ODD+D protocol ALABAMA-ABM

		Guiding questions	Our model
I) Overview	I.i Purpose	I.i.a What is the purpose of the study?	Method comparison (ABM vs. multi-objective landscape optimization), analysis of policy instruments and their influence on landscape configurations
		I.i.b For whom is the model designed?	Scientists
	I.ii Entities, state variables and scales	I.ii.a What kinds of entities are in the model?	Plots, farmers
		I.ii.b By what attributes (i.e. state variables and parameters) are these entities characterised?	**Farmers**: plots owned, total yield from owned plots, income, income threshold **Plots**: location, owner (farmer), soil fertility, proximity to river, land cover (river, intensive grassland or extensive grassland), number of neighbouring extensive plots, profit potential given management options, yield, realized profit/contribution margin (yield + agri-environmental payments)
		I.ii.c What are the exogenous factors / drivers of the model?	Payment levels (payment for extensive grassland, agglomeration bonus, bonus for extensive grassland along river), design of water quality bonus, relative productivity intensive vs extensive grassland, landscape persistence
		I.ii.d If applicable, how is space included in the model?	GIS (virtual landscape) via raster files
		I.ii.e What are the temporal and spatial resolutions and extents of the model?	Yearly time steps, 100 years, grassland allocation decisions are made once a year; one grid cell represents one plot, model landscape comprises 15x15 cells, up to 10 farms (with randomly assigned plots)
	I.iii Process overview and scheduling	I.iii.a What entity does what, and in what order?	1. Initialization: import raster files and translate them into patch attributes; allocate patches (=plots) to farms (randomly selected, same number of plots per farm); calculate income of farms from initial landscape configuration; (optional) set income threshold for each farmer (random from range between average income from initial landscape minus 1 standard deviation to average income plus 3 standard deviations) 2. Check income threshold reached: if farmer's income (from last year) is above threshold, no further changes in management of her plots are made 3. Potential profit calculation: calculate potential profit for each plot (intensive & extensive) given current land allocation (i.e. other farm's plots as managed in

			previous year) and including base payment and boni; includes a correction for increasing agglomeration bonus by switching neighbouring own plots to extensive 4. Allocation: allocate extensive/intensive management to a limited number of plots (given specification of landscape persistence: either a predefined number of randomly selected plots or a predefined number of plots with highest potential for income increase) 5. Yield calculation: calculate each plot's yield given allocation 6. Agglomeration: check how many neighbouring plots are managed extensively 7. Reception of payments: calculate payments received by each plot 8. Calculation of income: calculate total yield and income for each farm 9. Calculation of agri-environmental payment budget 10. Evaluate ecosystem services (ES): translate landscape configuration into ES realizations (R models) [after 100 ticks]
II) Design Concepts	II.i Theoretical and Empirical Background	II.i.a Which general concepts, theories or hypotheses are underlying the model's design at the system level or at the level(s) of the submodel(s) (apart from the decision model)? What is the link to complexity and the purpose of the model?	It's a relatively simple model trying to show that for heterogeneous landscapes, you need spatially differentiated incentives. The farmers' behaviour is boundedly rational in a very simple sense (income threshold).
		II.i.b On what assumptions is/are the agents' decision model(s) based?	Simple microeconomic model with minimal bounded rationality (satisficing): below threshold income maximizing, myopic farmers; above threshold continuation of last chosen strategy (i.e. management allocation pattern).
		II.i.c Why is a/are certain decision model(s) chosen?	Simplicity.
		II.i.d If the model / a submodel (e.g. the decision model) is based on empirical data, where does the data come from?	NA
		II.i.e At which level of aggregation were the data available?	NA

	II.ii Individual Decision Making	II.ii.a What are the subjects and objects of decision-making? On which level of aggregation is decision-making modeled? Are multiple levels of decision making included?	Subjects: farmers / Objects: management (extensive or intensive grassland) on plot level Farmers decide on plot-level
		II.ii.b What is the basic rationality behind agents' decision-making in the model? Do agents pursue an explicit objective or have other success criteria?	Income maximization up to threshold
		II.ii.c How do agents make their decisions?	Income function
		II.ii.d Do the agents adapt their behavior to changing endogenous and exogenous state variables? And if yes, how?	NA
		II.ii.e Do social norms or cultural values play a role in the decision-making process?	NA
		II.ii.f Do spatial aspects play a role in the decision process?	Agglomeration bonus and bonus for extensive grassland in proximity to river depend on spatial patterns (and play a role in farmers' decisions)
		II.ii.g Do temporal aspects play a role in the decision process?	NA
		II.ii.h To which extent and how is uncertainty included in the agents' decision rules?	Agents do not know how other agents will decide in the current period, they only know the allocation in the last period
	II.iii Learning	II.iii.a Is individual learning included in the decision process? How do individuals change their decision rules over time as consequence of their experience?	NA
		II.iii.b Is collective learning implemented in the model?	NA
	II.iv Individual Sensing	II.iv.a What endogenous and exogenous state variables are individuals assumed to sense and consider in their decisions? Is the sensing process erroneous?	Payment rates, soil fertility, land-use allocation in last period; no errors
		II.iv.b What state variables of which other individuals can an individual perceive? Is the sensing process erroneous?	Land-use allocation in last period; no errors

		II.iv.c What is the spatial scale of sensing?	Local (neighbouring plots)
		II.iv.d Are the mechanisms by which agents obtain information modeled explicitly, or are individuals simply assumed to know these variables?	Not modelled.
		II.iv.e Are costs for cognition and costs for gathering information included in the model?	Not explicitly; implicitly, cognitive burden is the reason for income threshold beyond which farmers cease to make new decisions
	II.v Individual Prediction	II.v.a Which data uses the agent to predict future conditions?	Extrapolation from last period
		II.v.b What internal models are agents assumed to use to estimate future conditions or consequences of their decisions?	NA
		II.v.c Might agents be erroneous in the prediction process, and how is it implemented?	Since they only consider neighbouring plots, they cannot take into account reactions of other farmers to changes in land allocation farther away
	II.vi Interaction	II.vi.a Are interactions among agents and entities assumed as direct or indirect?	Indirect
		II.vi.b On what do the interactions depend?	Spatial distances (neighborhood)
		II.vi.c If the interactions involve communication, how are such communications represented?	NA
		II.vi.d If a coordination network exists, how does it affect the agent behaviour? Is the structure of the network imposed or emergent?	NA
	II.vii Collectives	II.vii.a Do the individuals form or belong to aggregations that affect, and are affected by, the individuals? Are these aggregations imposed by the modeller or do they emerge during the simulation?	NA
		II.vii.b How are collectives represented?	NA

	II.viii Heterogeneity	II.viii.a Are the agents heterogeneous? If yes, which state variables and/or processes differ between the agents?	In one variant of the model (where BOUNDED-RATIONALITY = "heterogeneity"), they have different income thresholds.
		II.viii.b Are the agents heterogeneous in their decision-making? If yes, which decision models or decision objects differ between the agents?	See above.
	II.ix Stochasticity	II.ix.a What processes (including initialization) are modeled by assuming they are random or partly random?	Farmers' income thresholds are generated randomly. Also, the assignment of plots to farmers is random. In one model variant (where PERSISTENCE = "random"), the plots on which farmers are allowed to change management in each period are chosen randomly.
	II.x Observation	II.x.a What data are collected from the ABM for testing, understanding, and analyzing it, and how and when are they collected?	The land-use allocation is translated in a measure of biodiversity (based on configuration of extensive grassland plots) and water quality (based on proximity of extensive/intensive grassland plots from river); grass production is calculated by summing the production of each plot; also, the budget needed to finance the agri-environmental payments is calculated.
		II.x.b What key results, outputs or characteristics of the model are emerging from the individuals? (Emergence)	Landscape pattern
III) Details	II.i Implementation Details	III.i.a How has the model been implemented?	Windows 10, Netlogo 6.0.4, R 3.5.1
		III.i.b Is the model accessible and if so where?	https://github.com/BartoszBartk/magenta
	III.ii Initialization	III.ii.a What is the initial state of the model world, i.e. at time t=0 of a simulation run?	Landscape imported from raster files (allocation pattern of management + soil fertility gradient), 10 farmers with randomly distributed plots and (variant) randomly assigned income thresholds.
		III.ii.b Is initialization always the same, or is it allowed to vary among simulations?	Distribution of plots among farmers is random, and has limited influence on results. Income thresholds are always dependent on mean income from initialized landscape, and as such vary among simulations.
		III.ii.c Are the initial values chosen arbitrarily or based on data?	Arbitrarily.
	III.iii Input Data	III.iii.a Does the model use input from external sources such as data files or other	Landscape (raster files): soil fertility distribution (Gaussian), sinusoidal river along east–west axis

		models to represent processes that change over time?	
	III.iv Submodels	III.iv.a What, in detail, are the submodels that represent the processes listed in ‘Process overview and scheduling’?	R models: yield model, habitat index model, water quality model NetLogo submodels: bonus calculation, budget calculation.
		III.iv.b What are the model parameters, their dimensions and reference values?	See table below.
		III.iv.c How were submodels designed or chosen, and how were they parameterized and then tested?	Based on literature, highly stylized (see below).

Submodels:

- Agricultural yield (AY) modelled as a function of production intensity level *P* (with the value of 1.5 for extensive grassland and 2 for intensive grassland) and soil fertility *F*, summarized over all 200 grassland grid cells *i*:
  $$AY = \frac{\sum_{i=1}^{200} \sqrt{P_i(1+F_i)} - 296.8974}{45.9032}$$
  AY is normalized to range between 0 (all extensive) and 1 (all intensive). Within NetLogo, an analogous yield model is used for each grassland plot, without normalization.
- Habitat index (HI) was estimated as total area of the two largest patches of extensive grassland (A_{2X}) divided by 200 (i.e. the number of grassland cells):
  $$HI = \frac{A_{2X}}{200},$$
  assuming that both increasing extent and connectivity of extensive grassland is beneficial for biodiversity. Patches were defined as contiguous extensive grassland cells using the 4-neighbor rule (King’s case). HI can range between 0 (all intensive) and 1 (all extensive).

- Water quality (WQ) was a function of Euclidean distance (D) of intensive grassland cells i to their respective closest river cells:
  $$WQ = 1 - \frac{\sum_{i=1}^{I} \frac{1}{D_i}}{0.8635082} \text{ if } I > 0 \text{ or } WQ = 1 \quad \text{if } I = 0$$
  where I is the total number of intensive grassland cells. Decreasing the number of intensive grassland cells and/or increasing their distances to the river would thus increase WQ, which is normalized to range between 0 (all intensive) and 1 (all extensive).
- Agglomeration bonus is calculated by multiplying the bonus level with the share of extensive neighbouring plots.
- Water quality bonus is normally extended if the plot is within a predefined proximity to river (see parameter DIST) or (optionally, mainly for testing purposes) by a function following the WQ function.

Parameters

Parameter	Name in NetLogo model	Values/range
Persistence	persistence	“random”, “profit”
Limit of changeable plots per period per farmer	change-lim	1–20
Rationality type	bounded-rationality?	TRUE, FALSE
Income threshold type (only if bounded-rationality = TRUE)	bounded-threshold	“heterogeneity”, “uniform”
Type of water bonus	water-bonus	“simple”, “as ES model”
Number of agents	no-agents	1–10
Base payment level	base-p	0–0.25
Agglomeration bonus level	bonus-agg	0–0.25
Water quality bonus	bonus-wat	0–0.25

Distance from river of plots rewarded with water quality bonus	dist	0, 1, 2
Income threshold of each agent	income-thresh	random