Drechsler, M. (2023):
Improving models of coordination incentives for biodiversity conservation by fitting a multi-agent simulation model to a lab experiment
J. Behav. Exp. Econ. **102** , art. 101967

Improving models of coordination incentives for biodiversity conservation by fitting a multi-agent simulation model to a lab experiment

Martin Drechsler, Helmholtz Centre for Environmental Research – UFZ, Permoserstr. 15, 04318 Leipzig, Germany, and Brandenburg University of Technology Cottbus-Senftenberg, Germany; martin.drechsler@ufz.de.

Abstract
Coordination incentives (CI) like the agglomeration bonus that reward the spatial agglomeration (or other spatial patterns) of biodiversity conservation measures are gaining increasing attention. Experiments on CI, accompanied by statistical analyses, reveal insights into the behaviour of human subjects. However, the scope of statistical models is limited and one may, as in other sciences like physics or ecology, gain additional insights by fitting mechanistic process models to the experimental data. I present the first application of this type in the context of CI and fit a multi-agent simulation model to a seminal experiment on the agglomeration bonus. Comparing two basic approaches for the decision making of the model agents, reinforcement learning and using expectations about the future, reveals that the latter is much better able to replicate the observations of the experiment. Improved models of agent behaviour are indispensable in the model-based assessment of CI for the conservation of biodiversity.

Highlights
- A multi-agent model for actors under coordination incentives is developed
- The model is used to replicate the dynamics of a seminal lab experiment
- Fitting the model to the observations reveals what drove players' decisions

Key words
Agent-based model, agglomeration bonus, biodiversity conservation, coordination incentives, inverse modelling, (land use), pattern-oriented modelling.

JEL codes
C6, C9, Q24, Q28, Q57.

1 Introduction

To halt the continuing loss of biodiversity, policies need to address both the loss and the fragmentation of species habitats. On private lands, regulations and land-use planning have limited applicability, since many biodiversity conservation measures can be implemented only on a voluntary basis. Conservation activities on private lands are therefore generally induced by economic instruments like conservation payments or tradable land-use permits (de Vries and Hanley 2016). In the face of scarce budgets, these instruments should be cost-effective, so that biodiversity is conserved at least costs.

A challenge in the design of conservation payment schemes is the spatial heterogeneity of the conservation costs. Unless the conservation agency has knowledge of these costs, it is not possible to explicitly target payments to ecologically valuable sites. To address the problem of habitat fragmentation, coordination incentives have been proposed, such as the agglomeration bonus of Parkhurst et al. (2002) in which each conserving landowner receives a spatially homogenous base payment, and on top of this a bonus for each adjacent conserved land parcel. This bonus offsets the added cost that arises if a more costly but connected land parcel is conserved rather than a less costly but isolated one (Drechsler et al. 2010).

Since 2002, an increasing amount of research papers have been published about the agglomeration bonus. These comprise modelling studies (Drechsler et al. 2010, Bell et al. 2016, Delacote et al. 2016, Ifthekar and Tisdell 2017, Dijk et al. 2017), empirical analyses (Bell et al. 2018, Krämer and Wätzold 2018, Huber et al. 2021) and lab experiments (Parkhurst et al. 2002, Parkhurst and Shogren 2007, Banerjee et al. 2012, 2014, 2017, Parkhurst et al. 2016, Kuhfuss et al. 2018, 2022); for a comprehensive review on the literature on CI, see Nguyen et al. (2022).

While the agglomeration bonus idea is intriguing, the empirical analyses have identified a number of obstacles that may hamper the effectiveness and cost-effectiveness of this instrument. Probably the most critical obstacles are the strategic uncertainty the landowners are facing under an agglomeration bonus, and the landowners' limited cognitive abilities (Parkhurst et al. 2002, Parkhurst and Shogren 2007).

The strategic uncertainty arises from the above choice between connected high-cost and isolated low-cost land parcels, which places the landowners into a coordination game (van Huyck et al. 1990, Heinemann et al. 2004). Due to the bonus, the conservation of a costly land parcel adjacent to the land parcel of another landowner is more profitable (compared to the conservation of a less costly but isolated land parcel) but this is *only if* the other landowner conserves as well. And whether the neighbour conserves or not is generally not known a priori, since their decision is dependent on the size of the bonus, their conservation cost, and the land-use in their neighbourhood. Conserving the adjacent land parcel thus is termed the payoff-dominant strategy while conserving an isolated land parcel is termed the risk-dominant strategy (Parkhurst and Shogren 2007). Risk-averse decision makers will favour the latter over the former, which can hamper coordination and the formation of spatially agglomerated pattern of conservation activities.

The second reason for such coordination failure identified by Parkhurst and Shogren (2007) is that even in their stylised model landscape in which four players managed altogether 100 land parcels, each player could choose from tens of thousands of land-use strategies, leading to thousands of Nash equilibria. Given the cognitive limitations of the average human brain, this complexity may prevent the players from finding the (most) payoff-dominant, coordinated Nash equilibrium, even if the strategic uncertainty played no role.

In the lab experiment of Parkhurst and Shogren (2007) the players were able to establish desired spatial patterns of conserved land parcels with high probabilities. While being encouraging, it is not clear how far this result carries over to landscapes with other spatial and economic features, and it is difficult to predict how the subjects in the experiment would perform under alternative incentive designs that differ from the tested ones.

To improve the transferability of the experimental results to other cases and help develop better (simulation) models of coordination incentives, it would be desirable to better understand the factors behind the players' observed decisions, or in other words, to have a look into the players' minds. Such knowledge then could be included in agent-based models to assess the performance of coordination incentives like the agglomeration bonus (Iftekhar and Tisdell 2016, Bell et al. 2018, Drechsler 2021, Drechsler et al. 2022).

Agent-based models have numerous applications in economic research, including their integration with economic experiments (Duffy 2006, Mignot and Vignes 2020). Iftekhar and Tisdell (2016), e.g., modelled agents in a conservation auction who could submit joint bids with their neighbours, which would increase their payoffs and the spatial agglomeration of conservation efforts. The authors assumed that the model agents learn from their payoffs from previous bidding rounds via reinforcement learning (Erev and Roth 1998, Charpentier et al. 2021) in order to derive their decisions. Here a successful strategy is repeated with higher probability than other strategies, but in addition, the agents test alternative similar strategies and adopt them if they lead to a higher payoff.

While the model of Iftekhar and Tisdell (2016) is of a generic nature, so that plausible ranges for their model parameters are assumed in an ad-hoc manner, agent-based models can be parametrised with data from economic lab experiments in order to consider human behaviour more specifically. Or the dynamics observed in agent-based models can be compared to and tested in lab experiments, such as Benito-Ostolaza (2015) who experimentally replicated Schelling's (1971) famous model of social segregation and could confirm some but not all of the predictions of that rather stylised model.

Despite the potential of agent-based models as complements of experimental research, integration of these two strands of research is still rare, while the majority of economic experiments is evaluated statistically. In the context of coordination incentives, Parkhurst and Shogren (2007), e.g., fitted a statistical model to the observed coordination success of the four players that included predictors such as "complexity" (number of neighbouring players affecting the own payoff) and "experience" (round of play in the session). Banerjee et al. (2012) predicted current player decisions from own and neighbours' past decisions. And Parkhurst et al. (2002) fitted a (probabilistic) Markov model to the time series of player decisions to predict the steady state of the dynamics.

Statistical models use observable parameters (such as previous player decisions) as predictors. On a sufficient data base they can provide excellent predictions of the future dynamics of a system. For instance, the statistical model of Banerjee et al. (2014) predicts how communication and information provision affect the coordination of players; and with a similar approach, Banerjee et al. (2017) assessed the impact of transaction costs on coordination success.

However, the transferability of statistical models to other systems with different settings, conditions and constraints is often limited. Here an alternative are mechanistic (not to be confused with "mechanic"), process-based models that describe the behaviour and interactions of the entities (players) through mathematical equations or rules (in the present context, see as examples the agent-based models of Iftekhar and Tisdell (2016), Bell et al. (2018) and Drechsler (2021)). To illustrate the complementarity of statistical and mechanistic models, consider that one can forecast the weather by fitting a statistical (e.g., neural network) model to observed atmospheric data, or by fitting a mechanistic model built on the physical laws of thermodynamics. Both approaches have their pros and cons, where an advantage of mechanistic models is that they can generally be better adapted to other situations.

Fitting a mechanistic process model to observation data is called inverse modelling (Hartig et al. 2011). Since the task here is to define patterns of the system's spatio-temporal dynamics (such as the temporal mean or variance of a variable, or it spatial distribution) and identify parameter values of the process model so that the simulated patterns match the observed patterns, the approach is also termed "pattern-oriented modelling" (Grimm et al. 2005).

As an example from experimental economics, Shank et al. (2015) fitted an agent-based simulation model to observed dynamics of a public good game experiment in which players contribute, at some individual cost, to some public good whose level depends on the players' contributions. Among other results, the model in which players showed preferences both for their own and for the group payoffs fitted the data best.

While (at maximum) a handful of studies have fitted agent-based models to experimental data, none exists to the author's knowledge in the context of coordination incentives – which obviously involve particular challenges in the identification of appropriate decision models. The present paper provides a first application in this context.

As a test case, the above-mentioned study of Parkhurst and Shogren (2007) is taken. Two types of decision models are tested. The first is the above-mentioned reinforcement model of Erev and Roth (1998) that has by now a long history of several decades of application and is used in many economic contexts in which actors learn and adapt (Charpentier et al. 2021). The second model is novel and somewhat complementary to the first. Rather than basing decisions on past experiences, it considers expectations of the future, in particular the expected future decisions of the other agents. This model, termed henceforth the "future-expectations model", is motivated by the fact that in a coordination problem like the present one, with a large number of Nash equilibria, the actors' beliefs of what other actors will do play a central role (Russell et al. 1990, Hellwig 2002, Neumann and Vogt 2009). Therefore the future-expectations model explicitly considers the players' beliefs of how the future land-use decision of a neighbour depends on the current land use.

The dynamics induced by both agent decision models are simulated, and the emerging land-use dynamics are compared to those observed in the experiment of Parkhurst and Shogren (2007). It is analysed how the parameters of the two decision models affect the modeled land-use dynamics, and which values of the model parameters lead to the best “model fit”, i.e. the highest agreement between modeled and observed dynamics.

As a main result, I show that the reinforcement-learning model is unable to replicate the data, while for the future-expectations model one can find combinations of parameter values so that the model output agrees with the experimental data. The evaluation of these parameter combination allows for some conclusions about the factors that may have driven the decisions of the subjects in the experiment.

2 Methods

2.1 The experiment of Parkhurst and Shogren, and its model simulation

Parkhurst and Shogren (2007) arranged four players on a square grid with ten by ten land parcels, so that each player owned one quarter of the model landscape (Fig. 1). Each land parcel could be used for economic purposes (e.g., intensive agriculture) or for conservation. Conservation caused a loss of revenue (henceforth termed “conservation cost”) that was lowest at the northern and southern borders of the landscape and highest in the middle, while not depending on the east-west location of the land parcel.

2	2	2	2	2	2	2	2	2	2
4	4	4	4	4	4	4	4	4	4
6	6	6	6	6	6	6	6	6	6
8	8	8	8	8	8	8	8	8	8
10	10	10	10	10	10	10	10	10	10
10	10	10	10	10	10	10	10	10	10
8	8	8	8	8	8	8	8	8	8
6	6	6	6	6	6	6	6	6	6
4	4	4	4	4	4	4	4	4	4
2	2	2	2	2	2	2	2	2	2

Figure 1. Model landscape of the experiment. Each player owned one quarter of the land parcels. The numbers represent the conservation costs (forgone agricultural revenues when the land parcel is conserved).

Each player was offered a spatially homogenous base payment $bp = 3$ (per land parcel) for the conservation of up to six land parcels. Under this setting, profit-maximising players would conserve the low-cost land parcels in the north and the south. To induce agglomerated conservation further in the middle of the landscape, the authors offered, in addition to the base payment, three agglomeration bonuses:

- an own-border bonus *obb* that was paid for each border between own conserved land parcels (which would, e.g., sum up to four times *obb* if the five adjacent land parcels on the outer row were conserved),
- a row-border bonus *rbb* that was paid for each border a conserved land parcel had to a conserved land parcel of another player that lied north or south, and
- a column-border bonus *cbb* that was paid for each border a conserved land parcel had to a conserved land parcel of another player that lied east or west.

Three target patterns of conserved land parcels (Fig. 2) were incentivised through an appropriate combination of bonuses (identified by the authors through a numerical game-theoretic analysis). Note that the authors also considered a fourth pattern with conservation in the four corners of the landscape, but this is trivial and the solution not unique, so it is not considered in the present analysis. The bonus levels used by Parkhurst and Shogren (2007) were

- *obb* = 16, *rbb* = 13, *cbb* = 8 for the *core*
- *obb* = 8, *rbb* = 16, *cbb* = 0 for the *corridor*
- *obb* = 19, *rbb* = 16, *cbb* = 16 for the *cross*.

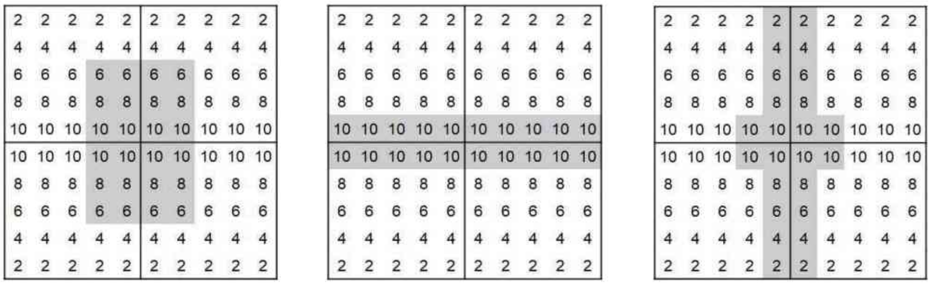

Figure 2. Three target patterns (*core*, *corridor* and *cross*: panels a–c) incentivised by Parkhurst and Shogren (2007) through appropriate settings of the agglomeration bonuses.

Each experiment went over ten rounds. In some treatments, several sets of ten rounds were coupled, so that the subjects played, e.g., ten rounds to achieve the *corridor*, followed by ten rounds for the *core* and then another ten rounds again for the *corridor*.

In each round, Parkhurst and Shogren recorded, for each player, whether these conserved the five or six targeted land parcels (Fig. 2) or not. If a player conserved "correctly", this was denoted with a score of 1; otherwise a score of zero was denoted. The maximum score over ten rounds and four players that could be achieved per experiment therefore was 10 × 4 = 40. Other, supplementary indicators recorded by the authors are not considered in the present analysis. On average (see below), the players reached a score of 14.5 (36 %) in the *core* experiments, a score of 28 (70 %) in the *corridor* experiments, and a score of 12 (30 %) in the *cross* experiments.

2.2 Modelling the experimental dynamics

The present simulations try to mimic the experiments of Parkhurst and Shogren (2007), starting from an economically used landscape and running over ten simulation rounds. Two models of the players' decision behaviour are considered (cf. the Introduction). The first is the "reinforcement learning model" by Erev and Roth (1998) that is used in many economic contexts (Charpentier et al. 2022), with an application in the context of CI by Iftekhar and Tisdell (2016). The second is the novel "future-expectations model" that acknowledges the relevance of beliefs in complex coordination games (Russell et al. 1990) and explicitly considers the model players' (agents') beliefs of how a neighbour's future land-use decision depends on the current land use. In the development of the models, a number of simplifications is introduced for numerical simplicity or due to the absence of required data.

1) Communication between model players (considered as a treatment by Parkhurst and Shogren (2007)) is not considered, so each model agent decides independently, based on their information about their own and the other players conservation costs (Fig. 1) and the land-use pattern in the entire landscape of the previous round.

2) As described above, Parkhurst and Shogren (2007) carried out several 10-round experiments in a sequence, where the players could learn from previous 10-round experiments, which affected their decision behaviour in later experiments. In particular, e.g., the dynamics observed in the *core* experiment differed by whether the experiment was the first, second or third in a set of experiments. These "long-term" dynamics are ignored in the present model, and the experimental results of Parkhurst and Shogren (2007) are averaged over the positions in the set (so the above score of 14.5 observed in the *core* experiments is an average over all core experiments reported in Parkhurst and Shogren (2007), regardless of whether the individual experiment was first, second or third in a set.

3) Given the payment levels described above, a real player's profit may be maximised by conserving five (in the *corridor* experiment) or six land parcels (in the *core* and *cross* experiments). In Parkhurst and Shogren (2007), the players were allowed to decide in each experiment how many land parcels (up to six) to conserve. For simplicity (but with negligible loss of generality), the present model agents can, in contrast, not choose whether to conserve five or six land parcels. Instead, in the *core* and the *cross* experiments (Fig. 2a and c) it is assumed that the agents always

conserve six land parcels, while in the *corridor* experiments (Fig. 2b) they always conserve five land parcels.

4) In the experiment of Parkhurst and Shogren (2007), the players could choose among all mathematically possible land-use strategies (25 × 24 × 23 × 22 × 21 = 6,375,600 different strategies for five conserved land parcels and 25 × 24 × 23 × 22 × 21 × 20 = 127,512,000 different strategies for six conserved land parcels). In the decision models below, the players explicitly attach a value to each strategy and sample a preferred strategy based on these values. Given the structure of the models, especially the future-expectations model which contains some dynamic optimisation, the consideration of so many strategies would lead to unacceptable computation times. Therefore I simplify by assuming that the agents are aware that it is always profitable to agglomerate the own conserved land parcels to collect at least the own-border bonus. Under this assumption is is plausible to consider only strategies in which each land parcel is connected to at least one other own land parcel by an east-west or a north-south border (so that a chess rook could move throughout the five- or six-parcel network). The number of these strategies turns out to be 346 when conserving five and 1,400 when conserving six land parcels (Appendix A).

5) Lastly, Parkhurst and Shogren (2007) also measured other indicators, in addition to the above introduced score, such as how often the emerged habitat network was not of the desired shape but contiguous; whether under allowed communication among players the played strategy equalled the announced one; and the "economic efficiency" defining the "percentage of available program rents earned by the group".

2.3 The reinforcement-learning model

The core of the agent-based simulation model is the decision model of the agents (players). The first variant to be considered is based on the reinforcement learning approach developed by Erev and Roth (1998) which has three parameters characterising the agent's learning behaviour: the initial value q_0 of the "propensity" q that determines how strongly the current decision depends on the associated payoff, the "recency" parameter φ that measures how strongly the present decision is influenced by the present conditions compared to the conditions further in the past, and the "experimentation" parameter ε that determines how willingly the agent explores and adopts alternative strategies.

Adopted to the present case, the reinforcement-learning model looks as follows (for a related application, see Iftekhar and Tisdell (2017)). The simulation starts in round $t = 1$ with the setting of parameter q_0, and for each agent the initial propensities of all $N = 346$ (1400) strategies a_1, …, a_N are set at

$$q_{i,n(i)}(t=0) = q_0 \ (n = 1, \ldots, N) \quad . \tag{1}$$

The probability of choosing strategy a_n is

$$p_{i,n(i)} = \frac{q_{i,n(i)}}{\sum_k q_{i,k}} \tag{2}$$

Equation (2) models some form of bounded rationality so that the strategy with the highest propensity is not chosen with certainty but only with a rather high probability. To allow for a continuous transition towards "perfect rationality", I scale the propensities q to the interval [0, 1] and raise them to some power κ:

$$q'_{i,n(i)} = \left[\frac{q_{i,n(i)} - \min_n(q_{i,n(i)})}{\max_n(q_{i,n(i)}) - \min_n(q_{i,n(i)})} \right]^{\kappa} \tag{3}$$

which is inserted in eq. (2) instead of q. By choosing the "rationality parameter" $\kappa = 1$ one obtains the original outcome of eq (2), and with increasing κ the strategy with the highest propensity receives increasingly higher weight and is chosen with an increasingly higher probability, so that for $\kappa \rightarrow \infty$ always the action with the highest propensity is chosen (Fig. 3). For each of the four agents, these probabilities form a probability distribution from which strategies $A_i(t = 1)$ (I = 1, …, 4) are drawn randomly.

For the next round, $t' = t + 1$, for each player i the payoff $R_i(t)$ associated with strategy $A_i(t)$ is calculated using the conservation costs and the payment and bonuses introduced in section 2.1. Due to the row- and column-border bonuses, the payoff R_i depends on the strategy A_i of player i as well as the strategies of the other players. So R_i is a function of all four players' strategies: $R_i(t) = R_i(A_1(t), A_2(t), A_3(t), A_4(t))$.

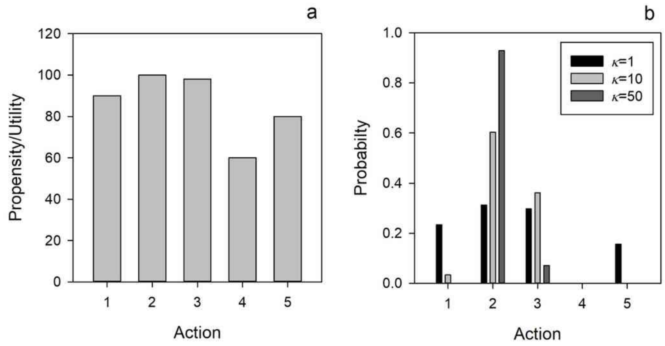

Figure 3. Propensities q (eq. 1) or utilities u (eq. 9) of five fictitious strategies (panel a) and associated probabilities (eqs. (2) and (3)) or eq. (11), respectively, for three levels of the "rationality parameter" κ (panel b).

As noted, in reinforcement learning the agents experiment with slightly varying strategies within some neighbourhood of the previous strategy. To define the neighbourhood $M_k(d)$ of size d around strategy A_k I introduce a distance, or spatial adjacency, $\delta(k, k')$ for all other strategies k' relative to strategy k. For this I consider the north-south and east-west coordinates (x_{kj}, y_{kj}) of the land parcels $j = 1, \ldots, J$ (with $J = 5$ or 6, respectively) conserved under strategy A_k, as well as the coordinates of the J land parcels conserved under the other $N - 1$ strategies, $(x_{k'j}, y_{k'j})$ ($k' \in \{1, \ldots, N\} \setminus k$). For each pair (k, k') I calculate the Euclidean distance between the land parcels within each of the J^2 pairs formed by taking one conserved land parcel from strategy A_k and the other from strategy $A_{k'}$, and take the average over these J^2 distances:

$$\delta(k,k') = \frac{1}{J^2}\left(\sum_{j=1}^{J}\sum_{j'=1}^{J}\left(x_{kj}-x_{k'j'}\right)^2+\left(y_{kj}-y_{k'j'}\right)^2\right)^{1/2} . \qquad (4)$$

The neighbourhood $M_k(d)$ then contains all actions k' that have a distance $\delta(k', k) \leq d$. The magnitude of the set $M_k(d)$ is denoted by $|M_k(d)|$. According to Erev and Roth (1998) the change in the propensity $q_{n'}$ ($n' = 1, \ldots, N$) induced by the payoff R_i from the previous round is given by

$$E_{i,n(i)} = \begin{cases} R_i(1-\varepsilon) & n'(i) = n(i) \\ R_i \cdot \varepsilon / M_{n(i)} & n(i) \neq n'(i) \in M_{n(i)} \\ 0 & n'(i) \notin M_{n(i)} \end{cases} , \qquad (5)$$

which means that the propensity of the strategy $n(i)$ chosen in the previous round is increased by the associated payoff weighted by $(1 - \varepsilon)$, while the "rest" of the payoff is distributed evenly among the propensities of the other strategies $n'(i)$ in the neighbourhood $M_{n(i)}$ around strategy $n(i)$. By this, a small ε gives a high weight to the previous strategy, while a large ε gives a high weight to the other strategies in the neighbourhood. The propensities for the next round then are

$$q_{i,n(i)}(t') = (1-\varphi)q_{i,n(i)}(t) + E_{i,n(i)}, \qquad (6)$$

where for $\varphi = 0$ the propensities of the previous round are fully considered, while for $\varphi = 1$ only the current payoffs are considered to determine the next propensities. From the new propensities of eq. (6) the probabilities $p_{i,n(i)}$ are calculated by eq. (2) to sample the strategies $A_i(t')$ for the next round, whose payoffs are calculated to update the propensities for the following round, and so on, until t = 10. The behavioural parameters q_0, d, ε, φ and κ are assumed identical for all four model agents.

2.4 The future-expectations model

The future-expectations model is somewhat complementary to the reinforcement-learning model by considering expectations about future payoffs rather than the knowledge of past payoffs. The motivation of the approach was the question, why a player would conserve one of the costly land parcels in the middle of the model landscape (Fig. 1). One answer could be that the players expect that if everybody conserved land parcels in the middle rows everybody would be better off than by conserving in the outer rows; and that each player expects the other players to have the same expectation.

This interaction of expectations however, leads to a circularity, such that player 1 may have an expectation of what player 2 will do, who has an expectation of what player 1 will do, including player 1's expectation of what player 2 will do, what might in turn be included in the expectations of player 1, and so on – which is difficult or even impossible to formalise correctly. To avoid this, I reconsider the decision problem as a temporal one and use the well-known economic concept of stochastic dynamic optimisation (Dixit and Pindyck 1994). If a player expects that the conservation of a land parcel at the boundary of their property in the current round would stimulate a neighbouring player to conserve their adjacent land parcel in the following round, s/he would altogether loose in the first round but would gain (together with the other player) in the next and possibly all following rounds. To formalise, I denote by

$u \in [0, 1]$ the "believed responsiveness", i.e. subjective belief (probability) of a model agent that if s/he conserves a land parcel at the property boundary in the current round, a neighbouring agent will conserve the adjacent land parcel in the following round.

In addition, it is plausible to assume some fidelity of the players' actions, so that if a player conserves a particular land parcel in the current round s/he may be likely to conserve it in the next round, too. To formalise, I denote by

$v \in [0, 1]$ the "believed fidelity", i.e. subjective belief (probability) of a model agent that if another agent conserves a particular land parcel sh/e will conserve the same land parcel in the following round, too.

To join the two probabilities, consider two land parcels at a property boundary where one land parcel belongs to some agent A and the other to neighbouring agent B. Denote by $x_A(t)$ and $x_B(t)$ the states of the two land parcels in round t, where $x = 1$ represents conservation and $x = 0$ represents economic use. Agent A's subjective probability of agent B conserving their land parcel in the next round $t + 1$ then is modelled as

$$\Pr\left(x_B(t+1) = 1\right) = u x_A(t) + v x_B(t) - uv x_A(t) x_B(t), \tag{7}$$

so that the belief of A that B will conserve their land parcel in the next round is highest if A and B conserve in the current round, and increases with increasing u and v. In terms of mathematical logic, eq. (7) represents an OR operator so that (for sufficiently large u and v) either $x_A(t) = 1$ or $x_B(t) = 1$ is sufficient for a high probability of obtaining $x_B(t + 1) = 1$.

To model an agent's strategy $A_i(t)$ I assume that the agents maximise some utility derived from the payoff associated with the current action and the payoff associated with the (optimised) strategy $A_i(t + 1)$ in the next round. To develop the approach, consider agent 1 in the upper left of the model landscape (Fig. 1) as the focal agent (for the other agents the arguments are analogous). The agent chooses in round t a strategy $A_1(t)$ for the current round and a strategy $A_1(t + 1)$ for the next round to maximise the inter-temporal utility

$$U_1(t) = w U_1^{(\mathrm{curr})}\left(A_1(t)\right) + (1-w) U_1^{(\mathrm{next})}\left(A_1(t+1)\right). \tag{8}$$

The inter-temporal utility is the weighted sum of the current and the next-round utilities, $U_1^{(\text{curr})}$ and $U_1^{(\text{next})}$, with weights $w \in [0, 1]$ and $1 - w$, respectively, and is maximised through stochastic dynamic optimisation (Dixit and Pindyck 1994).

The utility $U_1^{(\text{next})}$, accrues from the focal agent's strategy and associated payoff in round $t + 1$. This payoff depends on the strategies of the other agents in $t + 1$ which are uncertain to the focal agent in the current round t.

The other agent's strategies in round $t + 1$are not only uncertain to the focal agent in the current round t, but they also depend, via eq. (7) on the current land use, i.e. on the focal agent's strategy $A_1(t)$ as well as the other agents' strategies $A_2(t)$, $A_3(t)$ and $A_4(t)$. To consider the various dependencies and develop the utility for round $t + 1$, assume for the moment that the focal agent, at the time of his/her current decision $A_1(t)$, knows the other agents' strategies $A_2(t)$, $A_3(t)$ and $A_4(t)$, so that all four $A_i(t)$ are known.

From this information the focal agent builds for each of the neighbours' land parcels (note that only the land parcels adjacent to the focal landowner's property boundary are relevant) the probability (eq. 7) of being conserved in round $t + 1$.

Based on these probabilities, the focal agent samples the land-use on the neighbours' land parcels (conserved or in economic use) $l = 5$ times and for each of the obtained land-use patterns calculates the payoff (as a function of the conservation costs, base payment and the three bonuses) for each of his/her $N = 346$ (1400) possible strategies a_n introduced above. This considers the focal agent's uncertainty about the neighbour's land use. The rather small number of $l = 5$ was chosen to keep computation time acceptable but may also be regarded as considering the real players' cognitive limitations ("bounded rationality").

The associated payoffs are denoted as $R_1^{(\text{next})}(l', n)$ ($l' \in \{1, \ldots, l\}$. Assuming a risk-neutral agent (but see the Discussion), the utility of strategy a_n is calculated by the mean of the payoffs over all l':

$$u_1^{(\text{next})}(a_n) = \frac{1}{l}\sum_{l'=1}^{l} R_1^{(\text{next})}(l, n) \tag{9}$$

A perfectly rational agent would choose the a_n that maximises $u_1^{(\text{next})}$. A bounded rational agent will choose that strategy only with a high probability. To determine the probability $p_1(a_n)$ of the focal agent choosing strategy a_n, similar to eq. (3) I scale the utilities $u_1^{(\text{next})}(a_n)$ to the interval [0, 1] and raise them to some power κ:

$$u_1'(a_n) = \left[\frac{u_1(a_n) - \min_n(a_n)}{\max_n(a_n) - \min_n(a_n)}\right]^{\kappa} \tag{10}$$

to obtain the probability of choosing strategy a_n as

$$p_1(a_n) = \frac{u_1'(a_n)}{\sum_k u_1'(a_k)} . \tag{11}$$

As in the reinforcement-learning model, increasing κ models a continuous transition towards "perfect rationality" by giving the strategy with the highest utility u_1 an increasingly higher weight, so it is chosen with an increasingly higher probability. Using the probabilities of eq. (11), a strategy a_{n^*} is sampled randomly and the associated utility $u_1(a_{n^*})$ (eq. 9) is taken as the utility $U_1^{(\text{next})}$ in eq. (8):

$$U_1^{(\text{next})}\left(A_1(t+1)\right) = u_1(a_{n^*}) . \tag{12}$$

Finally, the above assumption has to be relaxed that the focal agent knows, in the current round t, the actions of the other agents in that round. But although these are not known with certainty, they can be predicted probabilistically from the known actions of the *previous* round, $A_1(t-1)$, $A_2(t-1)$, $A_3(t-1)$ and $A_4(t-1)$. From these, the focal agent can, analogous to the analysis of round $t+1$ above, via eq. (7), determine for each land parcel the probability that it will be conserved in the current round t. As above, the focal agent randomly samples $l = 5$ times which of the neighbours' land parcels are conserved in round t and which are not. For each of these $l' = 1, \ldots, l$ cases the agent calculates the current payoff $R_1^{(\text{curr})}(l', n)$ for each of its own strategies ($n = 1, \ldots, N$), as well as the next round's utility $U_1^{(\text{next})}$ – which depends, as introduced above, on the current strategies, $A_1(t)$, $A_2(t)$, $A_3(t)$ and $A_4(t)$. Similar to above, a utility for the current round t is calculated via

$$u_1^{(\text{curr})}(a_n) = \frac{1}{l}\sum_{l'=1}^{l} R_1^{(\text{curr})}(l, n) \tag{13}$$

and identified by $U_1^{(curr)}(A_1(t))$ of eq. (8). To determine the chosen strategy $A_1(t)$ of the focal agent in the current round, the utility $U_1(t)$ of eq. (8) is calculated for each of the strategies a_n ($n = 1, \ldots, N$), the obtained utilities are scaled to the interval [0, 1], raised to the power of κ. The probabilities are calculated as in eq. (11) and a strategy is sampled (i.e., following the same procedure as for round $t + 1$ described above).

This procedure is carried out for each round $t = 1, \ldots, 10$ for all four model agents, assuming identical behavioural parameters (u, v, w, κ) for all agents. For the initial round $t = 1$, it is assumed that there has been no conservation previously, so a focal agent's belief, i.e. subjective probability of the other agents conserving any land parcel, is zero.

2.5 Model analysis

Each of the three combinations of bonuses (section 2.1) is simulated for about $L = 10{,}000$ model parameter combinations, each built by randomly and independently drawing parameter values from the ranges given in Table 1. To motivate these ranges, according to preliminary analyses, the returns R_i are of the order of magnitude of 100. Due to eqs. (5) and (6), small initial propensities at the chosen lower bound of $q_0 = 1$ are thus strongly influenced by R_i, while large propensities at the chosen upper bound of $q_0 = 10{,}000$ are weakly influenced by R_i. The considered numerical range of q_0 thus represents a wide range of possible (relative) impacts of R_i on strategies' propensities.

A value of $d = 1$ represents a small spatial neighbourhood, while $d = 5$ covers almost the entire model landscape (Fig. 1). The role of κ and a motivation for its upper bound of 100 is given in Fig. 3. The other model parameters, φ, ε, u, v, and w, range within the unit interval [0, 1] by their definition.

For each simulation, in each of the ten rounds it is recorded for each agent whether it is conserving the targeted land parcels (Fig. 2), and the total number of scores S is calculated by summing over rounds and agents (so that the maximum of S is $10 \times 4 = 40$).

To determine the influences of the model parameters, S is plotted as a function of each parameter (ignoring, or essentially averaging over, the respective other parameters). In addition, the distribution of S over all L parameter combinations is determined.

For the future-expectations model (not for the reinforcement-learning model, for reasons to become apparent below) a second analysis is carried with the model parameters drawn from the restricted ranges given in Table 1, "Step 2". Here I am interested in the parameter combinations that fit the model best to the data, i.e. lead to the highest agreement between the simulated score S_{sim} and the observed score S_{obs} in the lab experiments by Parkhurst and Shogren (2007) (cf. section 2.1).

Table 1: Ranges of the model parameters (q_0 and κ are distributed logarithmically).

Model Parameter	Notation	Range	
Reinforcement-learning model			
Common logarithm of initial propensity (eq. 1)	$\lg(q_0)$	[0, 4]	
Recency parameter (eq. 6)	φ	[0, 1]	
Experimentation parameter (eq. 5)	ε	[0, 1]	
Neighbourhood size (eq. 4)	d	[0, 5]	
Common logarithm of rationality parameter (eq. 3)	$\lg(\kappa)$	[0, 2]	
Future-expectations model		Step 1	Step 2
Believed responsiveness (eq. 7)	u	[0, 1]	[0.9, 1]
Believed fidelity (eq. 7)	v	[0, 1]	[0, 1]
Weight of current utility (eq. 8)	w	[0, 1]	[0, 0.2]
Rationality parameter (eq. 10)	$\lg(\kappa)$	[1, 2]	[1.5, 2]

In particular I am plotting the distributions of model parameters, so that

1. for the *core* experiments $|S_{sim} - 14.5| \leq 0.1$
2. for the *corridor* experiments $|S_{sim} - 28| \leq 0.1$
3. for the *cross* experiments $|S_{sim} - 12| \leq 0.1$
4. ($|S_{sim} - 14.5| \leq 2$ for *core*) AND ($|S_{sim} - 28| \leq 2$ for *corridor*) AND ($|S_{sim} - 12| \leq 2$ for *cross*).

To detect pairwise interactions between the four model parameters I also determine all 3 + 2 + 1 = 6 bi-variate distributions of S.

Due to the stochasticity in the land-use decisions, S is subject to random variation. To use computation time more efficiently (still occupying a few dozens of PCs for several weeks), rather than replicating simulations of each model parameter combination and averaging the results, I determine S once for each of the L parameter combinations as described and smooth the results using a moving-window average. For this I rescale all model parameters to a range from zero to one (considering instead of κ its common logarithm lg(κ)). Each of these unit ranges is split into 20 even intervals, obtaining 21 nodes (0, 0.05, …, 1) for each parameter and a grid of 21^4 nodes y_0 that cover the 4-dimensional model parameter space.

For each node $y_0 \equiv (u_0, v_0, w_0, \lg(k_0))$ I identify all combinations $y \equiv (u, v, w_0, \lg(\kappa_0))$ of model parameters whose Euclidean distance e to y_0 is below some value e_{max}

$$e = \left((u-u_0)^2 + (v-v_0)^2 + (w-w_0)^2 + \left(\lg(\kappa) - \lg(\kappa_0)\right)^2 \right)^{1/2} \leq e_{\max}, \quad (14)$$

and take the average $\hat{S} \equiv \hat{S}(y_0)$ of the $S(y)$ of these parameter combinations.

With $e_{max} = 0.12$, eq. (14) represents the equation of a 4-dimensional sphere whose radius is about one fourth of the extension of the 4-dimensional unit cube that represents the model parameter space. The value of 0.12 was chosen as a compromise between having enough data points in the sphere for a meaningful average (so choosing e_{max} not too small) and detecting the local influence of the model parameters on S (so choosing e_{max} not too large).

3 Results

3.1 The reinforcement-learning model

With the reinforcement-learning model, simulated conservation efforts are only hardly coordinated in the entire model parameter space: For the *core* experiments, only about 0.5 percent of the L parameter combinations yield non-zero scores S, with $S = 10$ being the maximum observed; for the *corridor* experiments, 6 percent of the S are non-zero, mostly with a value of $S = 1$ and an observed maximum of 11; and for the *cross* experiments, about 0.3 percent of the S are non-zero, with a maximum 10. None of these scores comes close to the values observed in the lab experiments of Parkhurst and Shogren (2007) (cf. section 2.1). Therefore, this model appears unable to explain the observed players' behaviour and will not be evaluated any further.

3.2 The future-expectations model: performance and parameter influences

The future-expectations model, in contrast, is able to yield scores S from the full range between zero and 40 (Fig. 4). In most parameter combinations the score is zero, but a few parameter combinations in the *core* experiments and a considerable number in the *corridor* experiments lead to high scores close to the maximum of 40. In the *cross* experiments, scores up to 20 are observed. Comparison of these results indicates that the corridor is the easiest target pattern to establish by the model agents, followed by the core and, closely, the cross – which very well reflects the outcome of the lab experiments which had resulted in average scores of 28, 14.5 and 12, respectively (cf. sections 2.1 and 2.2).

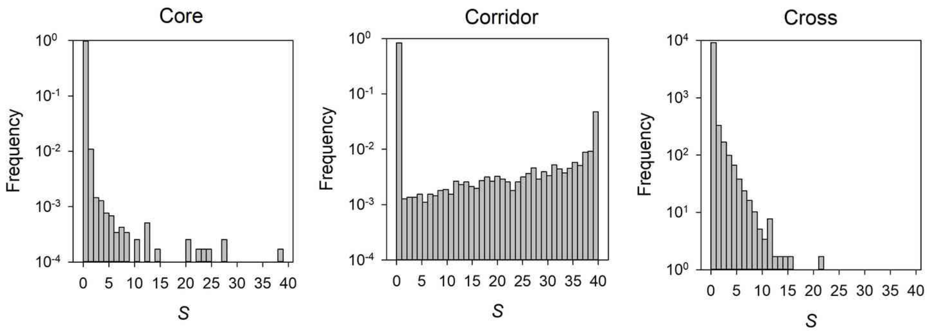

Figure 4. Frequency distributions of the score *S* for the three target patterns, *core*, *corridor* and *cross*, in logarithmic scale.

Now turn to the influence of the model parameters on the score *S*. For all three target patterns, *S* increases with increasing believed responsiveness *u*, decreasing weight of the current utility *w* and increasing rationality κ (Fig. 5). The influence of the believed fidelity *v* on *S* is weak and ambiguous. And a close look reveals that the weight *w* that maximises the score is not exactly but only close to zero (especially in the *core* experiments, Fig. 5, top row).

For all three target patterns (panels a–d) the experimental data are best fitted through a large believed responsiveness of *u* between 0.9 and 1.0, a rather large believed fidelity *v* between 0.8 and 1.0 and a low weight of current payoffs *w* between 0.0 and 0.1. For the rationality parameter κ the results are less clear, so that suitable lg(κ) are found in the full considered range between 1.5 and 2.0 (1.7 and 2.0 for the *cross* experiments) – which corresponds to κ between ca. 30 and 100; and results differ somewhat between target patterns.

If the model was fitted simultaneously to the data of all three target patterns (panels e–h) (note that here the allowed deviation between simulated and observed score was larger because otherwise no suitable model parameter combination could be found) the results are similar but clearer: values of $u \approx 0.95$, $v \approx 0.9$, $w \approx 0.025$ and $\lg(\kappa) \approx 1.7$–1.9 lead to the best fit.

There are a few pairwise interactions of the model parameters with respect to the best fit (Appendix B). In the *core* simulation experiments, decreasing *u* requires a decreasing *w* to obtain a good fit (Fig. B1, upper middle panel), so *u* and *w* are inversely “substitutable”; and in the *cross* simulation experiments, u and v and u and κ are "substitutable" so that a decrease in u can be compensated for by an increase in v and κ, respectively, to obtain a good fit.

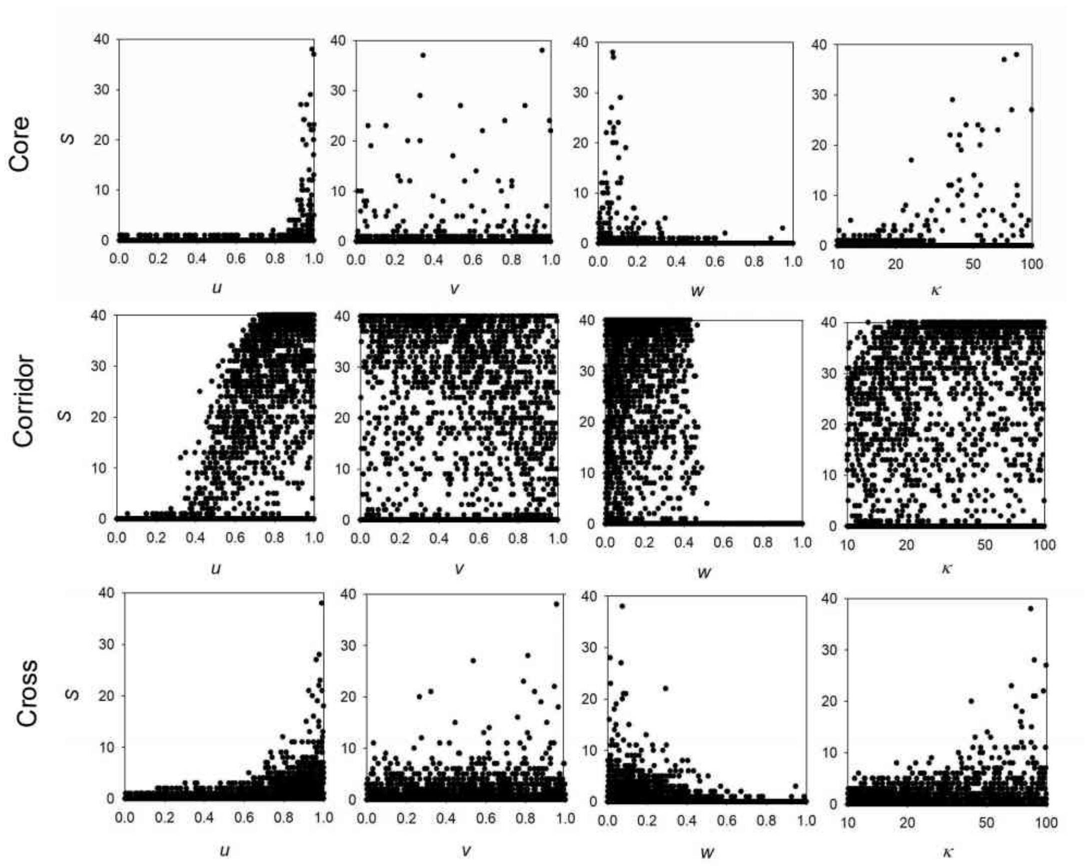

Figure 5. Scatter plot of score S versus model parameter (over all L parameter combinations) for the three target patterns, showing the influences of believed responsiveness u, believed fidelity v, weight of current utility w and rationality parameter κ.

3.3 The future-expectations model: fitting to the experimental data

Based on these observations, for the following analyses the ranges of the model parameters were restricted to the ranges that lead to relatively high scores of about $S > 10$ in all three target patterns (Table 1, "Step 2"). Again, the model was run for all three target patterns for about $L = 10{,}000$ model parameter combinations. The model parameter combinations were selected (after averaging as described in section 2.5) that led to the best fit to the experimental data (section 2.5). The results are shown in Fig. 6.

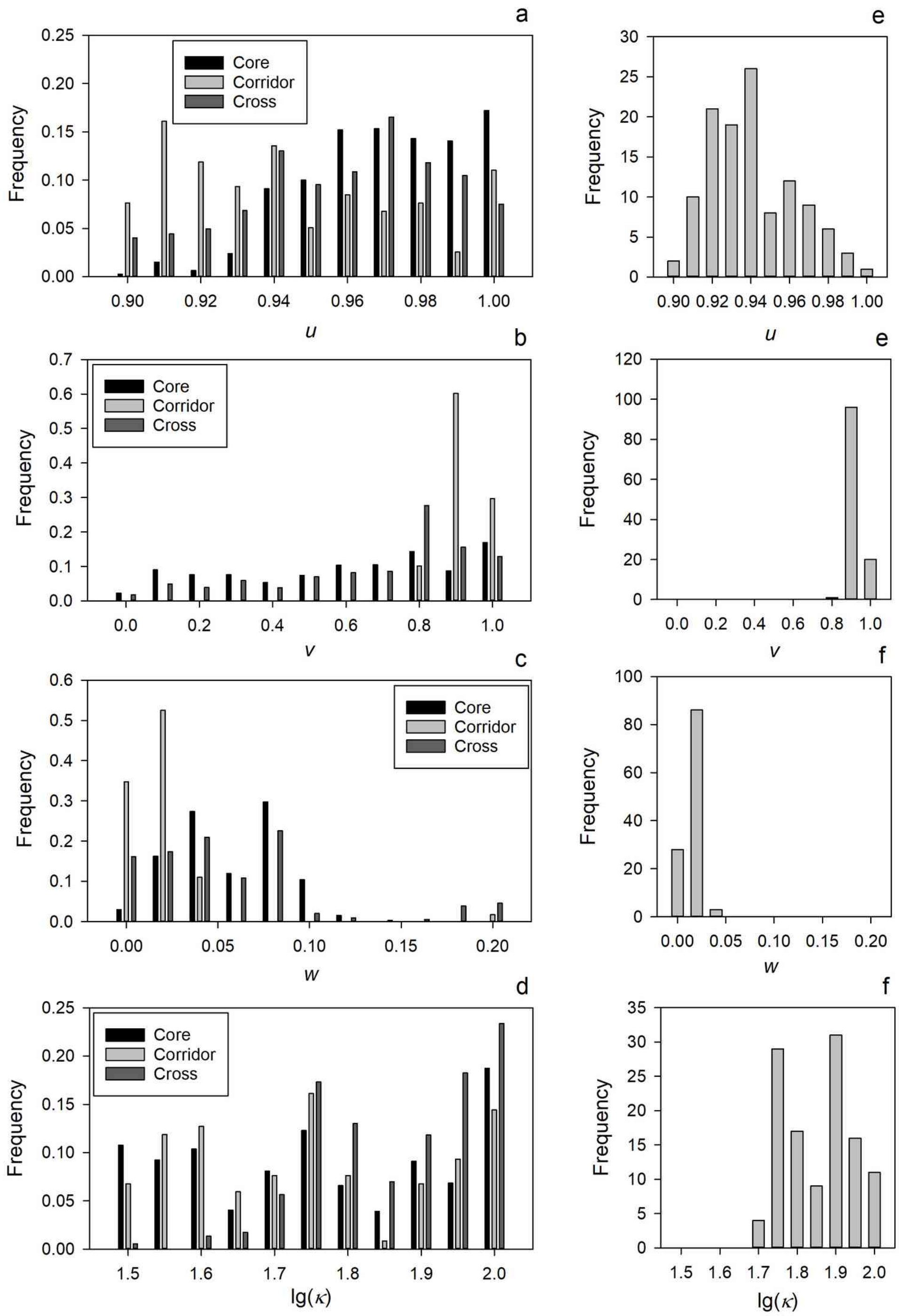

Figure 6. Distributions of model parameter values (u, v, w, κ), i.e, the proportion of model parameter combinations that include the particular value of parameter u, v, w and κ, respectively, that fulfil the constraints on the score S introduced in section 2.5. The rationality parameter κ is considered by its common logarithm. Panels a–d: each target pattern fitted individually (constraints (1)–(3)); panels e–f: simultaneous fit of all three target patterns (constraint (4)).

4 Discussion

The present paper attempts to replicate results of Parkhurst and Shogren's (2007) lab experiment on the agglomeration bonus in an agent-based simulation model. In particular, two alternative decision-making concepts are compared: learning from the past via reinforcement learning, and deriving expectations of the future through subjective beliefs and stochastic dynamic optimisation.

The first and most obvious outcome of the analysis is that model agents whose land use is based on reinforcement learning are not able to coordinate their land use and the experimental data cannot be replicated. In contrast the model based on future expectations could be fitted quite well to the data. This was not only individually for each of the three target patterns of Parkhurst and Shogren (2007), the *core*, the *corridor* and the *cross*, but also simultaneously for all three patterns. And in addition, the simulation confirmed the experimental result that the *corridor* was easiest to establish by the players, followed by the *core* and then closely by the *cross*.

For all three target patterns, the likelihood of coordination success increases with increasing believed responsiveness u (the player's belief that another player's neighbouring land parcel will be conserved in the next round if the own land parcel is conserved in the current round), increasing believed fidelity v (the player's belief that another player's neighbouring land parcel will be conserved in the next round if that land parcel is conserved in the current round), decreasing weight w of the current round's payoff (thus, increasing weight of the next round's payoff), and increasing rationality parameter κ (where for small $\kappa \approx 1$ the probability of choosing a particular land-use strategy is given by the strategy's relative payoff compared to that of the other strategies, and for large $\kappa >> 1$ the strategy with the highest payoff is chosen almost with certainty).

The experimental data were fitted best for rather high $u \approx 0.95$ and $v \approx 0.9$, low $w \approx 0.025$ and $\kappa \approx$ 50–80, which indicates rather high believed responsiveness and fidelity, a high weight on future payoffs, and a rather high probability of ("rationally") choosing the best land-use strategy.

The large fitted values for u and v indicate that spatial coordination requires expectations that the other players will coordinate as well, especially if they (v) and/or oneself (u) had chosen a coordinated strategy in the current round. This agrees well with experimental results in a two-by-two coordination game by Neumann and Vogt (2009) which stress the role of the "player's guess their partner choose the same strategy".

Given the relevance of payoff- and risk-dominance in coordination games, as they are induced by the agglomeration bonus, one may miss here the consideration of risk aversion. In fact, this was considered in preliminary model analyses by replacing the expected utilities in eqs. (9) and (13) by risk-utility functions. However, the parameter that measured the level of risk aversion only very weakly affected the simulated land-use dynamics and the players' coordination success. This, too, agrees with the experimental results of Neumann and Vogt (2009) who found the players' choices to depend more on the subjective beliefs of the other players' future decisions than their own risk attitude.

On the other hand, one could argue that in a short lab experiment the weight w of the current versus future payoffs, as introduced in the present model, does not really measure a time preference (as the formal similarity with a discount factor might suggest) but rather a risk attitude, so that a high weight on current (future) profits represents high (low) risk aversion. Altogether, the role of players' risk attitudes is an important issue for future research in coordination problems in general and coordination incentives in particular.

An interesting case is the rationality parameter κ which could not be fitted very precisely. A reason for this could be a change of its value over time in the lab experiment, so that in the initial rounds the players would be more uncertain and sometimes choose suboptimal strategies (cf. Fig. 3 in Parkhurst and Shogren (2007), modeled by smaller κ, and only in later rounds when they converged to the agglomerated target pattern would become more certain about the best strategy, modelled by an increasing κ.

Lastly, one may wonder why model agents using reinforcement learning were not able to coordinate, while in Ifthekar and Tisdell (2016) they were; and while Banerjee et al. (2012) were able to fit a time series model to the observed game dynamics that predicts current player decisions from past decisions. The reason seems to be the complexity of the present game. In both mentioned papers each player owned only a single land parcel which could be conserved or not (although in Ifthekar and Tisdell (2016) the players could then choose from several coalitions with their neighbours), so the strategy space is much smaller than that in Parkhurst and Shogren (2007). In fact, the difficulty in Parkhurst and Shogren's experiments seems to be that the risk-dominant strategy of conserving the outer rows (Fig. 1) is far away from the payoff-dominant strategy of conserving in the costly middle of the model landscape, physically separated by numerous suboptimal strategies that prevent a continuous shift from the risk-dominant to the payoff-dominant strategy.

To some extent this difficulty can be addressed by reinforcement learning that – similar to heuristic optimisation algorithms like "simulated annealing" (Kirkpatrick et al. 1983) – includes a random element (experimentation parameter ε, eq. (5)) in the decision process. This is like a hiker in foggy weather who would occasionally have to descend into a valley to have a chance of reaching the highest mountain top, rather than always walking upwards with a high risk of ending only on the hilltop nearby. In Parkhurst and Shogren (2007) the payoff-dominant, coordinated strategy could in principle be found by a sufficiently large experimental search space (parameter d in Table 1), but this implies the consideration of many suboptimal strategies, and the chance of two neighbours independently deciding for coordination is very small.

A more effective strategy to address coordination problems and strategic uncertainty is to learn not only from previous own rewards but also about the other players' behaviour. In these "belief-based models" (Camerer and Ho 1999) the agents (probabilistically) consider the strategy sets of the other players, which enlarges their own strategy sets – including not only actions that had been associated with high own past rewards (as in reinforcement learning) but also actions that would have yielded a high reward if the other player(s) had played a different strategy. However, since the agents still learn from the past, they are most likely to be unable to solve the present coordination problem, in which coordinated actions (that would induce own coordination) are simply not included in the other players' strategy sets.

In a similar manner one may relate the results of Alós-Ferrer and Weidenholzer (2006, 2008) who simulated the dynamics of multiple players interacting with each other according to a two-by-two coordination game. The model agents were able to find the coordinated, pay-off dominant state (Alós-Ferrer and Weidenholzer 2006), but this is again likely to be due to the players' small strategy spaces. Nevertheless, an interesting insight for the present context is the result of Alós-Ferrer and Weidenholzer (2008) that coordination was facilitated when coordinated states could be established first in a small neighbourhood before they would spread into the larger network. This may be relevant when coordination incentives are applied to a large number of landowners and where the question is not only, which coalitions of neighbouring conserving land owners are stable (Bareille et al. 2022), but what facilitates their establishment.

5 Conclusion

The present study proposes an alternative model for the behaviour of subjects facing a complex coordination problem with large strategy spaces. In the model-based replications of experiments with smaller strategy spaces (including those cited in the Introduction), model agents might be able to find observed coordinated states with reinforcement learning or belief-based models. However, real land use systems are generally more complex than considered in those experiments, so there seems to be a demand for models in which agents base their decisions not only on past experiences but also on expectations about the future. The developed future-expectations model is, of course, quite specific to the experiment of Parkhurst and Shogren (2007) and would require generalisation, but I hope that the ideas and arguments derived in the course of the model development and analysis well help to develop such a general model that is able to assess the performance of coordination incentives in various ecological and economic contexts.

References

Banerjee, S., de Vries, F.P., Hanley, N., van Soest, D.P., 2014. The impact of information provision on agglomeration bonus performance: An experimental study on local networks. *American Journal of Agricultural Economics* 96, 1009–1029.

Banerjee, S., Cason, T.N., de Vries, F.P., Hanley, N., 2017. Transaction costs, communication and spatial coordination in Payment for Ecosystem Services schemes. *Journal of Environmental Economics and Management* 87, 68–89.

Banerjee, S., Kwasnica, A.M., Shortle, J.S., 2012. Agglomeration bonus in small and large local networks: A laboratory examination of spatial coordination. *Ecological Economics* 84, 142–152.

Bareille, F., Zavalloni, M., Viaggi, D., 2022. Agglomeration bonus and endogenous group formation. *American Journal of Agricultural Economics*. https://doi.org/10.1111/ajae.12305.

Bell, A.R.,Parkhurst, G., Droppelmann, K., Benton, T.G., 2016. Scaling up proenvironmental agricultural practice using agglomeration payments: Proof of concept from an agent-based model. *Ecological Economics* 126, 32–41.

Bell, A.R., Ward, P. S., Mapemba, L., Nyirenda, Z., Msukwa, W., & Kenamu, E. (2018). Smart subsidies for catchment conservation in Malawi. *Scientific Data* 5, 180113.

Benito-Ostolaza, J.M., Brañas-Garz, P., Hernández, P., Sanchis-Llopis, J.A., 2015. Strategic behaviour in Schelling dynamics: Theory and experimental evidence. *Journal of Behavioral and Experimental Economics* 57, 134–147.

Camerer, C., Ho, T.-H., 1999. Experience-weighted attraction learning in normal form games. *Econometrica* 67, 827–874.

Charpentier, A., Élie, R., Remlinger, C., 2021. Reinforcement learning in economics and finance. *Computational Economics*, https://doi.org/10.1007/s10614-021-10119-4.

Cooper, R.W., DeJong, D.V., Forsythe, R., Ross, T.W., 1990.Selection criteria in coordination games: some experimental results. *The American Economic Review* 80, 218–233.

Delacote, P., Robinson, E. J., & Roussel, S. (2016). Deforestation, leakage and avoided deforestation policies: A spatial analysis. *Resource and Energy Economics* 45, 192–210.

de Vries, F.P., Hanley, N., 2016. Incentive-based policy design for pollution control and biodiversity conservation: A review. *Environmental and Resource Economics* 63, 687–702.

Dijk, J., Ansink, E., & van Soest, D. (2017). Buyouts and Agglomeration Bonuses in Wildlife Corridor Auctions. *Tinbergen Institute Discussion Paper.* DOI: 10.2139/ssrn. 2946850

Dixit, A.K., Pindyck, R.S., 1994. *Investment Under Uncertainty*. Princeton University Press.

Drechsler, M., 2021. Impacts of human behaviour in agri-environmental policies: How adequate is homo oeconomicus in the design of market-based conservation instruments? *Ecological Economics* 184, 107002.

Drechsler, M., Wätzold, F., Grimm, V., 2022. The hitchhiker's guide to generic ecological-economic modelling of land-use-based biodiversity conservation policies. *Ecological Modelling* 465, 109861.

Drechsler, M., Johst, K., Wätzold, F., Shogren, J.F., 2010. An agglomeration payment for cost-effective biodiversity conservation in spatially structured landscapes. *Resource and Energy Economics* 32, 261–75.

Erev, I., Roth, A.E., 1998. Predicting how people play games with unique, mixed strategy equilibria. *American Economic Review* 88, 848–881.

Grimm, V., Revilla, E., Berger, U., Jeltsch, F., Mooij, W.M., Railsback, S.F., Thulke, H.H., Weiner, J., Wiegand, T., DeAngelis D.L., 2005. Pattern-oriented modeling of agent-based complex systems: lessons from ecology. *Science* 310, 987–991.

Hartig, F., Calabrese, J.M., Reineking, B., Wiegand, T., Huth, A., 2011. Statistical inference for stochastic simulation models–theory and application. *Ecology Letters* 14, 816–827.

Hellwig, C., 2022. Public information, private information, and the multiplicity of equilibria in coordination games. *Journal of Economic Theory* 107, 191–222.

Heinemann, F., Nagel, R., Ockenfels, P., 2004. Measuring strategic uncertainty in coordination games. CESifo Working Paper, No. 1364, Center for Economic Studies and ifo Institute (CESifo), Munich. https://www.econstor.eu/bitstream/10419/18727/1/cesifo1_wp1364.pdf (last access Jan 22, 2022).

Huber, R., Zabel, A., Schleiffer, M., Vroege, W., Brändle, J.M., Finger, R., 2021. Conservation costs drive enrolment in agglomeration bonus scheme. *Ecological Economics* 186, 107064.

Iftekhar, Md.S., Tisdell, J.G., 2016. An agent based analysis of combinatorial bidding for spatially targeted multi-objective environmental programs. *Environmental and Resource Economics* 64, 537–58.

Kirkpatrick, S., Gelatt Jr, C.D., Vecchi, M.P., 1983. Optimization by Simulated annealing. *Science* 220 (4598), 671–680.

Krämer, J. E., & Wätzold, F., 2018. The agglomeration bonus in practice – an exploratory assessment of the Swiss network bonus. *Journal for Nature Conservation*, 43, 126-135.

Kuhfuss, L., Préget, R., Thoyer, S., de Vries, F.P., Hanley, N., 2022. Enhancing spatial coordination in payment for ecosystem services schemes with non-pecuniary preferences. *Ecological Economics* 192, 107271.

Kuhfuss, L., Préget, R., de Vries, F.P., Hanley, N., Thoyer, S., 2018. Nudging participation and spatial agglomeration in payment for environmental service schemes. Discussion Papers in

Environment and Development Economics, School of Geography and Sustainable Development, University of St Andrews. www.st-andrews.ac.uk/media/dept-of-geography-and-sustainable-development/pdf-s/DP%202017-11%20Kuhfuss%20et%20alpdf.pdf.

Neumann, T., Vogt, B., 2009. Do players' beliefs or risk attitudes determine the equilibrium selections in 2x2 coordination games? FEMM Working Paper No. 24, August 2009. Faculty of Economics and Management, Otto-von Guericke-University Magdeburg, Germany.

Nguyen, C., Latacz-Lohmann, U., Hanley, N., Schilizzi, S., Iftekhar, S., 2022. Coordination Incentives for landscape-scale environmental management: A systematic review. *Land Use Policy* 114, 105936.

Parkhurst, G.M., Shogren, J.F., 2007. Spatial incentives to coordinate contiguous habitat. *Ecological Economics* 64, 344–355.

Parkhurst, G.M., Shogren, J.F., Bastian, C., Kivi, P., Donner, J., Smith, R.B.W., 2002. Agglomeration bonus: An incentive mechanism to reunite fragmented habitat for biodiversity conservation. *Ecological Economics* 41, 305–328.

Parkhurst, G.M., Shogren, J.F., Crocker, T., 2016. Tradable set-aside requirements (TSARs): Conserving spatially dependent environmental amenities. *Environmental and Resource Economics* 63, 719–44.

Schelling,T. C.,1971. Dynamic models of segregation. *Journal of Mathematical Sociology*1(2), 143–186.

Shank, D.B., Kashima, Y., Saber, S., Gale, T., Kirley, M., 2015. Dilemma of dilemmas: how collective and individual perspectives can clarify the size dilemma in voluntary linear public goods dilemmas. *PLoS ONE* 10(3): e0120379.

Smith, A., 2015. Contribution heterogeneity and the dynamics of contributions in repeated public good games. *Journal of Behavioral and Experimental Economics* 58, 149–157.

van Huyck, J.B., Battalio, R.C., Beil, R.O., 1990. Tacit coordination games, strategic uncertainty, and coordination failure. *The American Economic Review* 80, 234–248.

Appendix A: Determination of the modelled actions

Here is is described how to determine those strategies, i.e. sets of five (for the corridor experiments) or six (for the *core* and the *cross* experiments) conserved land parcels, so that each of them shares at least one north-south or east-west border with another conserved land parcel of the same model agent.

To start with the 5-parcel strategies, I consider all possible 25 × 24 × 23 × 22 × 21 combinations of five conserved land parcels and exclude those in which there is at least one isolated conserved parcel (that has no conserved neighbours in the north, south, east or west, which is straight forward to identify). The only remaining combinations are: (i) one pair plus one triple (Fig. A1, shapes a, b), or (ii) the desired quintuple (shapes c, d, e). To distinguish cases (i) and (ii), I identify in the focal combination of those (conserved) land parcels that have *exactly one* (conserved) neighbour: in case (i) there can be four of these (shape a) or two (shape b), while in case (ii) there can be two (shape c), one (shape d) or zero (shape e). If there are four land parcels with exactly one neighbour one can thus conclude on case (i) and discard. If there are two land parcels with exactly only one neighbour, one has to distinguish between shapes b and c. The particularity that distinguishes the pair in shape b from shape c is that in the former each land parcel has as its neighbour the respective other conserved land parcel, while in shape c the land parcels with one neighbour are not adjacent to each other. This difference is straight forward to test numerically, which allows excluding shape b. After all combinations of types b and c have been excluded, a number of 346 combinations remain which are considered as the strategies in the main analyses.

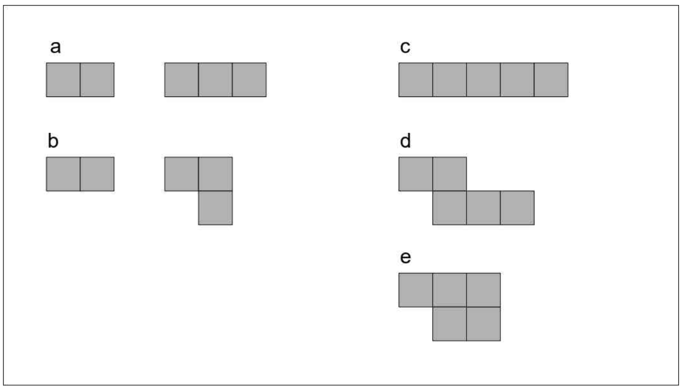

Figure A1. Typical shapes (the orientations may vary) of combinations of five land parcels which all have in common that each land parcel has at least one neighbour. For the roles of the different shapes, see the text.

Now turn to the 6-parcel strategies, and consider all possible 25 × 24 × 23 × 22 × 21 × 20 combinations from which again those are excluded that contain at least one isolated conserved parcel. The remaining possibilities are (Fig. A2): three pairs (shape a), a pair and a quadruple (shapes b–e), two triples (shapes f and g) and the desired sextuple (shapes h–l). Inspection reveals that all shapes except for the sextuple have at maximum five *inner* borders (bold lines in shape e) whose identification is straight forward by identifying all pairs of adjacent land parcels (not to be confused with the pairs in shapes a–e); while the sextuples have at least five inner borders (shapes h, i and k). The only shape that has five inner borders but is not a sextuple is shape e which consist of a pair and a block. Thus, to distinguish this shape from the sextuples with five inner borders, I test whether the shape contains a pair – which is identified as in the 5-parcel combinations above. Altogether, excluding all shapes with less than ten inner borders as well as shape e (and its variants) I obtain 1400 sextuples.

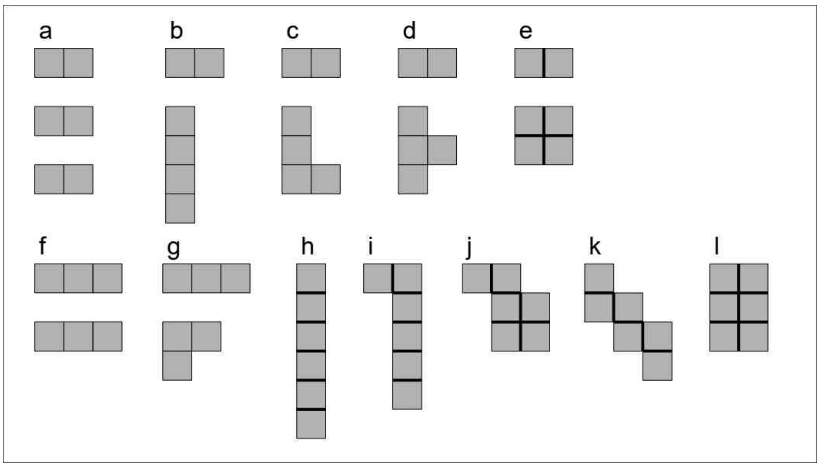

Figure A2. Typical shapes (the orientations may differ in other possible shapes) of combinations of six land parcels which all have in common that each land parcel has at least one neighbour. For the roles of the different shapes, see the text. The bold lines mark inner borders between adjacent land parcels.

Appendix B: Interactions of the model parameters in the future-expectations model

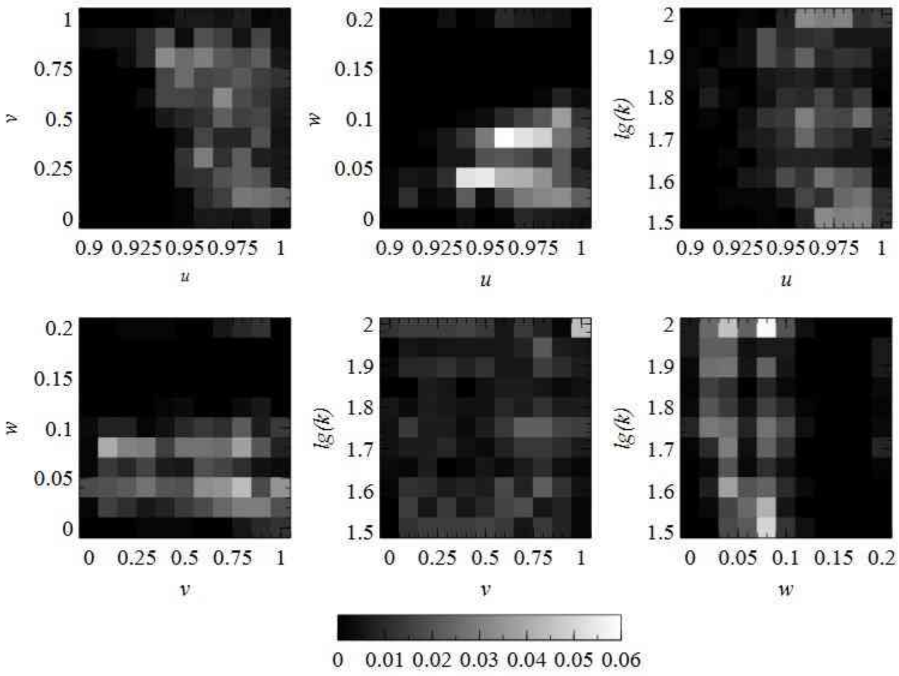

Figure B1. Two-dimensional distributions of model parameter values (u, v, w, $\lg(\kappa)$) that fulfil constraint (1) on the score S introduced in section 2.5, i.e. fit the model for the "core" experiments.

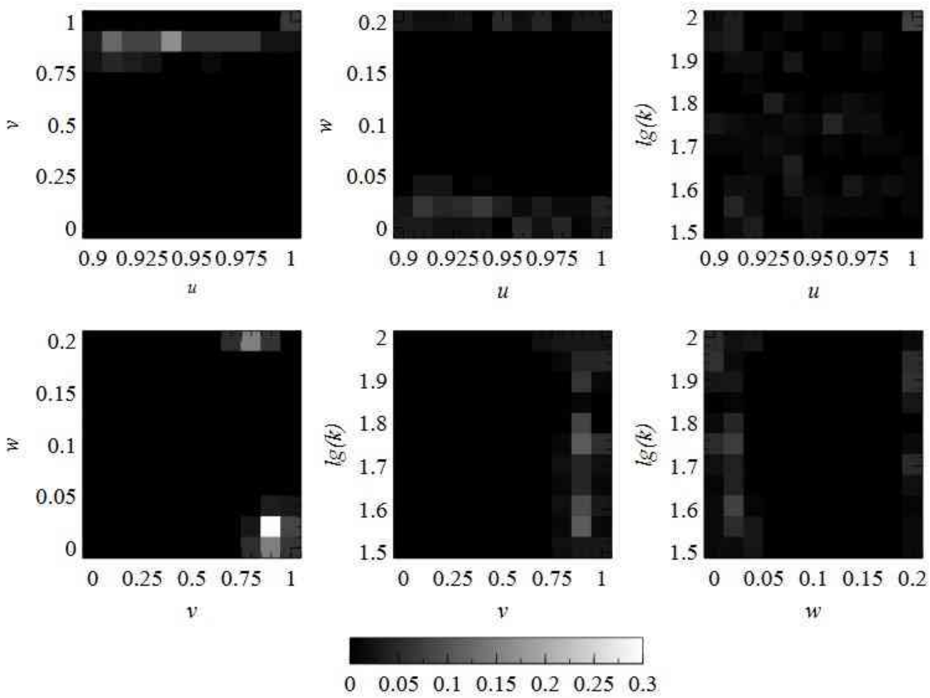

Figure B2. Two-dimensional distributions of model parameter values (u, v, w, $\lg(\kappa)$) that fulfil constraint (2) on the score S introduced in section 2.5, i.e. fit the model for the "corridor" experiments.

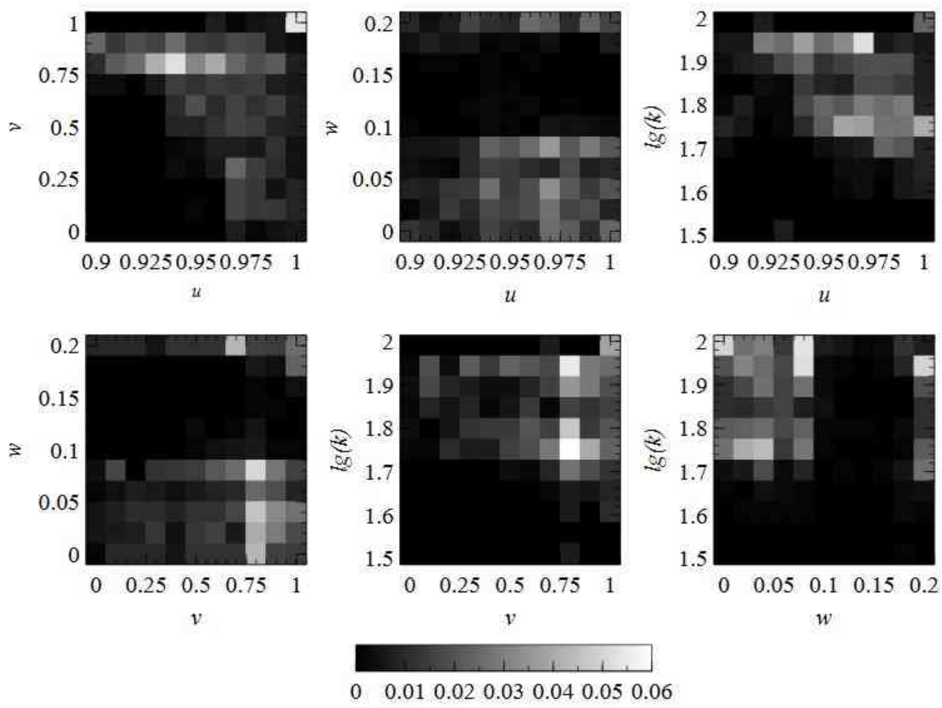

Figure B3. Two-dimensional distributions of model parameter values (u, v, w, $\lg(\kappa)$) that fulfil constraint (3) on the score S introduced in section 2.5, i.e. fit the model for the "cross" experiments.

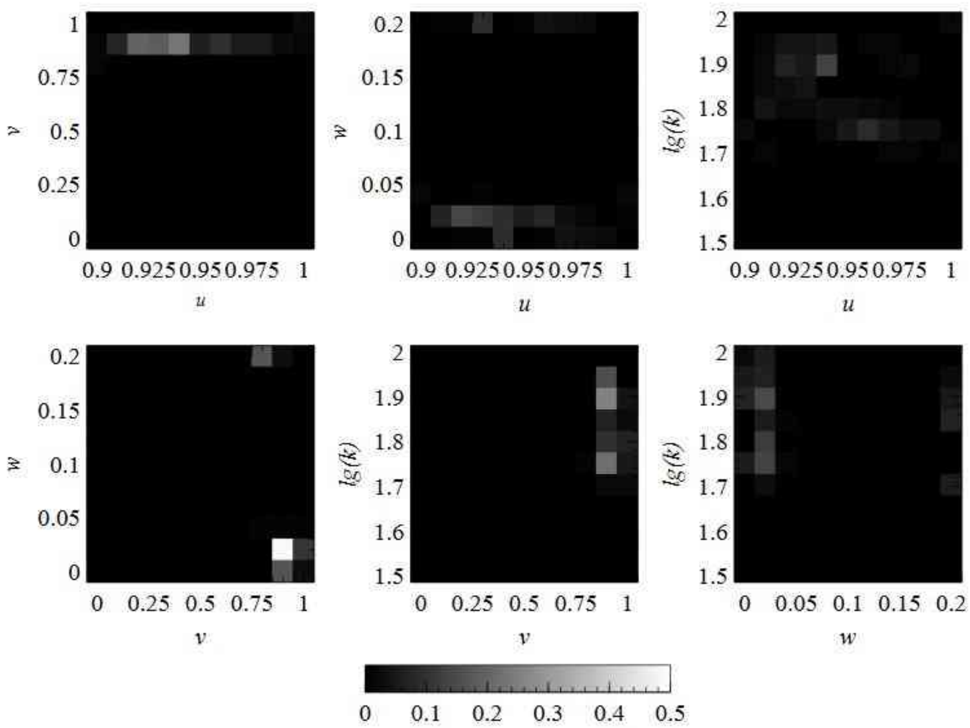

Figure B4. Two-dimensional distributions of model parameter values (u, v, w, $\lg(\kappa)$) that fulfil constraint (4) on the score S introduced in section 2.5, i.e. fit the model for all target patterns simultaneously.