CrossMark

Using Data on Social Influence and Collective Action for Parameterizing a Geographically-Explicit Agent-Based Model for the Diffusion of Soil Conservation Efforts

P. R. Van Oel[1] · D. W. Mulatu[2] · V. O. Odongo[1,3] · D. K. Willy[4] · A. Van der Veen[5]

Abstract

Social influence affects individual decision-making on soil conservation. Understanding the emergent diffusion of collective conservation effort is relevant to natural resource management at the river basin level. This study focuses on the effect of subjective norms and collective action on the diffusion of Soil Conservation Effort (SCE) in the Lake Naivasha basin (Kenya) for the period 1965–2010. A geographically-explicit Agent-Based Model (ABM) version of the CONSUMAT model was developed: the CONSERVAT model. In our model, we have represented heterogeneity in the physical environment and in the social network using empirical data. To parameterize the model, physical data, and social data from a household survey ($n = 307$) were used. Model simulation results show that it is possible to reproduce empirical spatiotemporal diffusion patterns of SCE levels which are quite sensitive to the way in which social survey data are used to initialize the model. Overall, this study demonstrates (i) that social survey data can effectively be used for parameterization of a geographically-explicit ABM, and (ii) that empirical knowledge on natural environment characteristics and social phenomena can be used to build an agent-based model at the river basin level. This study is an important first step towards including subjective norms for evaluating the effectiveness of alternative policy strategies for natural resource management.

1 Introduction

In many river basins, increasing resource demands lead to increased interconnectedness of human behavior and the water system [33, 47, 55]. The conservation of natural resources partly depends on the collective effort by individual actors. To effectively build on this collective action by individuals, management organizations need to take into account the way in which individuals make decisions in the context of their natural and social environments. The process of the diffusion of technologies or innovations over a social system is well-studied [40]. However, the roles of its drivers are not well-known, particularly in relation to water and soil conservation. The theory on the diffusion of innovations [42] describes how innovations diffuse over a network of socially-connected users. It has been applied in many fields including economics

✉ P. R. Van Oel
pieter.vanoel@wur.nl

D. W. Mulatu
dawitwoubishet@yahoo.com

V. O. Odongo
v.o.odongo@utwente.nl

D. K. Willy
kyalodaniel@gmail.com

A. Van der Veen
a.vanderveen@utwente.nl

[1] Water Resources Management Group, Wageningen University, PO Box 47, 6700 AA Wageningen, The Netherlands

[2] Environment and Climate Research Center (ECRC), Ethiopian Development Research Institute (EDRI), P.O.Box 2479, Addis Ababa, Ethiopia

[3] Egerton University, P.O. Box 536, Egerton, Nakuru, Kenya

[4] School of Agriculture and Enterprise Development, Kenyatta University, PO Box 43844, Nairobi 00100, Kenya

[5] University of Twente, PO Box 217 7500, AE, Enschede, The Netherlands

🌳 Springer

and marketing ([5], see also [22]). For capturing the important role of human behavior in socio-ecological systems, while accounting for the relevant heterogeneity of the physical and social environment, Agent-Based Modeling (ABM) or geo-simulation (when geographically-explicit) continues to gain popularity [2, 30, 41, 53]. Successful examples of ABM models for representation diffusion of innovation include Schwarz and Ernst [46] and Kuandykov and Sokolov [27] and articles in a special issue on the topic [7]. Relevant literature on applications to conservation issues does exist [4, 51], in particular in relation to agriculture ([3, 45], Van Duinen 2015). Methodological challenges include those related to (1) the design and parameterizing of agent decision models [49]; (2) verification, validation, and sensitivity analysis [6, 19]; and (3) the role of social networks and subjective norms. A subjective norm is the perceived social pressure to perform or not to perform the behavior [1]. In literature on social networks, it is argued that participation in social networks and distribution of motivations among populations are a function of network size, weak and strong ties, and elite influence [48]. Applied to the adoption of innovations, the probability of adoption is influenced by the characteristics of the innovation, the characteristics of the agents (including the position in the social network), and the characteristics of the environmental context (e.g., geographical settings).

The main purpose of the paper is to demonstrate that empirical data on subjective norms and participation in collective action can be used to parameterize a geographically-explicit ABM, simulating the diffusion of soil conservation effort. Next to demonstrating how our modeling approach allows for making use of these empirical data, we also evaluate the effect of selecting different approaches for model parameterization to encapsulate particular heterogeneous features of the local case-study environment. For these purposes our newly developed CONSERVAT model is introduced. The name CONSERVAT is inspired by earlier model applications labeled ANIMAT (simple robots, [61]) and CONSUMAT (consumer behavior, [13, 14], Kangur et al. in press). Using CONSERVAT as a modified version of the original CONSUMAT approach allows us to study the effects of social influence and collective action in a real-world physical environment.

1.1 The Lake Naivasha Basin

Lakes within endorheic basins often support unique ecosystems and are sensitive to variations in the inflow of nutrients and sediments, affected by human activity in upstream areas. For such lakes, excessive influxes of nutrients and sediments could lead to undesired rises in turbidity levels and sedimentation rates. Sedimentation in lakes and eutrophication may occur simultaneously as sediment loading often relates to phosphorus loading (e.g., [20]). These sedimentation rate

increases, due to soil erosion, often result from unsustainable agricultural practices. However, soil erosion (soil loss) from agricultural areas can be influenced through the adoption of soil conservation practices by farmers. Lake Naivasha (Fig. 1) is a fragile (and degraded) endorheic lake ecosystem supporting a rich variety of flora and fauna [11, 57]. Economic developments in the area are strongly connected to the lake ecosystem and the hydrological processes in the Lake Naivasha basin (3400 km^2; [34, 37]). The upstream part of the Lake Naivasha basin comprises the Gilgil, Malewa, and Turasha sub-basins (Fig. 1). The lake ecosystem is vulnerable to physical-chemical degradation by nutrients and sediments and over-abstraction of water resources (e.g. [8]). Many human activities depend on clean freshwater from the lake whereas both the availability and quality of freshwater vary considerably over space and time [36]. In addition to natural climatic variations, water availability is affected by recent investments in irrigated agriculture and by population growth [28, 39]. These developments have substantially increased water demand and unsustainable agricultural land use practices. In particular directly around Lake Naivasha, water demand (both groundwater and surface water) has increased, fueled by the large-scale production of cut flowers for export markets [56]. Challenges faced by smallholder farmers throughout the upstream parts of the basin are associated with unsustainable agricultural land use practices [60] and high rates of soil loss [38] and soil fertility loss partly counterbalanced by soil conservation efforts [60]. Drivers of erosion include seasonal cropping patterns and land fragmentation. Farming occurs on steep slopes that require land management practices to ensure sustainable agricultural use. The study by Willy and Holm-Müller [60] explored the effect of subjective norms and participation in collective action initiatives on the SCE in the Lake Naivasha basin. A subjective norm is the social pressure that makes people feel obliged to undertake a behavior (in our case, adoption of a technology) just because important people within their social networks feel they should do so or have already done it themselves. In Willy and Holm-Mueller (2013), subjective norms were found to have played a significant role in driving the diffusion of soil conservation technologies (Willy and Holm-Mueller 2013). Their study also showed that participation in collective action and subjective norms differ between local populations and that this affects SCE levels. In this study, we take these spatial differences into account and evaluate the effects of this heterogeneity.

2 Method

To assess the processes leading to soil erosion and subsequent sedimentation downstream, the relevant physical processes and social mechanisms need to be analyzed in an integrated

Fig. 1 The three upstream sub-basins of the Lake Niavasha basin: Gilgil (G), Malewa (M), and Turasha (T). The locations of 307 survey respondents [60] are also shown
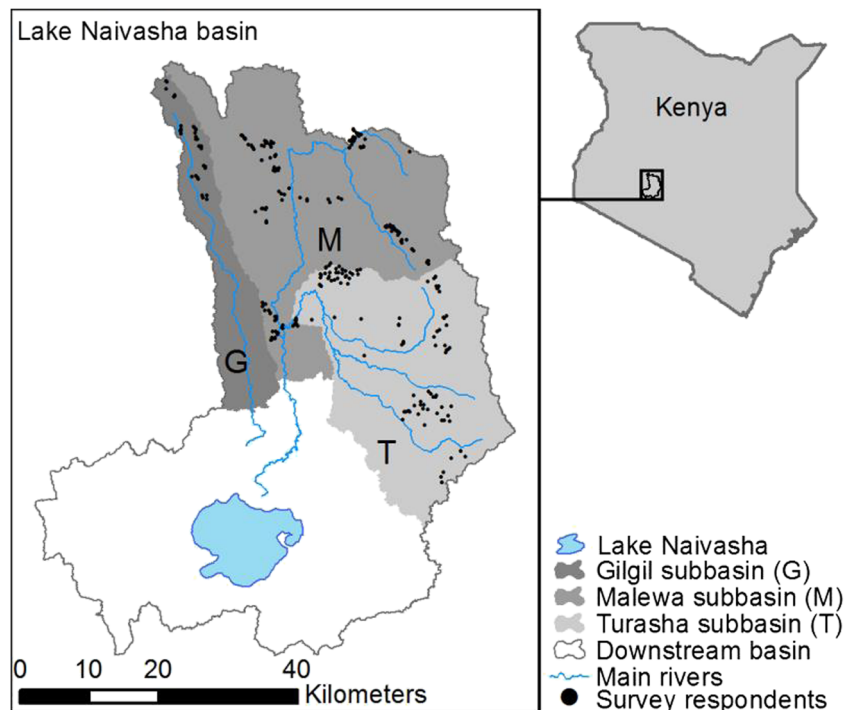
way. In the CONSERVAT model, basin level processes (i.e., sediment fluxes associated with erosion and sedimentation) emerge from micro-scale local processes that are influenced by the activities of local farm households (soil conservation practices in this case). Relevant processes (related to hydrology, soil loss, and human behavior) are represented in a geographically-explicit agent-based model. In this section, first the CONSERVAT model is introduced (Section 2.1), then the method for model parameterization is described (Section 2.2), followed by a description of model simulations and evaluation of results (Section 2.3).

2.1 The CONSERVAT Model

Our CONSERVAT model is described using standard elements of the ODD protocol [9]. Here we briefly introduce the model, while the full ODD description is included in the Appendix.

The CONSERVAT model is used to evaluate subjective norms on (i) the spatiotemporal diffusion pattern of Soil Conservation Effort (SCE) levels; and (ii) the resulting reduction in lake sedimentation rates for Lake Naivasha. In the model, smallholder farm-households are represented by agents at the local level (local catchments, see Fig. 2). Groups of neighboring agents (within a maximum distance of five kilometers) represent peers in the social environment of individual agents. The modeled natural environment consists of a spatial river basin structure with accompanying hydrological and soil loss processes. The spatial river basin structure includes local catchments for which soil loss is

calculated annually. Collections of adjacent local catchments form sub-basins for which rainfall-runoff is determined monthly. At the downstream end of the Lake Naivasha basin, a lake and a connected aquifer are modeled using a simple monthly water balance model [56]. The overall spatial structure of the CONSERVAT model is presented in Fig. 2. State variables include Farm-household strategy (repetition, optimizing, inquiring, and imitation) and Farmer choice (regarding the adoption of an elevated SCE level: from level 0 to level 1 or 2). State variables that are related to hydrological and soil loss processes include runoff in sub-basin streams (m^3 $month^{-1}$), water in the lake (m^3) and in the connected aquifer (m^3) and soil loss from farms and local catchments (tons $year^{-1}$), sediment inflow into the lake, (tons $year^{-1}$) and sedimentation on the lake bottom (mm $year^{-1}$).

Modeling the process of adoption of a SCE level using ABM is based on the CONSUMAT approach [13, 16]. It is well-known that in modeling, in the interaction between natural phenomena and social processes, there is a confusing amount of theories that can be chosen to describe social and individual behaviors. Not to say that there are conflicting theories [44]. Focusing on economics, mainstream economics is often narrowed down to rational behavior in a neo-classical paradigm. In Janssen and Jager [18], Jager & Janssen [12], and Kangur et al. [21], it is discussed how individuals are engaged in broader cognitive processes than only rationality, in deciding how to behave. On the one hand, individuals want to satisfy their needs, but on the other hand they are uncertain, with respect to the difference between expected and actual outcomes of their behavior. These two aspects lead to four
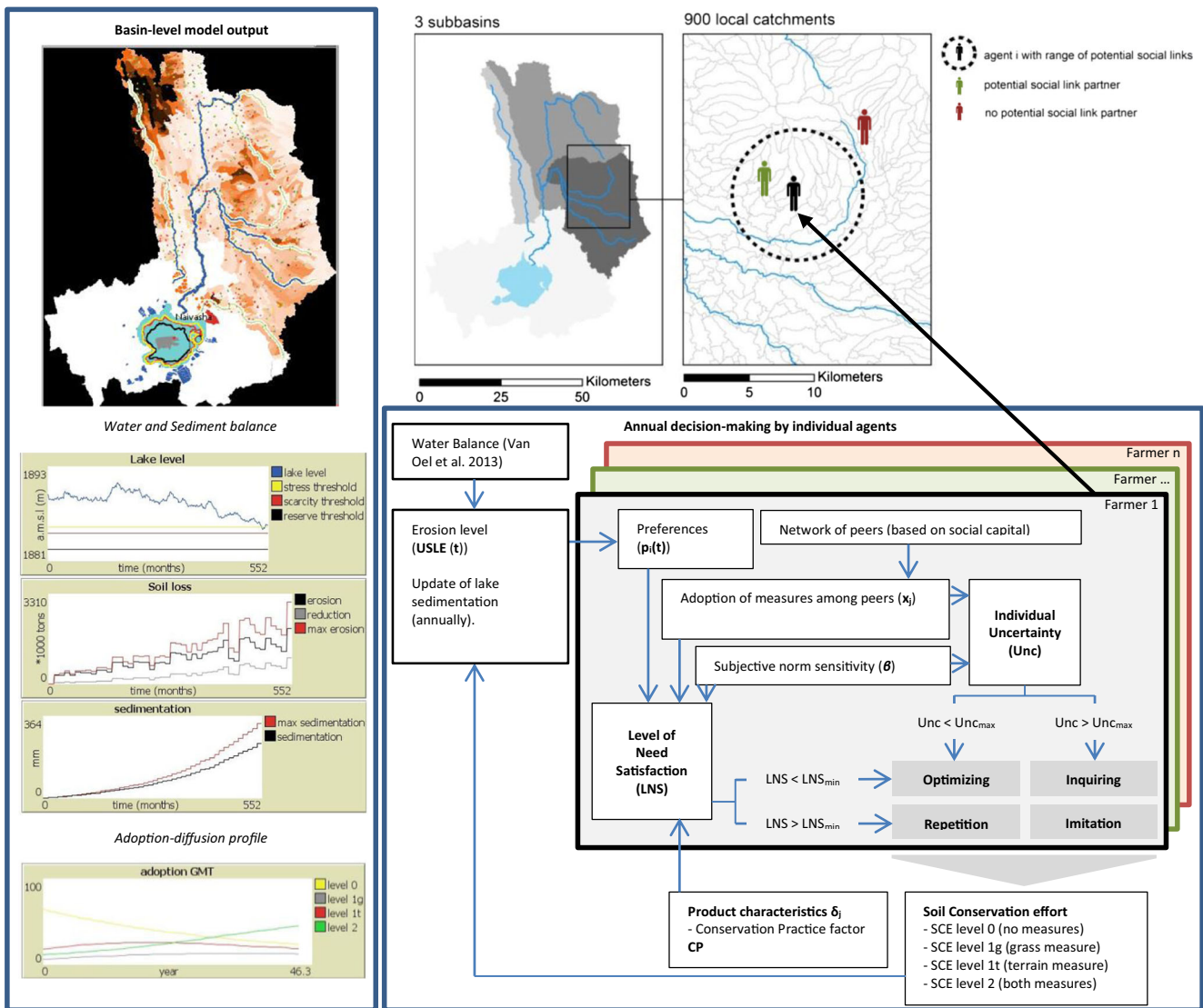
Fig. 2 Top right: The lake Naivasha basin with three sub-basins (left frame) and local catchments in which agents are located. Bottom right: The variables in the farmer box are fully explained in Table 1. The Unified Soil Loss Equation (USLE), accounts for Erosivity (R), Erodability (K), Topography (LS), Land Cover (C) and Conservation Practices (CP) as defined by Wischmeier and Smith [62]. The preference ($p_i(t)$) depends on recently experienced soil loss (personal preferences for soil conservation). The rule that is applied for updating agent preferences for soil loss reduction is shown in Table 2. Besides personal preferences, the adopted strategy (optimizing, repetition, inquiring, or imitation) also depends on characteristics of social networks. Left: Parts of the graphical user interface (GUI) of the CONSERVAT model including a dynamic geographical map, water levels in Lake Naivasha, soil loss in the basin, sedimentation in Lake Naivasha and an adoption-diffusion pattern. For a more detailed representation of the GUI, see the Appendix. The model is also available at OpenABM: https://www.openabm.org/model/4597/version/1/view.

possible behaviors: repetition (i.e., repeat past habits, imitation (i.e., copy others), inquiring (compare to others to decide), and optimizing (i.e., deliberation). Our approach builds on a version of the CONSUMAT model as described in Janssen [17]. Levels of uncertainty (Unc) and need satisfaction (LNS) provoke an agent to decide on the soil conservation effort to perform, according to one out of the four possible strategic behaviors (Fig. 2).

In our CONSERVAT model, both personal preferences for soil conservation and characteristics of social networks affect decisions of agents. Consequently, subjective-norm effects are

incorporated into agent decision making. The strategy applied by an individual agent depends on an expected Level of Need Satisfaction (*LNS*) and the Uncertainty (*Unc*) a person faces with regard to adopting an elevated SCE level. This uncertainty is influenced by the adopted SCE levels of peers. Agents compare their expected LNS and uncertainty, with a minimum required LNS (LNS_{min}) and maximum tolerated uncertainty level (Unc_{max}) and consequently choose one out of four strategies to apply: repetition, optimizing, inquiring and imitation.

In line with Wejnert [58], as discussed above, the need to conserve soil (LNS of a SCE level) is influenced by

characteristics of the adopter, characteristics of the product (the SCE level, resulting in soil conservation), and the environmental context:

- a personal preference for soil conservation of a farmer (p_i) in the context of the actual local soil loss situation, see Table Appendix 3 (Appendix)
- a product characteristic (δ_j), (i.e., the erosion reduction factor associated with a SCE level at a specific location)
- the share of peers employing a particular SCE level (x_j)

The level of uncertainty regarding a certain SCE level (Unc_{ij}) is also influenced by the share of peers employing that particular level of soil conservation effort (x_j). The following equations are used for determining the expected level of need satisfaction (LNS) and uncertainty (Unc) of farmer i and SCE level j:

$$LNS_{ij} = \beta_i\left(1 - /\delta_j - p_i/\right)x_j \qquad (1)$$

$$Unc_{ij} = (1 - \beta_i)(1 - x_j) \qquad (2)$$

where β_i reflects the insensitivity of a farmer to its social environment. Equation 1 reflects how personal preferences and those of peers affect personal satisfaction levels. Equation 2 describes how uncertainty levels of individuals are affected by the observed share of peers employing a particular SCE level (x_j). Agents with a high β are less sensitive to the choices of others than those with a low β. In our CONSERVAT model, agents can choose between four consecutive and increasing SCE levels:

- SCE level 0: No conservation measures applied
- SCE level 1 g: Grass measures applied (planting grass strips, mostly using Napier grass). The conservation practice factor (CP) is between 0 and 0.9 depending on the location.
- SCE level 1 t: Terrain measures applied (bench terraces or contour farming). The conservation practice factor (CP) is between 0 and 0.9 depending on the location.
- SCE level 2: Both measures are applied. The conservation practice factor (CP) is the combined effect of the two erosion reduction factors δ.

The CONSERVAT model is developed to run for the period 1965–2010 using a monthly time step. Over the simulation period 1965–2010, the agent population is kept constant at a level of 10,000 agents (households) for the entire Lake Naivasha basin. The majority of these agents is positioned in one of three sub-basins: Gilgil (G), Malewa (M), and Turasha (T), see Fig. 1. The chosen agent-population is based on the estimated average household size of around eight in 1965 [63], corresponding to a population of ~ 80,000 with their livelihoods being dependent on small-scale farming practices. No agents disappear and no new agents are created. Also, the social links between the agents and their peers remain unchanged. To still account for an increase in actual population over the period 1965–2010, the intensity of farmland use (and related soil loss) between 1965 and 2010 is assumed to gradually increase by a factor 3.5, along with actual population growth in the area: ~ 80,000 around 1965 and ~ 284,000 in 2009 [23–26].

2.2 Parameterization of the CONSERVAT Model

For the description of the parameterization method for human behavior, we use the framework proposed by Smajgl et al. [49], see Fig. 3 and Table 1. In Table 1, the methods applied are described along with a summary of the data used. The initialization of agent characteristics from our survey data is done in two alternative ways. Agent attribute values are either assumed to be homogeneous for all three sub-basins together (GMT) or specified heterogeneously by sub-basin (G, M, and T) based on corresponding sub-samples of the social survey data (Fig. 1). As stated in the "Introduction" section and as is shown in Table 1, our survey data show interesting differences between the three sub basins for the variables "Participation in collective action: CAPART", "Perception that soil erosion is a problem: PECEROYES", and "Sensitivity to subjective norms: SUBNORM". The differences for these variables are expected to affect the diffusion of SCE levels over the basin. Therefore, we compare a situation in which the entire agent population is regarded as homogeneous over the entire study area and a situation in which sub-sets of the household survey data-set are used to initialize populations of sub-basins.

For calibration of our CONSERVAT model the following two additional parameters have been introduced:

- Probability of adoption (prob_adoption): This fraction determines the number of agents that annually considers to alter (elevate) their soil conservation effort (SCE). Only those agents engage in decision-making according to the scheme shown in Fig. 2
- Soil loss calibration (RKLS_calib): In order to achieve realistic values of overall soil loss [62] the USLE parameters (RKLS) are multiplied by a factor. This is done to account for the effects of erosion-inducing activities that add to those accounted for in the USLE method. This also includes erosion from bare lands and roads surrounding farmland. A factor of 2.5 was found to be realistic and results in sedimentation rates that are in the same range of Stoof-Leichenring et al. (2011).

Fig. 3 Framework for parameterization of the CONSERVAT model, adapted from Smajgl et al. [49]. The implementation of these methods (M1–M5) is included on Table 1
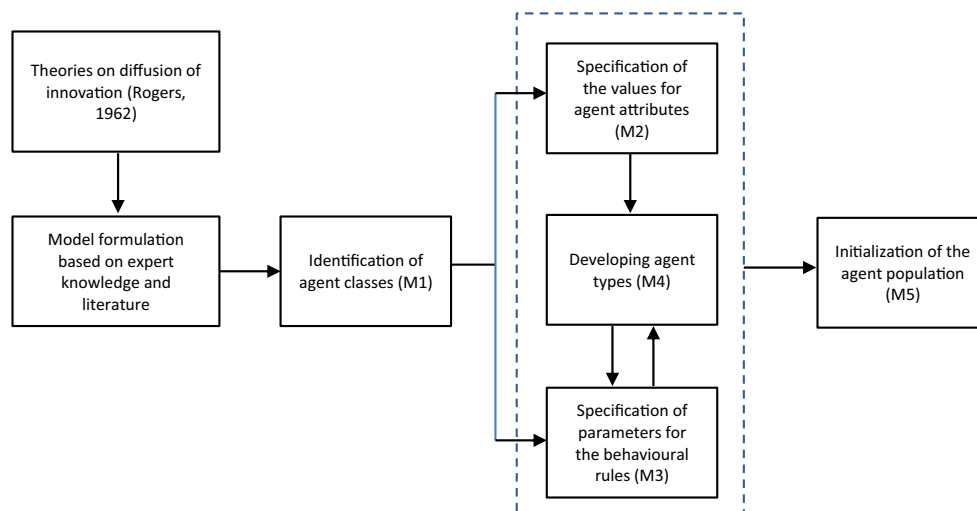
2.3 Model Simulations and Evaluation of Results

We test the effect of different ways for initializing the model using data from a household survey on spatiotemporal diffusion pattern of adopted SCE and on the sedimentation rate of Lake Naivasha. Note that in the previous section, spatial scale resolution (either lumped, i.e., covering the entire basin, or distributed to sub-basins) is an important issue for parameterization. Therefore, the following simulation specifications and initializations apply:

- Homogenous initialization: Differences among individuals in agent population (10,000 agents) are distributed homogeneously over the entire study area (all three sub-basins from Fig. 1, lumped: GMT). Initialization values used are those for GMT as specified in Table 1.
- Heterogeneous initialization: The agent population is initialized according to sub-basin specific information (G, M and T) as specified in Table 1. The survey-respondent population was categorized into three sub-basins (Fig. 1).

These model runs for homogenous initialization and heterogeneous initialization are performed to explore the effects of sub-basin specific information on diffusion of SCE levels. Specific details on the initialization of the CONSERVAT model can be found in the appendix. In addition, model runs for the two reference situations are carried out to determine the benchmarks of the model output on lake sedimentation:

Reference-no SCE: The entire agent population (10,000 agents) remains at SCE level 0 throughout the simulation period.

Reference-max SCE: The entire agent population (10,000 agents) remains at SCE level 2 throughout the simulation period. An important variable to judge the performance of

our model in the next section is to evaluate how our model simulation results compare to the reconstructed spatiotemporal diffusion pattern of adopted soil conservation effort [60] for the period 1966–2010. The reconstructed spatiotemporal diffusion pattern was categorized into three sub-basins (G, M, and T) for five consecutive periods of 10 years: 1960–1969; 1970–1979; 1980–1989; 1990–1999; and 2000–2009.

Next to comparing observed and simulated spatiotemporal diffusion patterns of SCE, we also compare the simulated sedimentation rate in Lake Naivasha with an independent data set on the observed sedimentation rate in Lake Naivasha [52].

Finally, to test the sensitivity of our model outcomes, a local sensitivity analysis has been applied in which the main model parameters are varied separately. To evaluate the influence of randomness, the model was tested for three arbitrarily chosen random-seeds. For all final simulation runs, the random seed has been fixed to a single value to avoid differences in initialization of model runs.

3 Results

3.1 Model Simulation Results on Diffusion Patterns

Results of our simulations are presented in Figs. 4 and 5. The outcomes presented in Fig. 4 show how the diffusion of SCE levels evolves over time. Figure 4 reveals a considerable difference between the results of homogeneous and heterogeneous initializations. This is the effect of using sample subsets of sub-basin populations for agent initialization (heterogeneous initialization) rather than using the entire household survey respondent set for initialization (homogeneous initialization). The difference is partly explained by the different ratios between the number of survey samples and the actual agent population that is used to represent the farmer

Table 1 Overview of methods used for parameterization of agent behavior in the CONSERVAT model. The distribution of the values for parameterization of agent attributes is either assumed homogeneous for all three sub-basins together (GMT) or heterogeneous, i.e., specified by sub-basin (G,M and T)

Details on the method applied, described following the framework by Smajgl et al. [49]	Sub-basin distribution for variables from the household survey [60]			

Identification of agent classes (M1)

There are two main agent classes: agents with low participation in collective action initiatives and agents with high participation in collective initiatives. The distinction between these two groups of agents is based on a variable by Willy and Holm-Müller [60], indicating the participation in collective action by the respondents (social survey variable: CAPART). Agents can have a maximum of either three or six social links[a] constituting their social network

Participation in collective action initiatives (CAPART = 1)

	GMT[b]	G	M	T
	49%	41%	43%	64%

Specification of the values for agent attributes (M2)

A portion of the agent population does not perceive erosion to be a problem in their local environment. This proportion of agents is based on the proportion of respondents that have indicated erosion to be a problem (social survey variable: PECEROYES). In the CONSERVAT model agent decisions on increasing soil conservation effort (SCE) levels are influenced by their preference p (Fig. 2, Table Appendix 3) for limiting erosion

Perception that soil erosion is a problem (PECEROYES = 1)

	GMT[b]	G	M	T
	50%	59%	47%	51%

Specification of the parameters for the behavioral rules that agents follow (M3)

The sensitivity to subjective norms (i.e., the sensitivity to subjective norms rather than personal preference) is determined by β (see Fig. 2 and Eqs. 1 and 2). The values of β are chosen based on the social survey variable: SUBNORM

Respondent would adopt a "conservation technique" because those important to him/her think he/she should (SUBNORM)

	GMT[b]	G	M	T
1, strongly agree: $\beta = 0.20$	36%	38%	34%	40%
2, agree: $\beta = 0.40$	35%	32%	32%	40%
3, neutral/disagree: $\beta = 0.60$	18%	15%	24%	11%
4, strongly disagree: $\beta = 0.80$	11%	15%	10%	10%

Developing agent types (M4)

There is only one agent type for the entire modeled population. The initially adopted soil conservation effort levels differ by sub-basin. When these (and other) differences are accounted for in initializing SCE levels (heterogeneous initialization) this can be regarded as cluster analysis based on the spatial locations of survey respondents

Initially adopted SCE levels

SCE level	GMT[b]	G	M	T
0	64%	100%	75%	28%
1 g (grass)	6%	0%	8%	5%
1 t (terrain)	17%	0%	8%	29%
2	13%	0%	8%	19%

Initialization of the agent population (M5)

We up-scale to 10,000 households using our representative survey of 307 respondents, randomly selected from the entire study area population [60]. The differences in population density are accounted for by the relative share of the land use type "farmland" obtained from a land use classification of remotely-sensed data for 2011

Number of agents in CONSERVAT

	GMT[b]	G	M	T
	9252[c]	1483	3451	4318

Number of survey respondents that started farming before 2010

	GMT[b]	G	M	T
	296	37	152	107
Ratio respondents/agents	3.2%	2.5%	4.4%	2.5%

[a] The maximum amount of social links between individual agents, either three or six social links, is chosen on expert judgment. The distinction is made to separate between those respondents that indicted to participate in collective action (attributed a maximum of six links) and those that did not (attributed a maximum of three links)

[b] The combination of the three sub-basins: Gilgil (G), Malawa (M), and Turasha (T)

[c] This is less than 10,000 since 748 agents, located in the Karati sub-basin, are not included in this analysis

population. For the homogeneous initialization, the respondents from the Malewa (M) sub-basin are relatively over-represented and respondents from the other sub-basins are relatively under-represented (Table 1, under M5). This plays a role in all stages of parameterization of agent behavior in the CONSERVAT model (Table 1, Fig. 3). Another part is explained by the diffusion process itself, since SCE levels do not easily rise in places with low initial SCE levels.

In Fig. 5, model simulation results are presented for the entire upper basin (homogeneous initialization: aggregated basin GMT) and by sub-basin (heterogeneous initialization: sub-basins G, M, and T). A visual comparison between model simulation results and empirical data on the reconstructed adoption pattern of soil conservation technology shows that the simulation results are in the same range as the empirical data. There are however considerable differences between the

Fig. 4 Model output for
CONSERVAT simulations of the
diffusion of SCE levels 0, 1, and 2
for homogeneous and
heterogeneous initializations

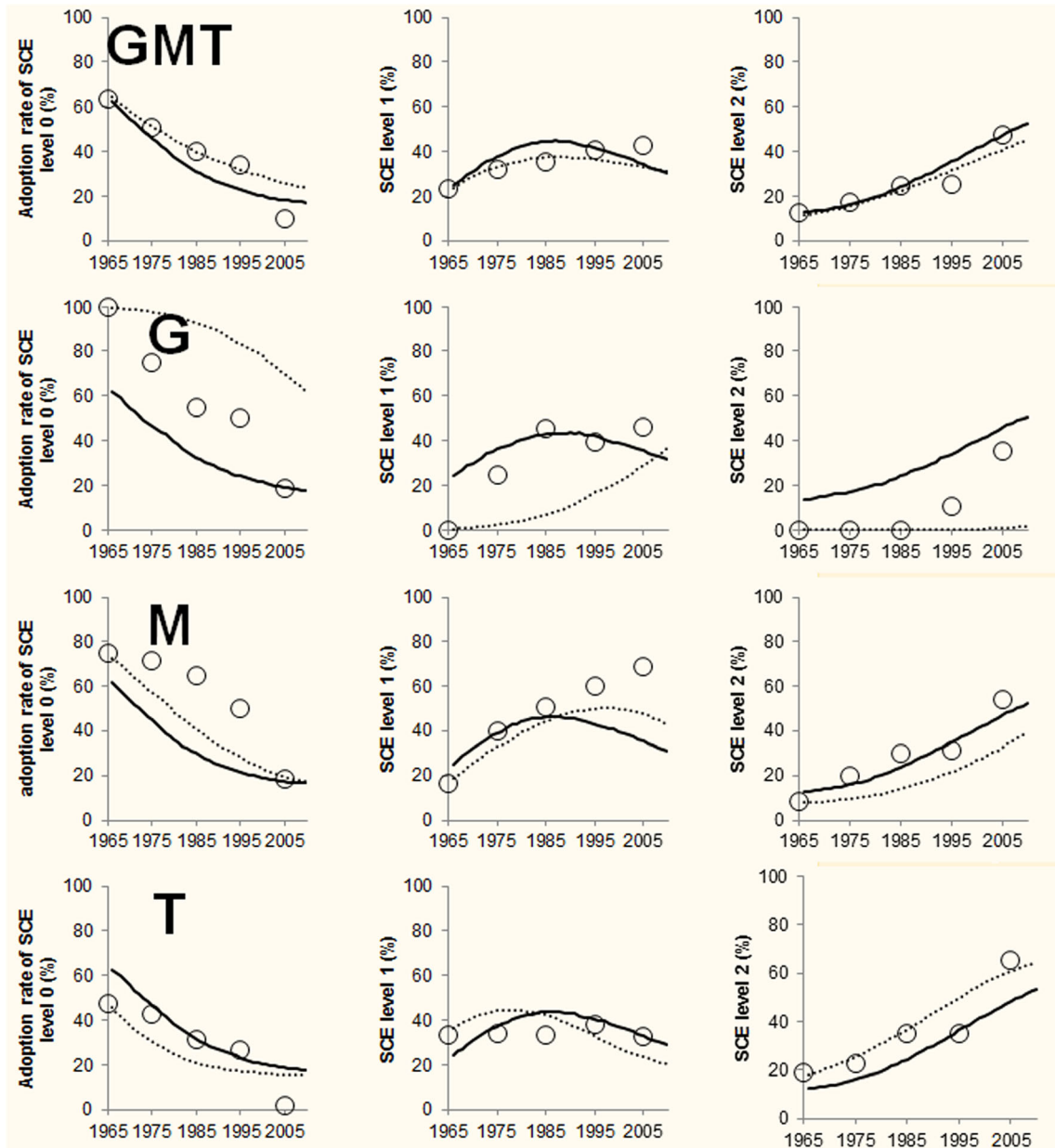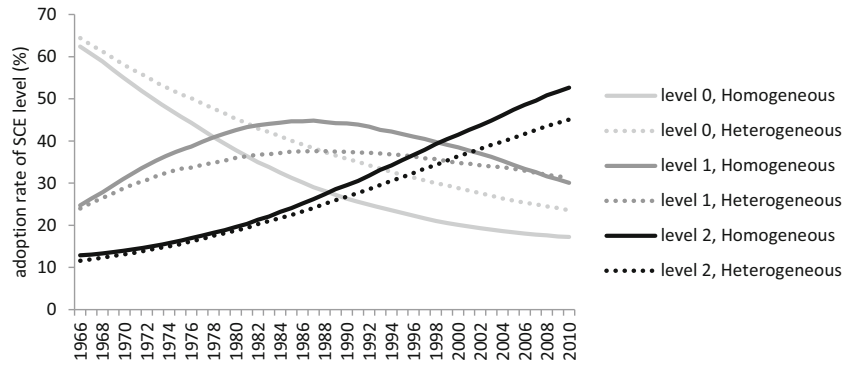**Fig. 5** Model results for the diffusion of soil conservation effort for different initial levels of effort: SCE 0, left panel, SCE 1, middle panel and SCE 2, right panel. Homogeneous initialization (solid line); heterogeneous initialization (dotted line); empirical data (circles; observed by [60]). First row, entire upper basin (GMT), followed by the sub-basins of Gilgil (G), Malewa (M), and Turasha (T)

three selected sub-basins, particularly in relation to the type of initialization applied.

The difference between simulation outcomes for homogeneous and heterogeneous initializations is particularly large for the Gilgil (G) sub-basin. This can be explained by the relatively large difference between initially adopted SCE levels of both initializations for the agent population. For the Gilgil (G) sub-basin, the diffusion of adoption is relatively slow and limited presumably due to below-average values for CAPART (implying less social links per agent), and a high sensitivity to the local subjective norm (SUBNORM) which in this case hampers the diffusion process since the initial adoption rate was 0% (Table 1).

For the Malewa (M) sub-basin, the diffusion of adoption is slower in case of heterogeneous initialization in comparison to homogeneous initialization. This is explained by the initial adoption rates of high SCE levels which is lower in case heterogeneous initialization. Also, the Malewa (M) sub-basin specific values for SUBNORM and CAPART (see Table 1) are less favorable for diffusion than is the case for Homogeneous initialization.

For the Turasha (T) sub-basin, the opposite is true from the Malewa (M) sub-basin. The initial adoption rate of high SCE levels is much higher in case of Homogeneous initialization. The initial adoption rates of high SCE levels in this sub-basin is also high in comparison to the other two sub-basins. In the case of the Turasha (T) sub-basin, the high values for sensitivity to the subjective norm (SUBNORM) and participation in collective action activities (CAPART) speed up the diffusion of a higher SCE level rather than to hamper it.

The physiological sensitivity of the area to erosion is highest in the Gilgil (G) sub-basin, as determined using the USLE approach [62]. The highest value for the people's perception of erosion to be a problem is also highest in that area (Willy and Holm-Muller, 2013). However, the diffusion of adoption of higher SCE levels is very slow in this part of the basin (Fig. 5). This applies both to our model output and the empirical data. This strongly suggests that farmers are more sensitive to subjective norms than to actual soil losses. We judge this to be an important finding.

3.2 Model Simulation Results on the Sedimentation Rate

In the data poor environment of Lake Naivasha Kenya, we have been able to compare our model simulation results to independent sample data on lake sedimentation (Stoof-Leichensring et al. 2011). In the CONSERVAT model, lake sedimentation results from soil loss due to land use practices in the upstream area. This emergent variable (sedimentation, or rather "reduced sedimentation") results from numerous individual soil conservation efforts.

If we define a situation of all agents adopting a SCE level 2 over the entire period of 1965–2010 as the maximum achievable reduction, then the actual adoption rate [60], as reproduced by our model, leads to 59% of this maximum achievable reduction. In Fig. 6, the maximum reduction is the solid line (Reference-Max-SCE). Our model simulation results for lake sedimentation (for both homogeneous and heterogeneous initializations) closely resemble the actual lake sedimentation rate [52], thus suggesting that the application

Fig. 6 Sedimentation rate of Lake Naivasha for the period 1965–2010. Our model results (for both homogeneous and heterogeneous initializations) resemble the sedimentation pattern that was found by a core sample inside Lake Naivasha [52] to a large extent
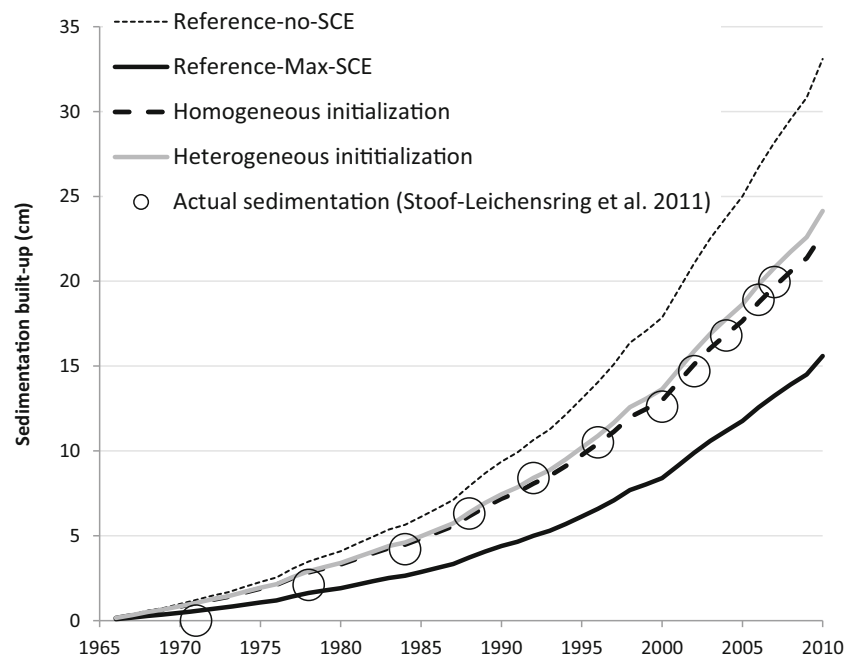
of soil conservation practices has substantially reduced lake sedimentation. Importantly, as a second result, comparing our model simulation results with the empirical sedimentation pattern [52] we show that our results catch the upward trend and change in observed lake sedimentation.

3.3 Sensitivity Analysis

The results of the local sensitivity analysis [10] are presented in Table 2. Varying the random seed has only very limited influence on simulation results for both SCE levels and sedimentation in Lake Naivasha at the end of the simulation period (2010). Both output variables are quite sensitive to variations in the calibration parameter *Prob_adoption*, whereas variations in *RKLS_calib* (the other calibration parameter) only affect sedimentation values.

In the CONSERVAT model decision-making is dominated by imitation, repetition or inquiring, depending on the settings chosen for LNS_{min} (level of need satisfaction) and Unc_{max} (individual uncertainty). For the selected settings (LNS_{min} = 0.05 and Unc_{max} = 0.05) imitation appears to be the dominant strategy. Substantially higher values for Unc_{max} result in lower

aggregate levels of SCE and consequent higher values of sedimentation. This is because the dominant agent strategy becomes *repetition*. For substantially higher levels of LNS_{min} the dominant agent strategy becomes *inquiring* which also leads to somewhat lower aggregate levels of SCE and consequent higher values of sedimentation.

The effect of initialization (homogeneous or heterogeneous) is also considerable as already shown (Figs. 4–6).

4 Discussion

The approach that was used for parameterization of the CONSERVAT model has been described using the framework of Smajgl et al. [49] including Methods M1, M2, M3, M4, and M5 as referred to in Fig. 3 and Table 1. The five methods applied are discussed below. For the identification of agent classes (M1), a very simple distinction was made between survey respondents [60] that indicated to participate in collective action activities and those that did not. Using this information, it was decided to allot the former with a maximum of six social links and the others only three. This rather pragmatic

Table 2 Results of the local sensitivity analyses. All parameters relate to homogeneous initialization, except for the last line

| Parameter (varied for lumped initialization) | Parameter (P) value | Sensitivity regarding SCE in 2010 | | Sensitivity regarding sedimentation (s) in 2010 | |
| | | % of maximum effort (reference-max-SCE) | | | |
		SCE value in 2010 (%)	(dSCE/SCE)/ (dP/P) $S^{-/+}$	s value in 2010 (mm)	(ds/s)/(dP/P) $S^{-/+}$
Random-seed	a (Netlogo random-seed 111)	67.7		228	
	b (Netlogo random-seed 222)	68.0	–	229	–
	c (Netlogo random-seed 333)	67.6	–	228	–
Prob_adoption	0.13	67.7		205	
	0.10 (0.13–0.03)	58.9	0.562	241	− 0.239
	0.16 (0.13 + 0.03)	73.0	0.340	221	− 0.143
RKLS_calib	2.50	67.7		228	
	2.25 (2.50–0.25)	67.7	0.000	205	1.000
	2.75 (2.50 + 0.25)	67.7	0.000	251	1.000
LNSmin	0.050	67.7		228	
	0.025 (0.050–0.025)	67.7	0.002	228	0.000
	0.100 (0.050 + 0.050)	67.8	0.002	228	0.000
	0.500 (0.050 + 0.450)	49.8	− 0.029	248	0.010
UNCmax	0.050	67.7		228	
	0.025 (0.050–0.025)	67.7	0.001	227	− 0.007
	0.100 (0.050 + 0.050)	65.4	− 0.034	231	0.023
	0.500 (0.050 + 0.450)	27.9	− 0.065	279	0.025
Initial values Homogeneous initialization	Parameters based on entire survey set	67.7		228	
Initial values Heterogeneous initialization	Parameters specified by sub-sample sets	60.8	–	241	–

procedure is admittedly quite arbitrary. Further exploration of the number of social links could yield insights in the propagation of the actual diffusion. In this respect, it would be quite interesting to explore the influence of alternative designs of social networks such as opinion-leader, village, or hierarchical networks [48].

In the CONSERVAT model, agents and their social network do not change during the period 1965–2010. However, in reality agent-specific attributes and behavior do change and also the social networks change. In order to maintain an acceptable level of simplicity and transparency of the model, the model was kept as simple as possible while still allowing to show how subjective norms affect diffusion of soil conservation efforts in the study area. This simplification of reality implies that an agent in the CONSERVAT model represents a number of actual farmers in a certain spatial unit.

A fundamental problem in social sciences and ABM is how to derive observations of a social system over time [19]. Fortunately, the survey by Willy and Holm-Müller [60] allowed us to reconstruct the diffusion of SCE levels for the Lake Naivasha basin for the period 1960–2010. However, respondent data on changes (over time) in sensitivity to subjective norms (SUBNORM), participation in collective activities (CAPART) and perceptions that soil erosion is a problem (PECEROYES) were not available for this study. For the specification of the value for the agent attribute for "the perception that erosion is a problem" (M2), some methodical considerations deserve attention. As we have assumed in our model it would be quite well conceivable that when erosion occurs it will increasingly be perceived as a problem. However, from the survey parameter PERCEROYES we cannot derive changes in perceptions by survey respondents over time. Moreover, it is questionable whether the local erosion experienced (compared to the previous 3 years) is the only important factor for farm-household preference in how much erosion has occurred. Farm-household preference could also be influenced by other factors such as repetitive occurrence of erosion events (consecutive wet years can trigger somebody to act eventually), the introduction of mechanical tools, access to funding, support by conservation organizations, and the introduction of market-based conservation practices in the basin.

Respondents from the Malewa sub-basin are over-represented in the farm-household survey. This skewed representation of sub-set respondents relative to the entire survey affects model results as seen from the difference between the outcomes (Table 1). The influence of different spatial scales for initialization (homogeneous and heterogeneous initializations; methods M4, M5) has been explored in this study. As seen in Table 2 it makes a lot of difference when other spatial units are used to aggregate survey data for model initialization. However, aiming at water management, soil conservation and the reduction of sedimentation, partitioning of a social survey into three sub-basins seems the logical choice. But, other spatial units could have been used as well.

Despite the limitations of the approach used, the results of this study suggest that using household survey data for parameterizing a basin-level ABM is an important element to practitioners in land and water management. Identifying spatial differences allows decision-makers to evaluate where interventions could be most effective. The added value of using an ABM in addition to the already informative survey data is that the innovation diffusion process (affected by subjective norms) is included in the modeling process. In order to increase added value of using ABM for land and water management, surveys should be carefully designed, with this purpose of the ABM in mind.

Two parameters for calibration were used: prob_adoption and RKLS_calib. The former parameter (prob_adoption) affects the speed of the diffusion process. This parameter does not affect the decision model of individual agents. It merely affects the number of agents that engage in decision making for a particular time step. Other factors that affect the speed of the diffusion process are the time-step used for decision-making (1 year) and the number of links to the peers. With regard to the other calibration parameter (RKLS_calib) erosion other than farmland areas (roads, wind) may also be very relevant, however uncertain. Moreover, the position of soil loss has not been explicitly considered when translated into sedimentation of Lake Naivasha. This means that erosion and sedimentation inside streams and rivers due to runoff variations (for example: flash floods) have not been considered.

Using a reconstructed time-series is not ideal. It is based on respondents that look back on their own agricultural history (sometimes all the way back to 1960). The accuracy of individual memories is debatable. Also, farmers that stopped farming before 2011 could have never been part of the survey. The people now looking back to 1960s were very young back then. It is arguably not representative for the farmer population around the 1960s. For more robust data, an additional survey would be needed and considerable additional effort would be necessary. This is however beyond the scope of the current study.

Also with regard to the observations for sedimentation in Lake Naivasha some critical remarks are justified. Whether the sediment core used [52] is representative for the sedimentation in the entire lake is questionable. Using only one particular position in the lake could easily lead to an under- or overestimation. Also, the process of sedimentation distribution over the lake bottom is likely to be influenced by lake level variations (hydraulic effects at mouth of the river) which have not been accounted for.

Although a sensitivity analysis has been performed for this study (Table 2) there are several aspects that could be further

explored. Adequate validation of the CONSERVAT approach could be done by applying quantitative methods and systematic sensitivity analysis (Lorscheid et al. 2011). In particular, alternative settings of Unc_{max} and LNS_{min} and their influence on strategic behavior (repetition, imitation, inquiring, and optimizing) of agents could be further explored. Similarly, initializing and running our model using three random-seed settings only, does not guarantee reliable model results. For this study a visual comparison (qualitative validation) served the purpose of demonstrating the usefulness of using farm-household survey data for parameterization and qualitative validation of agent-based models. However, follow-up studies are advised to put considerable effort in conducting thorough and systematic validation.

5 Conclusion

This study shows that household survey data on subjective norms and collective action can effectively be used to parameterize a geographically-explicit agent-based model for natural resource management. Particularly in data-poor environments, the use of social survey data for model-parameterization could advance the development of tools for evaluating the basin-level effects of management alternatives for soil conservation in both the natural and social domains.

The CONSERVAT approach that has been introduced in this paper allows for including both natural influence and social influences (i.e., subjective norms) on the diffusion of soil conservation efforts. It has been demonstrated that individual decisions on adopting soil conservation technologies are affected by subjective norms and that it is possible to reproduce this adoption process. Our demonstration also shows that different approaches for model parameterization result in relevant differences in model outcomes. This study is innovative in the sense that we incorporate empirical data on subjective norms in a geographically-explicit ABM that integrates the physical and social system for an environmental issue. This is an important addition to recent studies incorporating social networks [29, 54]. The model is developed as a first step to effectively use ABM for supporting decision-making on natural resource management (soil conservation).

Appendix ODD

ODD Overview

Purpose of the Model

The CONSERVAT model can be used to evaluate the effects of subjective norms among smallholder farmers in the Lake Naivasha basin on:

i)   The spatiotemporal diffusion pattern of Soil Conservation Effort (SCE) levels;
ii)   The resulting reduction in lake sedimentation rates for Lake Naivasha.

In the present study, we demonstrate that social survey data [60] can effectively be used for parameterization and validation of a geographically-explicit ABM, developed to simulate the effects of subjective norms on the diffusion of soil conservation effort.

Entities, State Variables and Scales

Entities Smallholder farm-households are represented by agents at the local level (local catchments, see Fig. 2 and Figure Appendix 7). Groups of neighboring agents (within a maximum distance of 5 km) represent peers in the social environment of individual agents. The modeled natural environment consists of a spatial river basin structure with accompanying hydrological and soil loss processes. The spatial river basin structure includes local catchments for which soil loss is calculated annually. Collections of adjacent local catchments form sub-basins for which rainfall-runoff is determined monthly. At the downstream end of the Lake Naivasha basin a lake and a connected aquifer are modeled using a simple monthly water balance model [56]. The overall spatial structure of the CONSERVAT model is presented in Fig. 2. State variables include Farm-household strategy (consisting of four human behavioral strategies: repetition; optimization; inquiring; imitation) and Farmer choice (regarding the adoption of an elevated SCE level). State variables that are related to hydrological and soil loss processes include Runoff in sub-basin streams (m^3 $month^{-1}$), water in the lake (m^3) and in the connected aquifer (m^3) and soil loss from farms and local catchments (tons $year^{-1}$), sediment inflow into the lake, (tons $year^{-1}$) and sedimentation on the lake bottom (mm $year^{-1}$).

The following scale-related characteristics apply:

–   Spatial extent: Lake Naivasha basin ($\sim 3400 km^2$)
–   Sub-basins (based on river gauging stations obtained from the Water Resources Management Authority, Naivasha, Kenya in 2012)
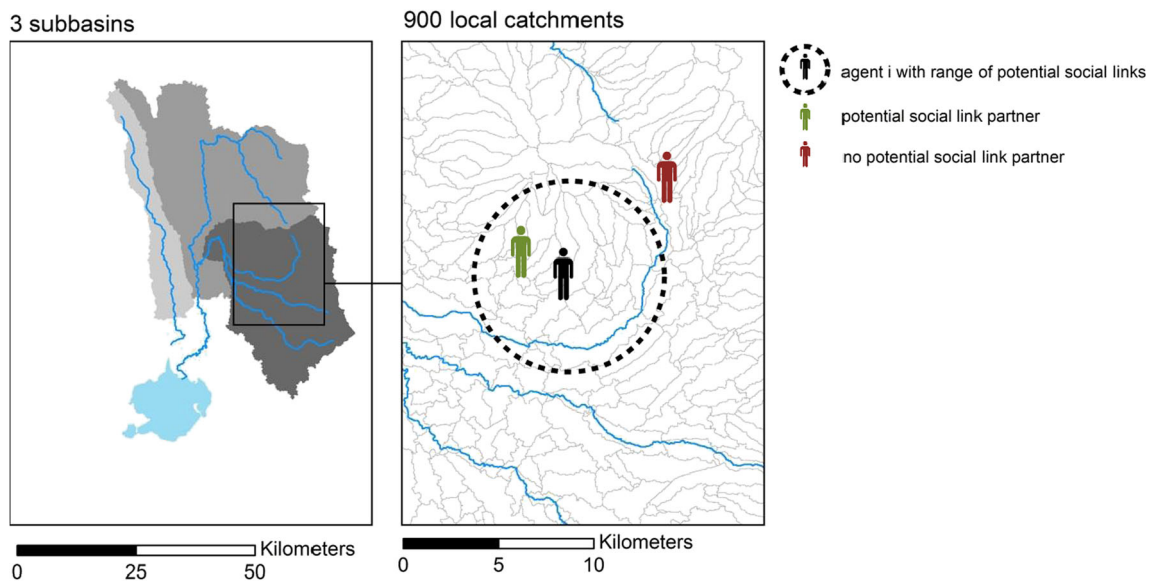
Fig. 7 The lake Naivasha basin with three sub-basins (left frame) and local catchments in which agents are located

– Local catchments are the smallest spatial unit in the model at which farmers-agents are located. For these local catchments environmental parameters are lumped (± 2 km^2)
– Average size of farm-household farmland (cropland): 1 ha
– Temporal extent: from 1965 to 2010
– Temporal resolution: 1 month for hydrological processes, 1 year for farmer decision-making; and for soil loss calculations.

Process Overview and Scheduling

The modeling sequence is the following. First, agents decide on the adoption of a SCE level using the CONSERVAT approach. Second, in the physical parameter update, natural run-off and soil losses are determined. Third, in the biophysical dynamics, the lake water balance and sedimentation are updated.

The CONSERVAT Approach Modeling the process of adoption of a SCE level using ABM is based on the CONSUMAT approach [13, 16]. In Janssen and Jager [18], it is presented how individuals are engaged in different cognitive processes in deciding how to behave. On the one hand, individuals want to satisfy their needs, but on the other hand they are uncertain, with respect to the difference between expected and actual outcomes of their behavior. These two aspects lead to four possible behaviors: repetition, deliberation, imitation, and social comparison. By including new insights from literature on (inter alia) interactions in groups, and social networks the four possible behaviors were recently coined as repetition, imitation, inquiring, and optimizing [12]. Our approach is adapted

from a CONSUMAT version as described in Janssen [17]. Levels of uncertainty (Unc) and need satisfaction (LNS) provoke an agent to decide on the soil conservation effort to perform, according to one out of four alternative possible strategic behaviors (Figure Appendix 8).

Like in the CONSUMAT models, in our CONSERVAT model personal preferences and characteristics of social networks affect decisions of agents. Consequently, subjective-norm effects are incorporated into agent decision-making. The strategy applied by an individual agent depends on an expected Level of Need Satisfaction (*LNS*) and the Uncertainty (*Unc*) a person faces with regard to adopting an elevated SCE level. This uncertainty is influenced by the adopted SCE levels of peers. Agents compare their expected LNS and Uncertainty, with a minimum required LNS (LNS_{min}) and maximum tolerated uncertainty level (Unc_{max}) and consequently choose one out of four strategies to apply: repetition, deliberation, inquiring, and imitation.

In line with Wejnert [58], the LNS of a SCE level (LNS_{ij}) is assumed to be influenced by characteristics of the adopter, characteristics of the product (the soil conservation effort), and the environmental context:

• a personal preference of farmer i (p_i) in the context of the actual local situation (i.e., actual local soil loss), see Table Appendix 3
• a product characteristic (δ_j), (i.e., the erosion reduction factor associated with a SCE level at a specific location)
• the share of peers employing a particular SCE level (x_j)

The level of uncertainty regarding a certain SCE level (Unc_{ij}) is also influenced by the share of peers employing that

Fig. 8 The Conservat approach applied in this study. The approach is a variation of the Consumat approach [13]. The variables in the farmer box are explained in Table 2. The Unified Soil Loss Equation (USLE), accounts for Erosivity (R), Erodability (K), Topography (LS), Land Cover (C) and Conservation Practices (CP) as defined by Wischmeier and Smith [62]. The preference (pi(t)) depends on recently experienced soil loss. The rule that is applied for updating agent preferences for soil loss reduction is shown in Table 2
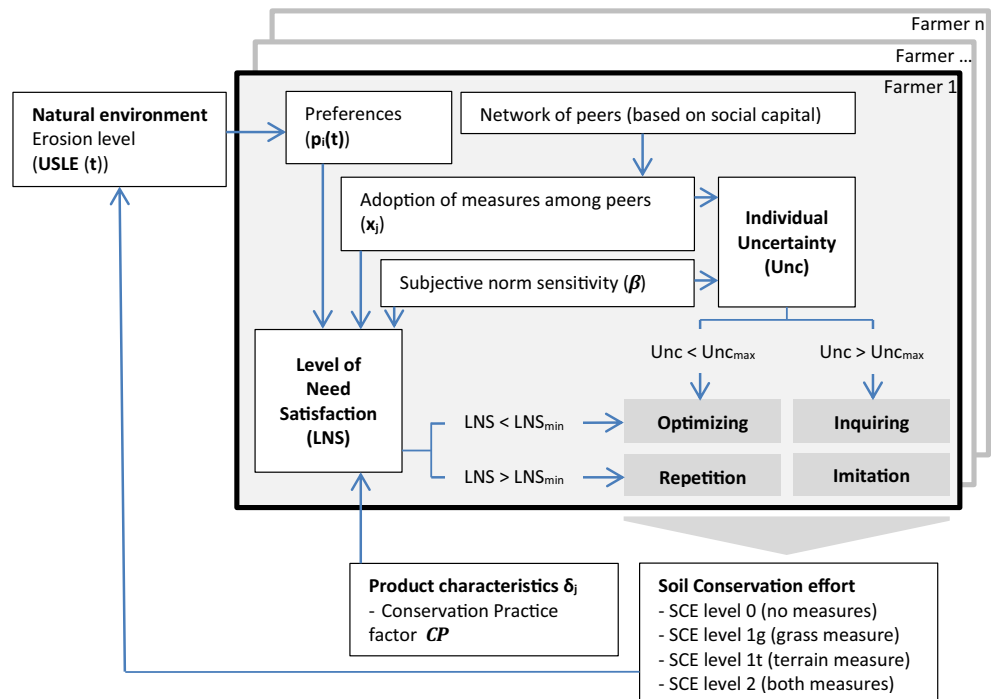
particular level of soil conservation effort (x_j). The following equations [17] are used for determining the expected level of need satisfaction (LNS) and uncertainty (Unc) of farmer i and SCE level j:

$$LNS_{ij} = \beta_i \left(1 - /\delta_j - p_i /\right) + (1 - \beta_i) x_j \quad (1)$$

$$Unc_{ij} = (1 - \beta_i)(1 - x_j) \quad (2)$$

where β_i reflects the insensitivity of a farmer to its social environment. Equation 1 reflects how personal preferences and those of peers affect personal satisfaction levels. Equation 2 describes how uncertainty levels of individuals are affected by the observed share of peers employing a particular SCE level (x_j). Agents with a high β are less sensitive to the choices of others than those with a low β. In our CONSERVAT model,

agents can choose between four consecutive and increasing SCE levels:

- SCE level 0: No conservation measures applied
- SCE level 1g: Grass measures applied (planting grass strips, mostly using Napier grass). The conservation practice factor (CP) is between 0 and 0.9 depending on the location.
- SCE level 1t: Terrain measures applied (bench terraces or contour farming). The conservation practice factor (CP) is between 0 and 0.9 depending on the location.
- SCE level 2: Both measures are applied. The conservation practice factor (CP) is the combined effect of the two erosion reduction factors.

To determine the farmers' level of need satisfaction (LNS) preferences are annually updated. Preferences are assumed to be influenced by actual local soil loss (soil erosion level in a

Table 3 Farmer preference updating rule. Farmer preference (p_i (t)) depends on recently experienced soil loss

Local soil loss condition (time t)	Preference value of farmer	
No severe erosion has been experienced (at time t) (soil loss (USLE) < average of last 3 years)	0	
	If erosion is not perceived to be a problem: PECEROYES = 0 [60]	If erosion is perceived to be a problem: PECEROYES = 1 [60]
Severe erosion has been experienced (soil loss (USLE) > average of last 3 years)	0	Random fraction of 1[a]

[a] As it is not very clear from the survey [60] what factors affect the farmer perception on the extent of erosion to be a problem, we decided to simply assign a random value

particular year, as compared to the average over the previous 3 years). This preference is annually updated (Figure Appendix 8) and depends on the actual erosion in the local area. Moreover, LNS is affected by variables that reflect sensitivity to subjective norms and social capital. These variables are taken from a survey by Willy and Holm-Müller [60] conducted in 2011 among 308 randomly selected farm-households in the study area. Our main perception variables from this survey are CAPART (indicating participation in collective action), PECEROYES (score on the statement that soil erosion is a problem), and the subjective norm variable SUBNORM (score on the statement: "I would adopt a technology just because those persons that are important to me think I should").

These variables (CAPART, PECEROYES, and SUBNORM) have been used for parameterization of the agent-population and have been used to implement the two basic model equations (agent-decision rules) and to define the number of peers (agents in social network) to be included in the procedure. The procedure assumes the same theoretical base as Janssen [17]. In terms of Eqs. 1 and 2, the household survey data are used to define the following parameters:

- The number of peers to be included to determine the adoption rate among peers (x_i): The household survey variable for participation in collective action (CAPART) has been used to divide the agent population into agents with three peers and agents with 6six peers.
- Preference ($p(t)$): The household survey variable for perception that erosion is a problem (PECEROYES) has been used to divide the agent population into agents that do not perceive erosion to be a problem and those that do. Those that do will update their preference variable $p(t)$ according to the erosion level experienced over a period of the past 3 years.
- The sensitivity to subjective norms (β): The household survey for subjective norm (SUBNORM) has been used to divide the agent population into four categories, specifying the agent's sensitivity to subjective norms.

For calibration of our CONSERVAT model the following two additional parameters have been introduced:

- Probability of adoption (prob_adoption): This fraction determines the number of agents that annually considers to alter (elevate) their soil conservation effort (SCE). Only those agents engage in decision-making according to the scheme shown in Figure Appendix 8
- Soil loss calibration (RKLS_calib): In order to achieve realistic values of overall soil loss [62] the USLE parameters (RKLS) are multiplied by a factor. This is done to account for the effects of erosion inducing activities that add to those accounted for in the USLE method. This also

includes erosion from bare lands and roads surrounding farmland.

ODD Design Concepts

The model design is such that decisions at the individual farmer household level affect emergent lake sedimentation (by inflow of sediments). Decisions of farmers on employing a certain SCE level (0, 1g, 1t, 2) are assumed to be based on personal preferences, the characteristics of the innovation and the characteristics of the social network [58]. The soil loss characteristics are specific for the geographic location at which agents are located. The characteristics of the social network is counted by the agents' social links with a maximum of either three or six randomly selected peers in a radius of 5 km (variable CAPART in the household survey of Willy and Holm-Müller [60], see Figure Appendix 7). The maximum number of social links between individual agents, either three or six social links, is chosen without sound empirical ground. However, the distinction is necessary to separate between those respondents that indicted to participate in collective action (attributed a maximum of six links) and those that did not (attributed a maximum of three links).

With regard to stochasticity, the initial rates of adoption are taken from the farm-household survey. For model calibration, a parameter for the probability of considering the adoption of an elevated SCE level (prob_adoption, random number) is introduced.

Overall observations from the simulations include adopted SCE levels, strategies adopted, LNS, uncertainty levels, reduced soil loss, the sedimentation rate in Lake Naivasha and the water level of Lake Naivasha (m above mean sea level). In Figure Appendix 9 a snapshot of the simple modeling interface is shown.

ODD Details

Initialization

The CONSERVAT model is developed to run for the period 1965–2010 using a monthly time step. For initializing the agent population and land use (in 1965), estimates are based on population census data [23–26] and land use classifications for the Lake Naivasha basin area for 1973 and 2011 [37], respectively. The size of cultivated cropland was estimated to be 1 ha which is in accordance with farm-household survey data of Mulatu et al. [35]; and Willy and Holm-Müller [60].

Over the simulation period 1965–2010, the agent population is kept constant: 10,000 agents (households) for the entire Lake Naivasha basin. The majority of these agents is positioned in one of three sub-basins: Gilgil (G), Malewa (M), and Turasha (T), see Fig. 1. The chosen agent-population is

Fig. 9 A snapshot of the simple modeling interface of the Naivasha CONSERVAT model. Implemented in Netlogo [59], version 5.0.3.

based on the estimated average household size of around eight in 1965 [63], corresponding to a population of ~ 80,000 with their livelihoods being dependent on small-scale farming practices. No agents disappear and no new agents are created. Also the social links between the agents and their peers remain unchanged. To still account for an increase in actual population over the period 1965–2010 the intensity of farmland use (and related soil loss) between 1965 and 2010 is assumed to gradually increase by a factor 3.5, along with actual population growth in the area: ~ 80,000 around 1965 and ~ 284,000 in 2009 [23–26].

For the description of the parameterization method for human behavior use is made of the framework proposed by Smajgl et al. [49], see Fig. 3 and Table 1. The terminology as used by Smajgl et al. [49] is adopted. Accordingly, in Table 1 the methods applied are described along with a summary of the data used. The initialization of agent characteristics from our survey data is done in two alternative ways. Agent attribute values are either assumed to be homogeneous for all three sub-basins together (GMT) or specified by sub-basin (G, M and T) based on corresponding sub-samples of the social survey data-set. As stated in the "Introduction" section and as is shown in Table 1, our survey data show interesting differences between the three sub-basins analyzed for the variables used (CAPART,

PECEROYES and SUBNORM). The differences for these variables are expected to affect the diffusion of SCE levels over the basin. Therefore we compare a situation in which the entire agent population is regarded as homogeneous over the entire study area and a situation in which sub-sets of the household survey data-set are used to initialize populations of sub-basins.

For the initialization of erosion, reduction factors at the level of local catchments Support Practice Factors (CP) as described by [62] have been used. CP varies with land slope percentage and is specified for alternative levels of soil conservation effort (Table Appendix 4). For SCE level 1g

Table 4 Erosion reduction factors of alternative levels of soil conservation effort. The CP factors are taken from Wischmeier and Smith [62]

Soil conservation effort (SCE) level	Local erosion reduction factor (1-CPfactor) (δ_j)
SCE level 0	0
SCE level 1g (grass)	$1 - CP_{strip\text{-}cropping}$
SCE level 1t (terrain)	$1 - CP_{contouring}$
SCE level 2	$1 - CP_{strip\text{-}cropping} * CP_{contouring}$

(planting of Napier grass and applying Grass strips in Willy and Holm-Müller [60]) the C*P* values for strip-cropping [62] are assigned. For SCE level 1t (Contour farming and Terraces in Willy and Holm-Müller [60]) *P* values for contouring [62] are assigned.

Input Data For calculating soil losses and the water balance of Lake Naivasha, two submodels are used: a hydrological sub-model and a soil loss sub-model.

Rainfall-Runoff Sub-Model Monthly rainfall data are used as input and are transformed into runoff using average rainfall-runoff coefficients for the different months of the year based discharge estimates and rainfall estimates at the level of the sub-basins based [31] for the period 1960–1985. For the Lake Naivasha water balance we use a model by Van Oel et al. [56]. This means that also water abstractions from the lake and the connected aquifer have been accounted for. Water abstractions from the upstream parts of the Lake Naivasha basin have not been accounted for in this version of the Conservat model.

Soil Loss Sub-Model Sedimentation of the lake has been accelerating from an average 2 mm year^{-1} in the 1960s to around 10 mm year^{-1} in the recent decade [52]. An important cause of this sedimentation increase is attributed to soil loss in the upper parts of the Lake Naivasha basin [38]. For simplicity it is assumed that runoff from the upper parts of the basin is responsible for all the sedimentation in Lake Naivasha, thus ignoring the possible influence of wind and local influxes from the banks. It is also assumed that soil erosion from cropland is responsible for all soil loss in the basin. This also means that erosion from roads and urban lands are simply ignored in this study and a calibration parameter (*RKLS_calib*) is justified. The Universal Soil Loss Equation (USLE) method [62] has been used for estimation of gross erosion rates in the different discretized cells of the local catchments. The estimation of soil erosion within a grid-cell is expressed as:

$$SE_i = R^* \, K_i^* \, LS_i^* \, C_i^* \, P_i \qquad (A1)$$

where SE_i = gross amount of soil erosion in cell i (MT ha^{-1} year^{-1}); R = rainfall erosivity factor (MJ mm ha^{-1} h^{-1} year^{-1}); K_i = soil erodibility factor in cell i (MT ha h ha^{-1} MJ^{-1} mm^{-1}); LS_i = slope steepness and length factor for cell i (dimensionless); C_i = cover management factor i (dimensionless), and P_i = supporting practice factor for cell i (dimensionless). To estimate the parameters, a range of studies were used including Roose [43] for R, Sombroek and Braun [50] for K, Mitasova et al. [32] for LS, Jain and Das [15] and Wischmeier and Smith [62] for C, and Wischmeier and Smith

[62] for *P*. Soil losses which are calibrated to agree with data on sedimentation data [52].

References

1. Ajzen, I., Fishbein, M., (1980). Understanding attitudes and predicting social behaviour. Englewood Cliffs, Prentice-Hall, New Jersey, USA.

2. An, L. (2012). Modeling human decisions in coupled human and natural systems: review of agent-based models. *Ecological Modelling, 229*, 25–36.

3. Berger, T. (2001). Agent-based spatial models applied to agriculture: a simulation tool for technology diffusion, resource use changes and policy analysis. *Agricultural Economics, 25*(2–3), 245–260.

4. Daloglu, I., Nassauer, J. I., Riolo, R., & Scavia, D. (2014). An integrated social and ecological modeling framework-impacts of agricultural conservation practices on water quality. *Ecology and Society, 19*(3).

5. Fagerberg, J., Fosaas, M., & Sapprasert, K. (2012). Innovation: exploring the knowledge base. *Research Policy, 41*(7), 1132–1153.

6. Filatova, T., Verburg, P. H., Parker, D. C., & Stannard, C. A. (2013). Spatial agent-based models for socio-ecological systems: challenges and prospects. *Environmental Modelling and Software, 45*, 1–7.

7. Garcia, R., & Jager, W. (2011). From the special issue editors: agent-based modeling of innovation diffusion. *Journal of Product Innovation Management, 28*(2), 148–151.

8. Gherardi, F., Robert Britton, J., Mavuti, K. M., Pacini, N., Grey, J., Tricarico, E., & Harper, D. M. (2011). A review of allodiversity in Lake Naivasha, Kenya: developing conservation actions to protect East African lakes from the negative impacts of alien species. *Biological Conservation, 144*(11), 2585–2596. https://doi.org/10.1016/j.biocon.2011.07.020.

9. Grimm, V., Berger, U., DeAngelis, D. L., Polhill, J. G., Giske, J., & Railsback, S. F. (2010). The ODD protocol: a review and first update. *Ecological Modelling, 221*(23), 2760–2768.

10. Hamby, D. M. (1994). A review of techniques for parameter sensitivity analysis of environmental models. *Environmental Monitoring and Assessment, 32*(2), 135–154.

11. Harper, D. M., & Mavuti, K. (2004). Lake Naivasha, Kenya: ecohydrology to guide the management of a tropical protected area. *Ecohydrology and Hydrobiology, 4*(3), 287–305.

12. Jager, W., & Janssen, M. A. (2012). An updated conceptual framework for integrated modeling of human decision making the Consumat II Workshop Compexitry in the Real World @ ECCS 2012 - from policy intelligence to intelligent policy 5–6 September 2012. Brussels.

13. Jager, W., Janssen, M. A., De Vries, H. J. M., De Greef, J., & Vlek, C. A. J. (2000). Behaviour in commons dilemmas: Homo economicus and Homo psychologicus in an ecological-economic model. *Ecological Economics, 35*(3), 357–379.

14. Jager, W., Janssen, M. A., & Vlek, C. (2001). Experimentation with household dynamics: the consumat approach. *International Journal of Sustainable Development, 4*(1), 90–100.

15. Jain, M. K., & Das, D. (2010). Estimation of sediment yield and areas of soil erosion and deposition for watershed prioritization using GIS and remote sensing. *Water Resources Management, 24*(10), 2091–2112.

16. Janssen, M., & Jager, W. (1999). An integrated approach to simulating behavioural processes: a case study of the lock-in of consumption patterns. *Journal of Artificial Societies and Social Simulation, 2*(2), 21–35.

17. Janssen, M. A. (2013). Introduction to agent-based modeling. In *Tempe*. USA: Open ABM.

18. Janssen, M. A., & Jager, W. (2001). Fashions, habits and changing preferences: simulation of psychological factors affecting market dynamics. *Journal of Economic Psychology, 22*(6), 745–772.

19. Janssen, M. A., & Ostrom, E. (2006). Empirically based, agent-based models. *Ecology and Society, 11*(2).

20. Jeppesen, E., Kronvang, B., Meerhoff, M., Søndergaard, M., Hansen, K. M., Andersen, H. E., et al. (2009). Climate change effects on runoff, catchment phosphorus loading and lake ecological state, and potential adaptations. *Journal of Environmental Quality, 38*(5), 1930–1941.

21. Kangur, A., Jager, W., Verbrugge, R., & Bockarjova, M. (2017). An agent-based model for diffusion of electric vehicles. *Journal of Environmental Psychology.* https://doi.org/10.1016/j.jenvp.2017.01.002.

22. Karakaya, E., Hidalgo, A., & Nuur, C. (2014). Diffusion of eco-innovations: a review. *Renewable and Sustainable Energy Reviews, 33*, 392–399.

23. KNBS. (1979). Population and housing census. Kenya National Bureau of Statistics (KNBS). Nairobi, Kenya.

24. KNBS. (1989). Population and housing census. Kenya National Bureau of Statistics (KNBS). Nairobi, Kenya.

25. KNBS. (1999). *Population and housing census*. Nairobi: Kenya National Bureau of Statistics (KNBS).

26. KNBS. (2009). *Population and housing census*. Kenya National Bureau of Statistics (KNBS), Nairobi.

27. Kuandykov, L., & Sokolov, M. (2010). Impact of social neighborhood on diffusion of innovation S-curve. *Decision Support Systems, 48*(4), 531–535.

28. Kuhn, A., Britz, W., Willy, D. K., & van Oel, P. (2016). Simulating the viability of water institutions under volatile rainfall conditions – the case of the Lake Naivasha Basin. *Environmental Modelling & Software, 75*, 373–387.

29. Manson, S. M., Jordan, N. R., Nelson, K. C., & Brummel, R. F. (2016). Modeling the effect of social networks on adoption of multifunctional agriculture. *Environmental Modelling & Software, 75*, 388–401.

30. Matthews, R. B., Gilbert, N. G., Roach, A., Polhill, J. G., & Gotts, N. M. (2007). Agent-based land-use models: a review of applications. *Landscape Ecology, 22*(10), 1447–1459. https://doi.org/10.1007/s10980-007-9135-1.

31. Meins, F. M. (2013). Evaluation of spatial scale alternatives for hydrological modelling of the Lake Naivasha basin, Kenya. (M.Sc.), University of Twente, Enschede, The Netherlands.

32. Mitasova, H., Hofierka, J., Zlocha, M., & Iverson, L. R. (1996). Modelling topographic potential for erosion and deposition using GIS. *International Journal of Geographical Information Systems, 10*(5), 629–641.

33. Molle, F., Wester, P., & Hirsch, P. (2010). River basin closure: processes, implications and responses. *Agricultural Water Management, 97*(4), 569–577. https://doi.org/10.1016/j.agwat.2009.01.004.

34. Mulatu, D. W., van Oel, P. R., & van der Veen, A. (2015). Firms' willingness to invest in a water fund to improve water-related ecosystem services in the Lake Naivasha basin, Kenya. *Water International, 40*(3), 463–482.

35. Mulatu, D. W., Van der Veen, A., & Van Oel, P. R. (2012). Upstream resource managers' preferences for water - related ecosystem services in the Lake Naivasha Basin, Kenya. ACES and ecosystem markets 2012: a community on ecosystem services linking science, practice and decision making, Fort Lauderdale, USA.

36. Ndungu, J. N., Monger, B. C., Augustijn, D. C. M., Hulscher, S. J. M. H., Kitaka, N., & Mathooko, J. M. (2013). Evaluation of spatio-temporal variations in chlorophyll-a in Lake Naivasha, Kenya: remote-sensing approach. *International Journal of Remote Sensing, 34*(22), 8142–8155. https://doi.org/10.1080/01431161.2013.833359.

37. Odongo, V. O., Mulatu, D. W., Muthoni, F. K., Van Oel, P. R., Meins, F. M., Van der Tol, C., Van der Veen, A. (2014). Coupling socio-economic factors and eco-hydrological processes using a cascade-modeling approach. Journal Of Hydrology.

38. Odongo, V. O., Onyando, J. O., Mutua, B. M., Van Oel, P. R., & Becht, R. (2013). Sensitivity analysis and calibration of the modified universal soil loss equation (MUSLE) for the upper Malewa Catchment, Kenya. *International Journal of Sediment Research, 28*(3), 368–383. https://doi.org/10.1016/S1001-6279(13)60047-5.

39. Odongo, V. O., Van der Tol, C., Van Oel, P. R., Meins, F. M., Becht, R., Onyando, J. O., & Su, Z. (2015). Characterization of hydroclimatological trends and variability in the Lake Naivasha Basin, Kenya. *Hydrological processes.*

40. Rao, K. U., & Kishore, V. V. N. (2010). A review of technology diffusion models with special reference to renewable energy technologies. *Renewable and Sustainable Energy Reviews, 14*(3), 1070–1078.

41. Rindfuss, R., Entwisle, B., Walsh, S., Li, A., Badenoch, N., Brown, D., . . . Verburg, P. (2008). Land use change: complexity and comparisons. Journal of Land Use Science, 3(1), 1–10. doi: https://doi.org/10.1080/17474230802047955.

42. Rogers, E. M. (1962). *Diffusion of innovations*. New York: Free Press.

43. Roose, E. J. (1976). *Soil erosion: prediction and control, use of the universal soil loss equation to predict erosion in West Africa*. USA: Ankeny.

44. Schlüter, M., Baeza, A., Dressler, G., Frank, K., Groeneveld, J., Jager, W., . . . Wijermans, N. (2017). A framework for mapping and comparing behavioural theories in models of social-ecological systems. Ecological Economics, 131, 21–35.

45. Schreinemachers, P., Berger, T., Sirijinda, A., & Praneetvatakul, S. (2009). The diffusion of greenhouse agriculture in northern Thailand: combining econometrics and agent-based modeling. *Canadian Journal of Agricultural Economics, 57*(4), 513–536.

46. Schwarz, N., & Ernst, A. (2009). Agent-based modeling of the diffusion of environmental innovations - an empirical approach. *Technological Forecasting and Social Change, 76*(4), 497–511.

47. Shadkam, S., Ludwig, F., Van Oel, P., Kirmit, Ç., & Kabat, P. (2016). Impacts of climate change and water resources development on the declining inflow into Iran's Urmia Lake. *Journal of Great Lakes Research, 42*(5), 942–952.

48. Siegel, D. A. (2009). Social networks and collective action. *American Journal of Political Science, 53*(1), 122–138.

49. Smajgl, A., Brown, D. G., Valbuena, D., & Huigen, M. G. A. (2011). Empirical characterisation of agent behaviours in socio-ecological systems. *Environmental Modelling and Software, 26*(7), 837–844.

50. Sombroek, W. G., & Braun, H. M. H. (1982). Exploratory soil survey report no. E-1. Republic of Kenya. Nairobi, Kenya.

51. Sopha, B. M., Klöckner, C. A., & Hertwich, E. G. (2013). Adoption and diffusion of heating systems in Norway: coupling agent-based modeling with empirical research. *Environmental Innovation and Societal Transitions, 8*, 42–61.

52. Stoof-Leichsenring, K. R., Junginger, A., Olaka, L. A., Tiedemann, R., & Trauth, M. H. (2011). Environmental variability in Lake Naivasha, Kenya, over the last two centuries. *Journal of Paleolimnology, 45*(3), 353–367. https://doi.org/10.1007/s10933-011-9502-4.

53. Torrens, P. M. (2010). Agent-based models and the spatial sciences. *Geography Compass, 4*(5), 428–448. https://doi.org/10.1111/j.1749-8198.2009.00311.x.

54. Van Duinen, R., Filatova, T., Geurts, P., & Van der Veen, A. (2015). Coping with drought risk: empirical analysis of farmers' drought adaptation in the south-west Netherlands. *Regional Environmental Change, 15*(6), 1081–1093.

55. Van Loon, A. F., Gleeson, T., Clark, J., Van Dijk, A. I. J. M., Stahl, K., Hannaford, J., Di Baldassarre, G., Teuling, A. J., Tallaksen, L. M., Uijlenhoet, R., Hannah, D. M., Sheffield, J., Svoboda, M., Verbeiren, B., Wagener, T., Rangecroft, S., Wanders, N., & Van Lanen, H. A. J. (2016). Drought in the Anthropocene. *Nature Geoscience, 9*(2), 89–91.

56. Van Oel, P. R., Mulatu, D. W., Odongo, V. O., Meins, F. M., Hogeboom, H. J., Becht, R., et al. (2013). The effects of groundwater and surface water use on total water availability and implications for water management: the case of Lake Naivasha, Kenya.

57. Van Oel, P. R., Odongo, V. O., Mulatu, D. W., Muthoni, F. K., Ndungu, J. N., Ogada, J. O., & Van der Veen, A. (2014). Supporting IWRM through spatial integrated assessment for the Lake Naivasha basin, Kenya. *International Journal of Water Resources Development, 30*(3), 605–618.

Water Resources Management, 27(9), 3477–3492. https://doi.org/10.1007/s11269-013-0359-3.

58. Wejnert, B. (2002). Integrating models of diffusion of innovations: a conceptual framework. *Annual Review of Sociology, 28*, 297–326.

59. Wilensky, U. (1999). NetLogo. Evanston, IL: center for connected learning and computer-based modeling, Northwestern University.

60. Willy, D. K., & Holm-Müller, K. (2013). Social influence and collective action effects on farm level soil conservation effort in rural Kenya. *Ecological Economics, 90*(0), 94–103. https://doi.org/10.1016/j.ecolecon.2013.03.008.

61. Wilson, S. W. (1987). Classifier systems and the animat problem. [machine learning]. *Machine Learning, 2*(3), 199–228. https://doi.org/10.1007/bf00058679.

62. Wischmeier, W., & Smith, D. (1978). Predicting rainfall erosion losses. Agricultural Handbook no. 537, US Dept of Agriculture, Science and Education Administration.

63. World bank. (2013). World DataBank, World development indicators.