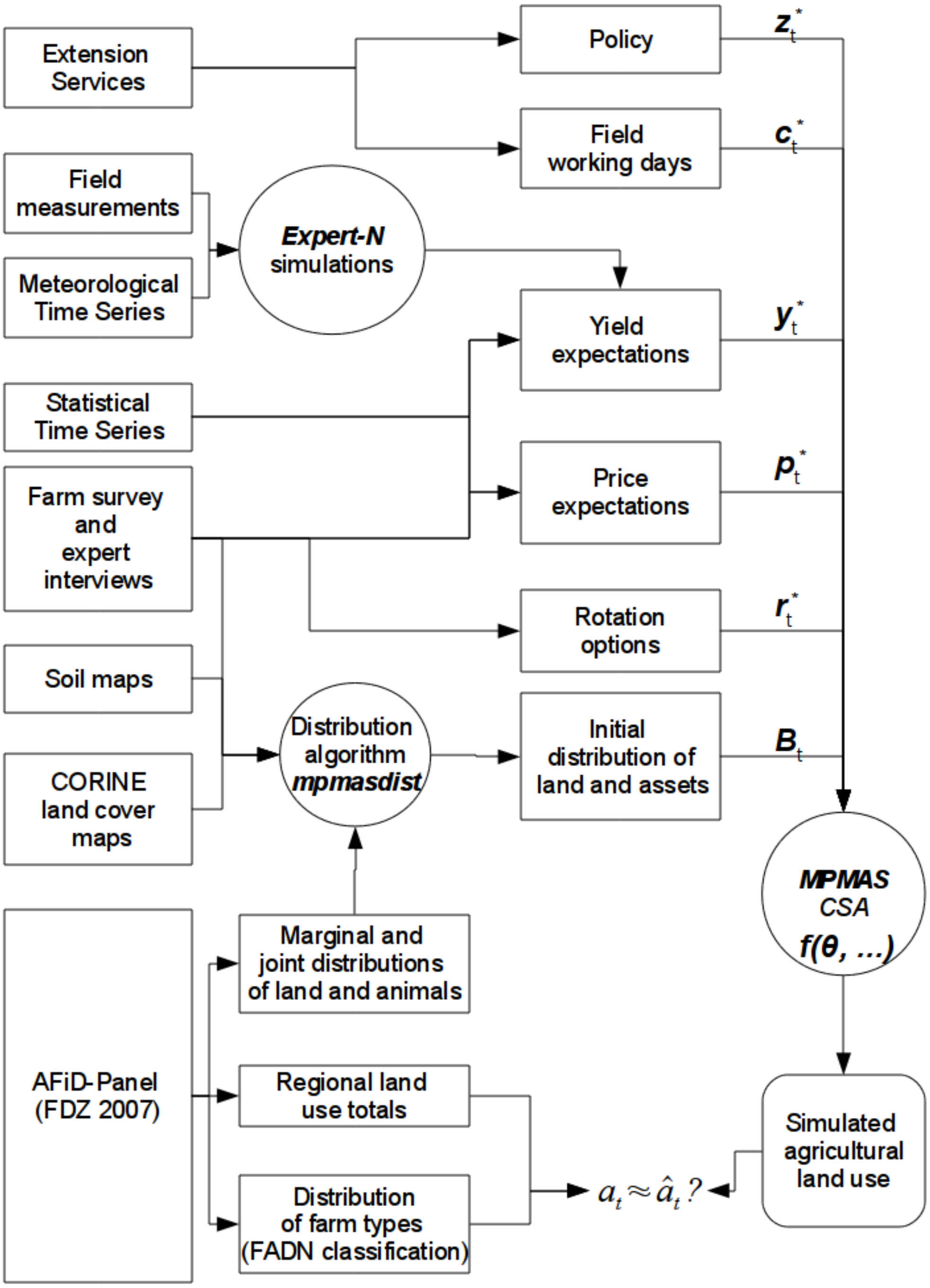


Figure 5.1: *Dataset for the empirical validation of short-term production decisions in the model* ($t \in \{1999, 2003, 2007\}$)

5.1 Calibration experiments

The process of calibration followed a sequential, Bayesian-like approach, in which we did not intend to identify a single, best parameter combination, but only to reduce the model uncertainty as far as considered possible without running the danger of overfitting. The whole process – which was as much about calibration as about technical verification of the model – can be subdivided into two phases:

The first phase can be characterized as an informal search for errors and significant omissions and comprised numerous iterative steps. Each step would usually include an elementary effects screening and then a full factorial including the most relevant parameters (i.e. those with the strongest effects on goodness-of-fit). The distribution of goodness-of-fit over the factorial as well as the predicted land use, farm type patterns and livestock numbers was then examined. Whenever the distribution did not cover the true value, i.e. there was no combination of parameters which was at all able to reproduce the observations, this prompted the identification of an error in the model implementation or a reconsideration of a theoretical or empirical aspect of the model, usually leading to the introduction of another parameter and its inclusion into the testing procedure.

The second and final phase was then the actual calibration, in itself not different from any of the steps of the first phase, but including again all parameters tested or introduced during the previous experiments globally varying them, and formally applying the calibration criteria. Only this final phase is reported here as it implicitly summarizes the previous process.

5.1.1 Parameter variation

Parameter settings tested during the calibration and validation experiments are listed in table 5.1.

A total of 48 different initial agent populations were generated for each year using four different seed values for the random generator, three different settings for the *gvpha* limit used during the random allocation of animals, two different values for *birth_factor_past* and two different values for *potsuc_prob_male*, the probability of a male descendant to pursue a career in farming or not.

Parameters related to crop yields include the four yield sets, the scaling factor for first year wheat (*wheat_normal*), a scaling factor for silage maize yields (*maize_yc*) to reflect the uncertainty of maize production in this boundary region, and to factors which include or exclude the production of whole-plant silage (*wps*) and scale whole-plant silage yields (*wps_coef*), as we consider this an innovation and we have no data on the diffusion of this technology in the observation years and little information on crop yields.

Two parameters affect the potential maize area of an agent: *maize_on_maize* controls the number of years maize can be grown after itself, and *maizerotlimit* constitutes the upper limit for the total share of maize in the crop rotation.

Parameters related to field working days include the KTBL climate region (*clregion*),

Table 5.1: *Parameter settings tested during the calibration and validation experiments*

Initial agent population	*gvpha*	2, 2.5, 3
	popseed	4 different seeds
	birth_factor_past	[1; 1.05]
	potsuc_prob_male	[0.5; 1]
Yields	yield set	urt, uad, umx, xn3
	maize_yc	[0.75; 1]
	wheat_normal	[1; 1.1]
	wps	yes/no
	wps_coef	[1;1.3]
Crop rotation	*maize_on_maize*	0, 1/2, 1
	maizerotlimit	[0.4;0.6]
Field working days	*clregion*	4, 5
	fielddayprob	60%, 80%
Contracted field work	*workforhirecoef*	[0; 1]
	proptohire	[0.5; 2]
Cattle feeding	*pasturelabor*	[1; 3]
	pastureloss	[0.1; 0.4]
	freshgrasslabor	[1; 3]
	freshgrassloss	[0; 0.2]
Markets	*trade_yf_cattle*	no/yes
	trade_smaize	no/yes
	nawarosale	no/yes
	biertreber	yes/no
	kwkyno	no
Manure	*high_manure_maize*	yes/no
	manure	[1; 1.5]
Farm household	*ihorizon_type*	4 different versions

and the certainty level of the assumed field working days (*fielddayprob*) .

The *workforhirecoef* scales the price for contracted field work between the maximum and the minimum of the range given in KTBL [2010], while the *proptohire* coefficient scales the availability of hired field work per hour with suitable weather (see 2.5.3).

The *pasturelabor* and *freshgrasslabor* coefficients scale the amount of labor necessary for pasturing, and the *pastureloss* and *freshgrassloss* coefficients indicates the share of pasturing, respectively harvest losses.

The *trade_yf_cattle* parameter controls whether farmers can buy young female cattle (3-months old calves and heifers) or he has to raise them himself. While in reality, of course, there is a market for female cattle, the model is not able to reflect the quantity effects on the market, and produces too many farms expecting to buy female cattle from the market without anyone producing them. Similarly, the *trade_smaize* parameter controls, whether farmers can sell silage maize on the market or whether it can only be considered for own production. The *nawarosale* controls, whether there is any demand for crops produced on set-aside land under the NaWaRo regulations, or whether these can only be used in own biogas plans. The *biertreber* parameter controls whether brewery by-products are generally available as fodder.

Two parameters are related to the maximum amount of manure that can be applied to a crop. The production activities, which consider manure use assume a standard amount of manure use, which effectively creates an upper limit of manure application to each crop. The ζ_{manure} parameter scales this upper limit on manure use of all production activities, in order to test whether the assumed standard amounts may be too low. The *high_manure_maize* is specific to silage maize production. It controls the inclusion of specific silage maize production activities, which assume a manure amount of 30 m^3 instead of the standard 20 m^3 (with the complementing mineral fertilization reduced).

Last, the *ihorizon_type* represents four different implementations of the influence of farm household composition on the production decisions of the farm: In the simplest version, the investment horizon is independent of the farmers age and the ζ_{H2ut} is equal to zero, i.e. the farm manager derives no utility from employing potential successors. In the second version, the investment horizon remains independent of the farmers age, but the farm manager derives utility from employing potential successors, i.e. ζ_{H2ut} is equal to one. In the third version, the investment horizon depends only on the age of the current household head, while in the fourth version it depends on the age of the successor with highest priority. ζ_{H2ut} is equal to one in both cases.

5.1.2 Observed land use decisions and goodness of fit criteria

To check the predictive accuracy of our model, we compared the simulated land use decisions to the panel observations using three indicators: the total land use in the study area, the total livestock numbers in the study area, and the classification of farms according to principal type of farm (PTOF) of the EU farm typology (Commission Decision 2003/369/EC). While the first two indicators reflect the aggregate response of the agricultural sector, the distribution of farm agents over principal type of farm (PTOF) classes reflected the combination of different production activities within one farm, and indicates whether the heterogeneity of farm setups has been well reproduced.

Total land use of full-time farms in the study area for the three observation years is shown in table 5.2. Due to the privacy constraints several original crop categories had to be grouped to aggregate categories. Not all crop categories have been included into the model. For goodness-of-fit comparison they have been associated with the closest representative in the model in terms of cultivation pattern and use.

The farm classification of the EU typology (Commission Decision 2003/369/EC) is based on the observed crop areas and livestock numbers of farm holdings. It is not used as input into the model, but calculated based on the simulated crop areas and stocking rates of agents, which are then compared to the farm type distribution calculated from the observed panel data. In a first step, the EU typology weighs farm activities by their potential contribution to farm income using standard gross margins (SGM). Farm types are then defined based on the share that different activity categories contribute to the total SGM of a farm. An overview of the farm type distribution obsered in the area is gien in table

Both, the distribution of total area over land uses and the distribution of farms over PTOF classes are restricted by an overall total and thus constitute categorical data: For these, we used the model efficiency based on the standardized absolute error [Voas and Williamson, 2001] as goodness-of-fit measure. For the animal numbers, we used the standard model efficiency.

5.1.3 Reducing the parameter space

The parameters described above (section 5.1.1) all reflect some uncertainty about the best representation of reality in our model. To reduce parameter uncertainty, we used a stepwise procedure to exclude parameter settings, which consistently perform inferior than their alternatives. To guard against overcalibration, we only excluded those settings, whose inferiority was consistently observed in all three observation years. As the observation years differ by several structural breaks (especially with respect to the policy setting, but also with respect to price levels), there is a reasonable chance that parameters fulfilling these conditions present a good choice also for scenario analysis.

There were two exceptions to this rule: First, parameters affecting the initial agent population could differ from year to year. Here, we were not so much interested in stable parameters, but rather looked for an agent population, which best represented the agent population in a given year. Second, in the case of the yield data, it was clear *a priori*, that we would use the ‘xn3’ set of yields during the scenario analysis: We used Expert-N to simulate future yields and therefore we had to use yields calculated by Expert-N also for the baseline. Due to the fact that the simulated yields were subject to considerable uncertainty, we also included the other yield sets in our calibration experiments to guard against overcalibrating the other parameters to this specific set of yields. Analogous to the condition, that a parameter setting was only excluded if it consistently performs inferior in all years, it was also excluded only if it consistently performed inferior for all

Table 5.2: *Overview of aggregate land use used for goodness-of-fit calculation*

		Area		
Statistic from FDZ	Model category	1999	2003	2007
Summer barley		4822	4266	4188
Oats, Mixed summer cereals		1290	1163	866
	Summer barley	6112	5429	5054
Winter barley, Mixed winter cereals	Winter barley	3153	2962	2784
Winter wheat		4965	4450	4997
Triticale		795	908	934
	Winter wheat	5760	5359	5931
Winter rapeseed	Winter rapeseed	2673	2220	2184
Silage maize	Silage maize	1621	1720	3030
Field grass		32	22	183
Clover, Alfalfa		2086	1197	1483
Other field forage		546	271	100
Fodder peas, fodder beans		373	279	149
	Field grass	3038	1770	1915
Fallow	Fallow	1047	1038	586
Total Arable		**23403**	**20497**	**21483**
Pasture	Pasture	507	486	298
Meadow		15906	14138	13868
Mown pasture		1140	1341	1297
	Meadow	17045	15479	15165
Total Grassland		**17552**	**15964**	**15463**
Total		**40956**	**36462**	**36946**

Table 5.3: *Overview of the farm classification according to the EU farm typology valid until 2009 (Commission Decisions 2003/369/EC). The classification is based on the shares of standard gross margin (SGM) of activity categories in the total farm SGM [BMELV, 2008]*

GTOF	PTOF	Code	Classification rule
Specialist field crops			Field crops >2/3
	Specialist cereals/oilseeds	13	Cereals/Oilseeds > 2/3
	General field crops	14	Cereals/Oilseeds $\leq$ 2/3
Specialist horticulture	Specialist horticulture	20	Horticulture > 2/3
Specialist permanent crops	Specialist permanent crops	30	Permanent crops > 2/3
Specialist grazing livestock			Grazing livestock/Grasslands > 2/3
	Specialist dairying	41	Dairy cattle > 2/3 & dairy cows > 2/3 of these
	Specialist cattle fattening/rearing	42	Cattle > 2/3 & dairy cows $\leq$ 1/10
	Cattle fattening/rearing/dairying combined	43	Cattle > 2/3 & dairy cows > 1/10 but not class 41
	Sheep, goats and other grazing livestock	44	Cattle $\leq$ 2/3
Granivore	Specialist granivores	50	Granivores > 2/3
Mixed cropping	Mixed cropping	60	1/3 < Field crops $\leq$ 2/3 or 1/3 < Horticulture $\leq$ 2/3 or 1/3 < Permanent crops $\leq$ 2/3 and Granivores $\leq$ 1/3 and Grazing lv. < 1/3
Mixed livestock			1/3 < Grazing lv. $\leq$ 2/3 or 1/3 < Granivores $\leq$ 2/3 and all others each $\leq$ 1/3
	Mixed livestock, mainly grazing	71	1/3 < Grazing lv. $\leq$ 2/3 and all others each < 1/3
	Mixed livestock, mainly granivore	72	1/3 < Granivores $\leq$ 2/3 and all others each < 1/3
Mixed			Others
	Field crops & grazing livestock combined	81	Grazing Lv > 1/3 and Field crops > 1/3
	Various crops & livestock combined	82	Others

SGM: Standard Gross Margin
GTOF: General Type of Farm
PTOF: Principal Type of Farm

Table 5.4: *Overview of farm type distribution for goodness-of-fit calculation*

GTOF	PTOF	Code	1999	2003	2007
Specialist field crops	Specialist cereals/oilseeds	13	40	25	18
	General field crops	14	0	0	5
Specialist horticulture	Specialist horticulture	20	14	11	9
Specialist permanent crops	Specialist permanent crops	30	7	5	5
Specialist grazing livestock	Specialist dairying	41	359	243	212
	Specialist cattle fattening/rearing	42	18	15	20
	Cattle fattening/rearing/-dairying	43	71	49	38
	Sheep, goats and other livestock	44	28	27	18
Specialist granivores	Specialist granivores	50	72	58	44
Mixed cropping	Mixed cropping	60	17	9	8
Mixed livestock	Mixed livestock, mainly grazing	71	74	40	24
	Mixed livestock, mainly granivore	72	42	22	20
Mixed	Field crops & grazing livestock	81	73	60	65
	Various crops & livestock	82	118	42	47

yield sets (Yield sets were therefore also never fixed in any of the screening steps) and for both the aggregate land use and disaggregate farm type goodness-of-fit measures.

Even if we had wanted to use only two factor levels for each of our parameters, $3 * 2^{26} > 200\mathrm{Mio.}$ runs would have been required to run a full factorial design for our setting. With a model run time of at least 30 minutes for one simulation period, this was infeasible. Rather we used two rounds of elementary effects screening and then ran a full factorial with the remaining parameters.

The first elementary effects screening used ten repetitions for each elementary effects and thus required $10 * (26 + 1) * 3 = 810$ model runs. The design was created in R using the morris() function of the 'sensitivity' package [Pujol, Iooss, and Janon, 2012], which includes the space filling improvements of Campolongo, Cariboni, and Saltelli [2007], and allows for choosing a different number of levels for each parameter. The later came very handy in our case, as the majority of our parameters was discrete, many with only two defined levels. We calculated the Morris sensitivity measures to assess the importance of each parameter in determining the three goodness-of-fit measures (see 5.1.2)

We then grouped the parameters into three groups: Parameters that showed little or no effect (low μ^* and low σ^*) on goodness-of-fit could be fixed at their theoretically most convincing values for the next steps as we could not hope to gain much insight on them in the calibration procedure.

The second group were parameters, for which a clearly superior setting could already be identified in the screening. This was indicated by a very low difference between the absolute value of μ and μ^* in the most simple case. The sign of μ then indicated whether the parameter had to be fixed at the lower or upper end of the range. This applied mostly in the case of binary parameters. In other cases, a closer analysis of the sample points, revealed that moving away from a certain value consistently deteriorated goodness-of-fit. These parameters could then be fixed at the identified value for the subsequent steps of the calibration.

We continued the analysis with the third group of parameters, which showed important, but ambiguous effects on goodness-of-fit, using a second elementary effects screening ($195 * 3 = 585$ runs) in order to reassess their importance after fixing the parameters of the first two groups. We repeated the same procedure as above and then, as a third step, ran a full factorial with the six most important unfixed parameters ($324 * 3 = 972$ runs).

Table 5.5 show the results of the parameter fixing. Parameters that could be fixed based on the empirical results are *pasturelabor*,*pastureloss*,*fielddayprob*, *trade_yf_cattle*, *trade_smaize*, *nawarosale*. Three different starting populations based on combinations of *gvpha* and *popseed* were selected reducing these two parameters to one new parameter *pop*. The potential range for *maize_yc* could at least be reduced. Parameters that were fixed temporarily due to insignificant effects on goodness-of-fit remain part of the reduced parameter space, since insignificant effects on goodness-of-fit do not rule out important influences on the effect of climate change or policy analysis.

Table 5.5: *Parameter fixing during the calibration and validation experiments*

Parameter group	Parameter	EE1	EE2	FF
Initial agent population	*gvpha* , *popseed*	-	-	3 comb. per year ⇒*pop*
	birth_factor_past	(1)	(1)	(1)
	potsuc_prob_male	(1)	(1)	(1)
Yields	yield set	-	-	(xn3)
	maize_yc	-	(0.75)	0.8-0.9
	wheat_normal	-	(1.05)	-
	wps	-	0	0
	wps_coef	(1.15)	(1.15)	(1.15)
Crop rotation	*maize_on_maize*	(3/4)	(3/4)	(3/4)
	maizerotlimit	-	(0.5)	(0.5)
Field working days	*clregion*	(4)	(4)	(4)
	fielddayprob	-	80%	80%
Contracted field work	*workforhirecoef*	-	-	-
	proptohire	(1)	(1)	(1)
Cattle feeding	*pasturelabor*	3	3	3
	pastureloss	0.4	0.4	0.4
	freshgrasslabor	(3)	(3)	(3)
	freshgrassloss	(0.1)	(0.1)	(0.1)
Markets	*trade_yf_cattle*	0	0	0
	trade_smaize	0	0	0
	nawarosale	0	0	0
	biertreber	-	(1)	-
	kwkyno	(0)	(0)	(0)
Manure	*high_manure_maize*	-	(1)	-
	manure	-	-	-
Farm household	*ihorizon_type*	(2)	(2)	(2)

Values show fixings applied based on the results of each experiment.
Values in parenthesis denote temporary fixing at theoretical values due to less significant effects on goodness-of-fit

5.2 Empirical validation

The reduced parameter space constituted the basis for further scenario and policy analysis. Before conducting any specific analysis, however, we compared simulation outcomes to observed land use, farm type and farm accounting data to allow a detailed impression of the empirical performance of the model and the reduced parameter space. We could not use the Morris designs of the calibration as a basis for this comparison: Applying the parameter fixing would have reduced the design to too few replications for our purpose. Also the full factorial design of step three was not suitable as it misses parameters, which might have an important influence on income and land use.

To measure the predictive skill that can be achieved with the calibrated parameter space, we ran an LHS sample of 100 runs over the 16 uncalibrated parameters for the year 2007 (19 -3 parameters: There was no market for surplus heat of biogas plants in the past and wheat silage was not considered a production option (rendering also the yield coefficient for wheat silage meaningless.)

In this sample, the $ESAE$ for total land use lies between 0.73 and 0.84, and the model efficiency for total livestock numbers was between 0.9 and 0.95.If we compare the predicted and observed areas for individual crops in 2007 as shown in tab. 5.6, we see a tendency to consistently overestimate wheat, rapeseed and fallow areas, and an underestimation of silage maize and summer barley areas. The bias in the silage maize area is consistent with the omission of silage maize trade as a result of the calibration.

Table 5.6: *Comparison of predicted and observed land use [ha] in validation runs for 2007. (Prediction shown as average and standard deviation over the 100 runs of a Latin hypercube sample of the unfixed parameters.)*

	Predicted			
	Mean		StDev	Observed
Winter wheat	6,982	±	736	5,931
Winter rapeseed	3,244	±	143	2,184
Summer barley	4,150	±	250	5,054
Winter barley	2,748	±	518	2,784
Fallow	1,766	±	140	586
Silage maize	1,938	±	369	3,030
Field grass	1,111	±	227	1,915
Meadow/Mown pasture	14,464	±	161	15,165
Pasture	692	±	115	298

The $ESAE$ for the farm type distribution lies between 0.62 and 0.71. The discrepancy in farm type distribution can be explained partly as a consequence of the omission of horticulture, fruticulture, and sheep rearing from the model (tab. 5.7). These agents falling into PTOF categories 20, 30, 44 are predominantly absorbed by the class of field crop specialists (13). We also notice a slight underestimation of ruminant specialists.

Table 5.7: *Comparison of predicted and observed farm classification in 2007. (Prediction shown as average and standard deviation of the number of agents in each category over the 100 runs of a Latin hypercube sample of the unfixed parameters.)*

	Predicted		
EU typology (PTOF)	Mean	StDev	Observed
Specialist cereals/oilseeds	81	± 6	18
General field crops	3	± 2	5
Specialist horticulture	0	± 0	9
Specialist permanent crops	0	± 0	5
Specialist dairying	188	± 5	212
Specialist cattle fattening/rearing	17	± 3	20
Cattle fattening/rearing/dairying	23	± 3	38
Sheep, goats & other grazing livestock	3	± 1	18
Specialist granivore	47	± 6	44
Mixed cropping	6	± 2	8
Mixed livestock, mainly grazing	30	± 2	24
Mixed liestock, mainly granivore	36	± 2	20
Field crops & grazing livestock	57	± 8	65
Various crops & livestock	42	± 4	47

Bibliography

Aurbacher, J., P.S. Parker, G.A. Calberto Sánchez, J. Steinbach, E. Reinmuth, J. Ingwersen, and S. Dabbert. 2013. “Influence of climate change on short term management of field crops – A modelling approach.” *Agricultural Systems* 119:44 – 57.

Berendonk, C. 2011. *Standortgerechte Weide- und Mähweidenutzung des Dauergrünlands - Grünlandanlage, Pflege, Düngung, Nutzung, Weidemanagement*. Kleve: Landwirtschaftskammer Nordrhein-Westfalen.

Berger, T., and P. Schreinemachers. 2006. “Creating agents and landscapes for multiagent systems from random samples.” *Ecology & Society* 11:19.

BMELV, ed. 2008. *Buchführung der Testbetriebe - Grundlagen zur BMELV - Testbetriebsbuchführung - Stand: Januar 2008*. Bundesministerium für Ernährung, Landwirtschaft und Verbraucherschutz -BMELV.

BMELV. 2006. *Die EU-Agrarreform – Umsetzung in Deutschland. Ausgabe 2006.*. Bundesministerium für Ernährung, Landwirtschaft und Verbraucherschutz.

Campolongo, F., J. Cariboni, and A. Saltelli. 2007. “An effective screening design for sensitivity analysis of large models.” *Environmental Modelling & Software* 22:1509–1518.

CLC2000. 2004. “CORINE Land Cover (CLC2000); Federal Environment Agency, DLR-DFD.”

CLC2006. 2009. “CORINE Land Cover (CLC2006); Federal Environment Agency, DLR-DFD.”

Day, R. 2008. “Micro foundations for meso and macro economic theory.” *Journal of Evolutionary Economics* 18:261–273, 10.1007/s00191-007-0084-2.

de Brabander, D., J. de Boever, J. Vanacker, C. Boucque, and S. Bottermann. 1999. “Evaluation of physical structure in dairy cattle nutrition.” In P. Garnsworthy and J. Wiseman, eds. *Recent Advances in Animal Nutrition*. Nottingham Univ. Press, pp. 111–145.

Dederer, M., and J. Messner. 2011. *Biogaserzeugung in Baden-Württemberg*. Boxberg: Bildungs- und Wissenszentrum Boxberg (Landesanstalt für Schweinezucht - LSZ).

destatis. 2011. "Arbeitskräfte - Landwirtschaftszählung 2010." Land- und Forstwirtschaft, Fischerei - Fachserie 3 No. Heft 2, Statistisches Bundesamt, Wiesbaden.

—. 2012a. "GENESIS online - Tabelle 12411-0008: Bevölkerung Deutschland."

—. 2012b. "GENESIS online - Tabelle 12612-0008: Geburtenziffern Deutschland."

—. 2012c. "GENESIS online - Tabelle 12613-0003: Gestorbene Deutschland."

—. 2012d. "GENESIS online - Tabelle 61221-0002: Index der Einkaufspreise landwirtschaftl. Betriebsmittel: Deutschland, Wirtschaftsjahr, Messzahlen mit/ohne Umsatzsteuer, Landwirtschaftliche Betriebsmittel."

Fachagentur nachwachsende Rohstoffe e.V. ???? "Abb. 27 Entwicklung Biogasanlagen."

FDZ. 2010. *AFiD-Panel Agrarstruktur 1999, 2003, 2007*. Kiel: Forschungsdatenzentrum der Statistischen Ämter des Bundes und der Länder.

Grimm, V., U. Berger, D.L. DeAngelis, J.G. Polhill, J. Giske, and S.F. Railsback. 2010. "The ODD protocol: A review and first update." *Ecological Modelling* 221:2760 – 2768.

Hartmann, A. 2008. "Wie viel Fläche wird für Biogas benötigt?" Statistisches Monatsheft Baden-Württemberg No. 7/2008, Statistisches Landesamt Baden-Württemberg.

Hazell, P., and R. Norton. 1986. *Mathematical programming for economic analysis in agriculture*. New York: Macmillan.

KTBL. 2008. *KTBL-Datensammlung. - Betriebsplanung Landwirtschaft 2008/09*. Darmstadt: Kuratorium für Technik und Bauwesen in der Landwirtschaft e.V.

—. 2010. *KTBL-Datensammlung. - Betriebsplanung Landwirtschaft 2010/11*. Darmstadt: Kuratorium für Technik und Bauwesen in der Landwirtschaft e.V.

Law, A.M. 2007. *Simulation Modeling & Analysis - International Edition*, 4th ed. Boston: McGraw-Hill.

LEL. 2012. "Datensammlung Kalkulationsdaten Marktfruchtbau Ernte 2012 in Baden-Württemberg; Version 01/2012."

—. 2010. *Material aus der Ernährungswirtschaft des Landes Baden-Württemberg: Getreide und Futtermittel 2008/2009*. Schwäbisch Gmünd: Landesanstalt für Entwicklung der Landwirtschaft und der ländlichen Räume (LEL).

—. 2011a. *Material aus der Ernährungswirtschaft des Landes Baden-Württemberg: Milch 2010*. Schwäbisch Gmünd: Landesanstalt für Entwicklung der Landwirtschaft und der ländlichen Räume (LEL).

—. 2011b. *Material aus der Ernährungswirtschaft des Landes Baden-Württemberg: Vieh und Fleisch 2010*. Schwäbisch Gmünd: Landesanstalt für Entwicklung der Landwirtschaft und der ländlichen Räume (LEL).

LfL. 2012. "Fachbeiträge aus dem Institut für Pflanzenbau und Pflanzenzüchtung."

LfL, ed. 2010. *Gruber Tabelle zur Fütterung der Milchkühe Zuchtrinder Schafe Ziegen*, 32nd ed. Freising-Weihenstephan: Bayerische Landesanstalt für Landwirtschaft.

—. 2011. *Gruber Tabelle zur Fütterung in der Rindermast: Fresser, Bullen, Ochsen, Mastfärsen, Mastkühe*, 16th ed. Freising-Weihenstephan: Bayerische Landesanstalt für Landwirtschaft.

LUBW. 2007. *Wasser- und Bodenatlas Baden-Württemberg*. Landesanstalt für Umwelt, Messungen und Naturschutz Baden-Württemberg.

Mann, S. 2007. "Wie entstehen HofnachfolgerInnen?" *Agrarwirtschaft/German Journal for Agricultural Economics* 56:161–165.

MLREV. 2011. "Verwaltungsvorschrift des Ministeriums für Ländlichen Raum, Ernährung und Verbraucherschutz zur Förderung der Erhaltung und Pflege der Kulturlandschaft und von Erzeugungspraktiken, die der Marktentlastung dienen (Marktenlastungs- und Kulturlandschaftsausgleich - MEKA III)." `https://www.landwirtschaft-bw.info/servlet/PB/show/1318342/28.02.2011%20-%20Verwaltungsvorschrift%20MEKA%20III%20-%20MLR%2025-8872.53.pdf`.

Müller, B., A. F., R. Drees, G. Dressler, G. J., C. Klassert, S. M., J. Schulze, H. Weise, and N. Schwarz. 2013. "Describing human decisions in agent-based models - ODD+D, an extension of the ODD protocol." *Environmental Modeling and Software* 48:37–48.

Pujol, G., B. Iooss, and A. Janon. 2012. *R Package 'sensitivity'*.

Saltelli, A., S. Tarantola, F. Campolongo, and M. Ratto. 2004. *Sensitivity analysis in practice. - A guide to assessing scientific models.*. Chichester:Wiley.

Schreinemachers, P., and T. Berger. 2011. "MP-MAS: An agent-based simulation model of human-environment interaction in agricultural systems." *Environmental Modelling & Software* 26:845–859.

Statististisches Landesamt Baden-Württemberg. 2012. "Struktur- und Regionaldatenbank." http://www.statistik.baden-wuerttemberg.de/SRDB/home.asp.

Voas, D., and P. Williamson. 2001. "Evaluating Goodness-of-Fit Measures for Synthetic Microdata." *Geographical & Environmental Modelling* 5:177–200.

WETTREG. 2010. "WR2010_EH5_1_A1B: UBA-WETTREG ECHAM5/OM 20C + A1B Lauf 1: 1961-2100." `http://cera-www.dkrz.de/WDCC/ui/Project.jsp?proj=WETTREG2010`.

Appendix A

Appendix

A.1 Symbology used in mathematical programming equations

A.1.1 General remarks

For the description of the equations of the mathematical programming problems, the following symbols are used in this model documentation: The letter x is used for decision variables, the letter b for capacities, the letter a for the coefficients of decision variables in a constraint equation and the letter c for the coefficients of the decision variables in the objective function. M stands for a high value representing infinity, which is chosen high enough to completely relax the associated restriction, but low enough not to cause any problems for the branch-and-bound algorithm.

The type of decision variable is indicated by the superscript, which usually refers to a set, denoted by a capital letter. The specific member of the set to which a decision variable is associated is denoted by a lowercase letter in the subscript. E.g. L denotes the set of land use activities. The $\mathbf{x}^L$ is the vector of areas of all land use activities and x_l^L is the area of land use activity $l \in L$.

Subsets of sets are denoted by adding lowercase letters to the capital letter denoting the superset, e.g. Lg denotes the set of all grassland and field grass production activities, which is a subset of L.

In some cases, a decision variable is associated with several sets, e.g. $x_{e,t,k}^{E,T,K}$ is the number of hours equipment e is used in work seasons t for work with weather sensitivity k . In other cases, several decision variables are associated to the same set and a small letter is added to the superscript in order to allow a distinction, e.g. x_g^{sG} and x_g^{bG} denote sales and purchase of good g, respectively.

A similar convention is used for coefficients and capacities. E.g. b_e^E denotes the number of equipments of type e owned, $a_{l,g}^{LG}$ the amount of good g yielded by land use activity l, and c_g^{sG} the sales price of good g.

The symbol $\forall$ to the right of a displayed constraint equation is used to indicate that this type of relation is repeated for each member of the indicated set(s).

Symbols marked with a tilde˜denote expected, rather than actual values.

A.1.2 List of symbols

Table A.1: *Overview of sets in the CSA decision model*

Symbol		Description
A		types of animal production
B		investment
C		cash
	Cf	fix costs, debt payments, depreciation,rental payments
	Cmc	minimum household consumption
D		feeding season
E		equipment
G		consumable goods
	Gb	goods that can be used for biogas production
	Ge	biogas electricity
	Gc	crops
	Gg	fresh grass (cut or pastured)
	Ggp	fresh grass (pastured)
	Gi	pure inputs
	Gm	maize
	Gn	non-traded intermediates
	Go	manure
	Gs	pure products
	Gt	traded intermediates
H		labor
J		
	Ja	animal groups for feeding
	Jf	crop groups for following position in rotation
	Jp	crop groups for preceding position in rotation
	Jr	crop groups for overall rotational limit
	Js	crop groups for rotational limit on self-following
	Jym	crop groups for MEKA diversification measures
K		weather sensitivity level
L		land use activity
	$Lb4$	grassland, with one conservation cut per season
	Lg	grassland and field grass production
	Lgp	grassland and field grass production used for pasture only
	Lgg	grassland, cross-compliance conformant
	$Lgg1$	grassland, one use per season
	$Lgg2$	grassland, two uses per season
	$Lmai$	maize

Table A.1: *Overview of sets in the CSA decision model [cont.]*

Symbol		Description
	Lmf	land use counted as main forage area
	$Loil$	oilseeds
M		infrastructure and machinery for animals
N		nutrients
	N_b	basic nutrients
	N_l	nutrients with a lower limit on dry matter share
	N_r	nutrients with two sided limit
	N_u	nutrients with an upper limit on dry matter share
O		services for animals
P		Tractive power class
S		soil types
T		work season
U		biogas production
V		types of services for animal production
W		types of field work
	Wh	types of work that can be contracted
Y		subsidy & policy related
	Ya	single farm payment entitlements, arable, EU MTR
	Yb	special male cattle premium, EU Agenda 2000
	Yc	crop premium, EU Agenda 2000
	Yd	milk quota
	Ye	extensification premium, EU Agenda 2000
	Yg	single farm payment entitlements, set-aside, EU MTR
	Yg	single farm payment entitlements, grassland, EU MTR
	Yk	slaughter premium cattle, EU Agenda 2000
	Ym	commitments eligible under MEKA
	Yo	mother cow premium, EU Agenda 2000
	Ys	set aside premium, EU Agenda 2000
	Yu	size limits for rewarding biogas electricity through EEG
	Yx	small manure biogas plant (EEG2012)
	Yy	year of establishment of biogas plant
	Yz	relationship of remuneration classes for biogas electricity (EEG 2012)
Z		Tractor class

A.2 Initialization of the agent population

The traditional approach to create agent populations for MPMAS as described in Berger and Schreinemachers [2006] is to estimate empirical cumulative distribution functions and use their inverse (the quantile function) to distribute values according to random draws from a uniform distribution, a standard procedure in Monte-Carlo simulations called *inverse transform method* [e.g. Law, 2007, p. 424]. To account for covariance among different variables, Berger and Schreinemachers [2006] suggest to classify agents according to the variable that shows highest correlation with all other resources (in agricultural applications usually farm size) and then estimate separate distribution functions for each class. This sampling procedure is implemented as a part of the MP-MAS executable and will henceforth be called the traditional MPMAS lottery algorithm. It loops over each agent and each variable, and independently draws a value from the distribution function of a variable in the agent's class. As the information used in the algorithm only incorporates the correlation of a single variable to all other variables, Berger and Schreinemachers [2006] suggest to generate a large number of samples and filter them using statistical comparison tests on population means and the correlation matrix.

While this works in populations with a relatively simple covariance matrix, it will become rather inefficient in more complex situations, especially if additionally a number of theoretical constraints has to be respected (see below). Further, correlation measures like Pearson's or Spearman's correlation coefficients capture only the linear, respectively monotonic part of the relationship between two variables.

More importantly, a neglect of covariance between variables will bias the resulting distribution [Saltelli et al., 2004] and if also the posterior distribution of outcome indicators. Filtering of populations using statistical tests does not overcome this problem: It will only exclude those samples, whose likelihood of being a good representation of the true population falls below a certain threshold, but not correct for the bias in the accepted populations, which will still be considered equally likely.

Ensuring the incorporation of as much available information about the covariance structure as possible into the sampling distributions and ensuring the unbiasedness of the sampling algorithm are thus a major requirement for using MPMAS simulations.

Apart from stochastic covariance between variables, there may also be deterministic theoretical constraints that have to be respected during the assignment of agent characteristics in order to avoid inconsistent agent populations. These theoretical constraints can be used in cases where joint distributions of variables cannot be observed (e.g. if distribution functions for water availability and apple plantations come from different sources, we can constrain apple plantations to be allocated only to agents, which have enough water to sustain them), but they may also be important if a joint distribution function is known: For example, shares of land classes need to add up to one or legal restrictions constrain the number of animals based on the size of the land holding. The restrictions provided by the estimated joint distribution may not be strict enough to ensure these relations. For example, it may assign a probability of 80% for the share of arable land to lie between 40-60%, and of 20% for it to lie between 0 and 40% if the share of grassland lies between 30 and 50% of the total agricultural area. If the grassland share of an agent was randomly determined to be 45%, arable shares of more than 55% could occur according to the JDF estimation, although they are clearly impossible from a theoretical point of view.

In the traditional MPMAS lottery algorithm, theoretical constraints are implemented using a kind of rejection sampling. The value drawn from the distribution function is tested for compliance with the theoretical constraints. If it complies, it is assigned to the agent, if not, a new value is drawn. This procedure is repeated until a suitable value has been found (or a pre-specified maximum number of iterations has been surpassed). As a consequence of this algorithm, the distribution in the agent population is biased towards 'less demanding' characteristics, because in effect the values are drawn from truncated distributions. A mixture of rescaling of input distributions and rejection of too inconsistent samples largely based on trial and error was used in the past in order to ensure consistency of populations [Berger and Schreinemachers, 2006].

One way to overcome this problem is to first draw N values from the estimated distribution ensuring that the whole of the distribution is properly represented, and then randomly distribute these values among the agents respecting the specified theoretical constraints.

In case of a single, one-sided constraint (e.g. the value assigned to the agent has to be smaller than a certain characteristic of the agent) has to be respected, the random allocation can follow a simple algorithm:

1. Order the drawn values from 'most demanding' to 'least demanding', i.e. in ascending order in case of a greater-than and in descending order in case of less-than restriction.
2. Starting from the 'most demanding', one can then randomly assign each value to one of the agents, for whom the constraints are fulfilled and who have not been assigned a value yet.

In case of a complex set of constraints, it will be necessary to use a matching algorithm. We suggest to use the Hungarian Method (or Kuhn-Munkres algorithm) in combination with a random cost component: Each potential combination of an agent with a value is associated with a cost. This cost consist of two parts: a deterministic and a random component. The deterministic component should be zero, when the constraint is fulfilled and positive if the constraint is not fulfilled. The cost can also be used to reflect the severeness of a bad match, in order to prefer slight violations of constraints over more severe ones in case a complete fulfillment of constraints is infeasible. The second cost component should be a random value, which is comparatively small compared to the deterministic part, such that it does not overrule constraint penalties. It is added to the deterministic component and ensures a unique and random solution of the matching problem, which usually contains many feasible value matches for many agents.