Environmental and Ecological Statistics (2019) 26:47–86
https://doi.org/10.1007/s10651-019-00419-2

Modelling relationships between socioeconomy, landscape and water flows in Mediterranean agroecosystems: a case study in Adra catchment (Spain) using Bayesian networks

Rosa F. Ropero[1] · Rafael Rumí[2] · Pedro A. Aguilera[1]

Received: 11 July 2017 / Revised: 6 February 2019 / Published online: 5 March 2019

Abstract

In Mediterranean areas, the co-evolution between social and natural systems has given rise to heterogeneous and complex systems of interactions called agroecosystems, in which strong relationships between socioeconomy, landscape and water flows have been identified. In this context, water resources management is a prominent area of research, particularly in semi-arid conditions, where a special set of challenges requires novel tools to deal with uncertainty, multiple sources of information and expert knowledge. In this paper, Bayesian Networks are proposed as a means to model the relationships between socioeconomy, landscape and water flows in a Mediterranean agroecosystem, studying its behaviour under two scenarios of change in land use trends: maintenance of traditional Mediterranean agriculture, and agricultural intensification through the development of greenhouses. Results show that an increase in the area of traditional agriculture would lead to better control of runoff and increased primary productivity, measured as green water flows. By contrast, agricultural intensification of the territory would provoke an increase in evaporation and water losses. Due to the versatility of Bayesian networks, results can be expressed not only as probabilities, but also using other metrics that can be computed from them. Accordingly, *Sensitivity Analysis to Evidence*, *Sensitivity Analysis to Parameters* and the *Kullback–Leibler divergence* were carried out. Bayesian Networks have demonstrated their ability to deal with uncertainty inherent to natural systems, combining expert knowledge, data from regional datasets and Geographical Information Systems, and automatic training algorithms giving robust and proper results.

Keywords Bayesian networks · Green and blue water · Kullback–Leibler divergence · Landscape change trends · Mediterranean agroecosystems · Sensitivity analysis

---

Handling Editor: Pierre Dutilleul.

Extended author information available on the last page of the article

1 Introduction

In Mediterranean areas, the co-evolution between society and natural systems that has developed over the course of time has led to a complex patchwork of traditional agriculture, natural areas and human infrastructures co-existing in a so-called agroecosystem (Sal and García 2007). In such a system there is a strong relationship between the socioeconomy and the structure of a territory which affects water flows and the provision of ecosystems services (Gordon et al. 2010; Rockstroem et al. 2010). Under semi-arid conditions this relationships is even more important; this is the case in southern Spain, where water supply is key and has become a prominent area of research. Thus, in the field of water management planning and policy, researchers are addressing several of the challenges (Casadei et al. 2016; Kersebaum et al. 2016; Phan et al. 2016; Teegavarapu 2010), (i) the uncertainty inherent to these (socio)natural systems, (ii) the existence of limited or incomplete data and (iii) the high number of variables. In addition, planning needs to include knowledge from experts and stakeholders, but their inclusion in the model training processes is usually difficult. These challenges are even more evident under the current framework of Global Environmental Change (GEC) (Hui et al. 2015).

One of the main impacts of GEC is related to the relationship between landscape and social and economic systems (Lambin and Meyfroidt 2010; Grau et al. 2003). There is wide recognition in the literature that socioeconomic changes impact on landscape structures and functions (Caillault et al. 2013; Rudel et al. 2009; Aranzabal et al. 2008; Foley et al. 2005; Schmitz et al. 2005) and have a direct influence on water systems (Maes et al. 2009; Scanlon et al. 2005). At the end of the 1990s, two new concepts were proposed to introduce the whole water cycle into water management plans and policies (Rockstroem 2000; Falkenmark 1997): the so-called *Green* and *Blue water* flows. *Blue water* is the amount of rainfall that exceeds the soil's storage capacity and flows into rivers, lakes and aquifers whilst *Green water* refers to the rainfall that infiltrates into the root zone of the soil to support the primary productivity of natural and agricultural systems through evapotranspiration (Falkenmark and Folke 2002). Both flow through natural subsystems across the landscape, participating in several ecological processes of energy and material transport (Willaarts et al. 2012). The characteristics of soil and the type and cover of vegetation determine the amount of water that evaporates back to the atmosphere, infiltrates into the soil or flows away as runoff.

Modelling agroecosystems is therefore becoming more complex (Bonneau et al. 2016; Irvine and Gitelman 2011), as the inclusion of this new concepts of *Green* and *Blue water* flows requires new tools that are capable of managing this complexity and uncertainty, and of providing a common framework to include information from different sources.

In the mid-1980s, Bayesian networks (BNs), was proposed for reasoning with uncertainty in knowledge-based systems. They were introduced in environmental and ecological modeling with some initials approximations at the end of the 1990s (Varis and Kuikka 1997), but it was not until the beginning of 2000s when they started to be considered as an appropriate tool in this field (Falk et al. 2015; Aguilera et al. 2011; Gitelman and Herlihy 2007; Pal et al. 2001). Focusing on water management, Phan et al. (2016) reviewed their application into this field taking into account the geographic distribution, data sources, software, model validation, climate change impacts, decision-making process and whether BNs had been integrated with other modeling tools or not. Their results demonstrated that BNs can be applied to a wide range of problems related to water source. They also confirmed that experts and stakeholders are involved in a high percentage of the reviewed studies, as compared to model training and validation being done directly from the data or by comparing them against other traditional models. In this way, BNs can be integrated with other models, *e.g.* with GIS techniques, though this approximation is still scarcely applied. All these points have been encouraging researchers to apply BNs in the context of Integrated Water Resource Management (Castelletti and Soncini-Sessa 2007a, b; Henriksen et al. 2007). This has led to the application of BNs in some European projects such as the FP5-MERIT (Bromley et al. 2005) or the NeWater (Henriksen and Barlebo 2008). However, as far as we know, applications of BNs in agroecosystems modelling are still scarce (Frayer et al. 2014; Baynes et al. 2011; Sadoddin et al. 2005; Joshi et al. 2001).

The aim of the study was to model the relationships between socioeconomic structure, landscape and water flows in a Mediterranean agroecosystem. A semi-arid catchment was selected and a BN model was trained using data from three different sources of information (*Andalusian Environmental Information Network,BalanceMED model* and *Andalusian Multiterritorial Information System*). The objective is also focused on show how BNs can be applied in the field of water resource modelling. Under the framework of GEC and taking literature into account, two main landscape change trends (maintenance of traditional croplands and agricultural intensification through greenhouses) were included as scenarios of future change, and their impact over the distribution of water flows were evaluated.

2 Materials and methods

2.1 Study area

The Adra catchment lies in a semi-arid part of south-eastern Spain (Fig. 1). It is located in the provinces of Almería and Granada, bounded to the north by the *Sierra Nevada* mountain range, to the south by the Mediterranean Sea, to the east by the *Sierra de Gádor*, and to the west by the *Sierra Filabres* mountain ranges. Extending over 74.400 ha, it supports an estimated population of 12.400 people over fourteen municipalities.

This catchment supports various agricultural land uses, including agroecosystems, with four areas being differentiated: (i) landscape in *Sierra Nevada* is characterized by patches of the original Mediterranean forest with oak, and scrubland resulting from several episodes of deforestation (García-Latorre and Sánchez-Picón 2001); (ii) in the *Sierra de Gádor* foothills, land uses comprise traditional croplands including olive and almond groves with patches of woodland and scrub, creating a complex and heterogeneous landscape; (iii) in the middle and west of the area a mixture of scrub and patches of woodland is found, configured by 19th century mining and deforestation of the natural forest (García-Latorre and Sánchez-Picón 2001); and, (iv) in the lower reaches, intensive agriculture with greenhouses is the main land use.

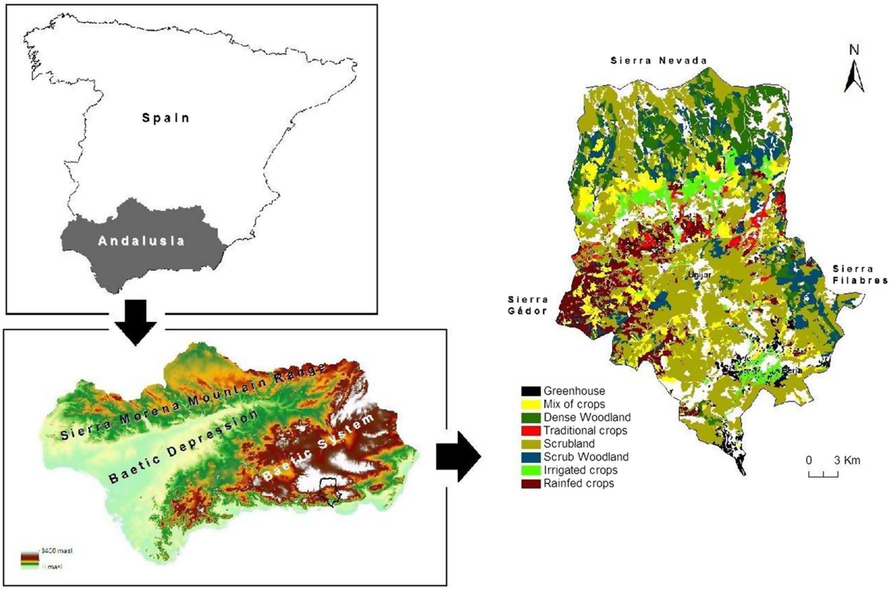

Fig. 1 Adra catchment is located in south-eastern Andalusia, in the south of Spain. Its landscape is configured by a mosaic of different land uses

The socioeconomic structure is related to this landscape pattern, with a gradual change from the upland mountain areas, with depopulated and ageing municipalities to the lower elevations where agricultural activity implies an important immigration rate and economic activity.

2.2 Data collection and pre-processing

The dataset comprises three groups of variables. Initially, information about the landscape was collected from the *Andalusian Environmental Information Network* as shapefile information and pre-processed with ArcGIS 9.3. A total of 79 land uses were initially collected, but using cartographic criteria,[1] this number was reduced. The total surface area of each land use in the study area was calculated, and those occupying less than 0.5% of the total were eliminated. This pre-processing resulted in just 17 different land uses being identified for the main study. The study area is quite heterogeneous, which gives a mosaic of different, interrelated land uses. Thus, each observation corresponds to a patch of land with an unique land use. In the data set, each row (observation) represents the size (in hectares) occupied by this land use, whilst the 16 remaining land use variables are equal to zero. This implies that the final dataset contains 67% values equal to zero (Table 1).

For each patch (observation) the *BalanceMED model* (Willaarts et al. 2012; Willaarts 2009) was applied to calculate *Green* and *Blue water* flows. It is a semi-deterministic model developed for quantifying the hydrological functioning of a

[1] Available in http://www.juntadeandalucia.es/medioambiente/site/rediam.

Table 1 Variables collected, their mean values, standard deviation (SD) and the percentage of zero values (% Zeros)

Variable	Mean	SD	% Zeros
Ageing (rate)	28.4	5.41	0.00
Emigration rate (%)	4.7	2.71	0.00
Immigration rate (%)	4.2	1.73	0.00
Scrubland (Ha)	3.84	21.68	77.02
Woodland and sparse scrubland (Ha)	10.92	1.36	90.58
Woodland and dense scrubland (Ha)	0.77	10.01	96.62
Dense woodland (Ha)	1.82	11.75	87.85
Riverbed Vegetation (Ha)	0.02	0.40	99.27
Abandonment woody crops (Ha)	0.25	8.99	97.93
Traditional woody crops (Ha)	0.30	3.18	93.53
Sparse vegetation (Ha)	0.35	5.79	95.92
Rainfed crops (Ha)	1.14	13.24	87.73
Grazing land (Ha)	0.65	9.13	95.00
Riverbed areas (Ha)	0.075	1.07	98.51
Heterogeneous crops (Ha)	0.74	5.48	92.93
Wood and grazing land (Ha)	0.21	3.48	98.83
Irrigated crops (Ha)	0.30	4.03	95.38
Greenhouses (Ha)	0.002	0.05	99.53
Human infrastructures (Ha)	0.097	0.66	93.95
Water areas (Ha)	0.006	0.15	99.31
Non-productive green water (mm)	82.8	64.98	6.70
Productive green water (mm)	187.8	78.71	8.44
Consumptive blue water (mm)	10.68	51.48	94.30
Runoff blue water (mm)	137.8	137.5	14.68

Mediterranean catchment using long time series of monthly rainfall and potential evapotranspiration data. The model assumes that a fraction of the total precipitation is intercepted by vegetation or soil and evaporates directly as Non-Productive Green Water (NPGW). Another fraction of the total precipitation can be intercepted on impermeable surfaces and then returned to the atmosphere as Consumptive Blue Water (CBW). The remaining precipitation reaches the soil and is taken up by plants and transpired, this portion is called Productive Green Water flow (PGW). When the infiltrated water exceeds the soil storage capacity, it can either percolate or drain as Runoff Blue Water (RBW). This model was used just for data collection purposes, so it was considered as a black box and no information about data uncertainty is available.

The socioeconomic variables included in the BN model take into account information from literature (Schmitz et al. 2005) and expert knowledge. Accordingly, only three variables were considered, which summarized the main socioeconomic behaviour of the catchment: percentage of people older than 65 years old, called the *Ageing* variable, and both emigration and immigration rates, calculated as the percentage of the total population that migrate out of or into the study area, respectively. The *Andalusian Multiterritorial Information System*[2] was the source of this information. The catchment contains a total of 14 municipalities; however, those covering less than 1% of the surface area were eliminated. The three were calculated for the remaining municipalities. Again ArcGIS 9.3 was used to merge this municipal information with the natural landscape data. Each observation (patch) was added into one municipality and merged with the corresponding municipal values for the three socioeconomic variables. However, some observations belong to more than one municipality; in these cases, a weighted mean was calculated from the percentage falling into each of the municipalities included in the patch.

The final dataset contains 17 land use, 3 socioeconomic and 4 water flow variables over 8017 observations.

Due to the high percentage of zero values (Table 1), all variables were discretized using the Equal Frequency method with three intervals. The discretization selected is devoted to partially solve the sparcity problem, in such a way that the first level of the variables includes the zeros, and the other two levels model the variables. Some other discretization methods were used, such Equal width and K-means, however they were discarded because the results obtained were worst in terms of scenarios modeling. Anyway, as it is shown in Ropero et al. (2018) in some cases discretizing the data does not imply worst results.

2.3 Model description

Bayesian networks are a powerful statistical tool that allows information from different sources, including expert knowledge, to be included. They can be used to study the conditional relationships between the components of the model. They are defined (Jensen and Nielsen 2007), specifically for discrete multinomial variables as: "a statistical multivariate model for a set of variables $\boldsymbol{X} = \{X_1, \ldots, X_n\}$, which is defined in terms of two components: (i) a direct acyclic graph in which each vertex represents one of the variables, linked by an edge which indicates the existence of statistical dependence between them configuring the qualitative part, and (ii) the quantitative part as the conditional probability distribution for each variable X_i, $i = 1, \ldots, n$, given its parents (pa) in the graph ($pa(x_i)$) expressed in Conditional Probability Tables (CPTs)".

Thus, they are based on graph theory and probability theory, and composed by a visual part representing the structure of interaction between the variables, and a set of CPTs that represent the strength of these relationships.

One of the main advantages of BNs is related to this qualitative component which allows to be easily understood by experts in environmental and water sciences experts who may be unfamiliar with the model's mathematical context (Kelly et al. 2013; Aguilera et al. 2011; Uusitalo 2007). In this way, BNs play an important part in the model training step by identifying relationships between the variables, giving values for the CPTs or even refining the structure previously trained from data (Aguilera et al. 2011). The structure of the network also means that, with no mathematical calculation involved, the variable(s) that are relevant (or not) for a certain one can be known (Pearl

[2] http://www.juntadeandalucia.es/institutodeestadisticaycartografia/sima/index2.htm.

1988). This allows us to simplify the joint probability distribution (PDF) of the variables necessary to specify the model. Thus, BNs provide a compact representation of the (PDF) over all the variables, defined as the product of the conditional distributions attached to each node, so that

$$p(x_1, \ldots, x_n) = \prod_{i=1}^{n} p(x_i \mid pa(x_i)). \tag{1}$$

Furthermore, since relationships between variables are expressed using conditional probability values, the output returned by this model is more expressive than other noon probailisitcs/stochastics models outputs (Uusitalo 2007).

In several applications, once the model is trained and validated, new information is received and needs to be included into the model with the aim of testing an scenario of change, such as the effect of management decisions (Stafford et al. 2016) or climate change impacts (Mantyka-Pringle et al. 2014). BNs allow this new information, or *evidence*, to be included into one or more variables, through the so-called *inference process* or *probabilistic propagation*, updating the CPTs of the remaining variables. If we denote the set of *evidenced* variables as **E**, and its value as e, then the inference process consists of calculating the posterior distribution $p(x_i|\mathbf{e})$, for each variable of interest $X_i \notin \mathbf{E}$:

$$p(x_i|\mathbf{e}) = \frac{p(x_i, \mathbf{e})}{p(\mathbf{e})} \propto p(x_i, \mathbf{e}), \tag{2}$$

since $p(\mathbf{e})$ is constant for all $X_i \notin \mathbf{E}$. So, this process can be carried out computing and normalizing the marginal probabilities $p(x_i, \mathbf{e})$, by this way:

$$p(x_i, \mathbf{e}) = \sum_{\mathbf{x} \notin \{x_i, \mathbf{e}\}} p_e(x_1, \ldots, x_n), \tag{3}$$

where $p_e(x_1, \ldots, x_n)$ is the probability function obtained from replacing in $p(x_1, \ldots, x_n)$ the evidenced variables **E** by their values **e**.

This method is highly inefficient due to the large number of value combinations it needs, and because of the storage problem it creates in the case of large networks. If the conditional independence structure represented in the graph is exploited, the number of computations needed in the inference process can be drastically reduced because of the decrease of parameters to estimate in the conditional probability tables. The use of BNs reduces drastically the number of parameters to estimate when modelling the joint probability distribution, due to the factorization of the joint distribution (Eq 1). As an example see Fig. 2. Assume that the domain of A = 2, B = 3, C = 2, D = 4 and E = 4; the joint probability table would have 2*3*2*4*4 = 192 entries, which means to estimate 191 parameters. However, by the factorization given by Eq. 1: P(A,B,C,D,E) = P(A)*P(B|A)*P(C|A)*P(D|B,C)*P(E|C), which reduces the number of parameters to be estimated to $1 + 4 + 2 + 18 + 6 = 31$. Recall that the number of parameter to estimate in each table is the number of entries -1. In this simple example the reduction in the size is drastic, even though the number of variables is small. In bigger problems,

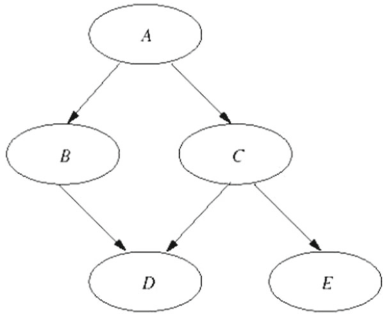

Fig. 2 Example of a Bayesian network with five variables

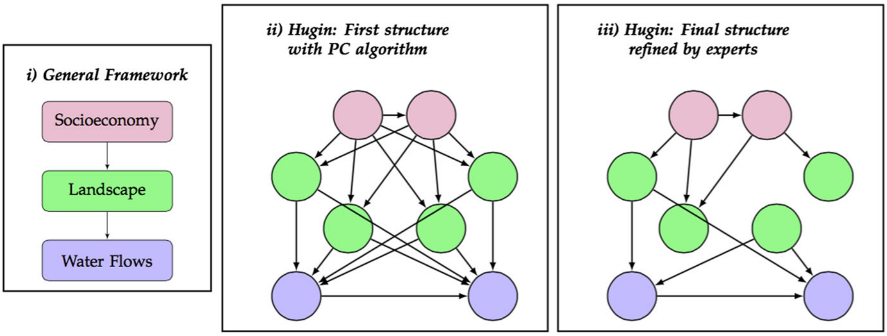

Fig. 3 Outline of the model training step

the joint probability table is easily unfeasible to compute without taking into account the dependencies of the different variables given in the network. This differences also have an impact when computing the marginal distribution of a given variable.

Based on this idea, several algorithms to compute the probabilities both exactly and approximately have been proposed (Fung and Chang 1990; Lauritzen and Spiegelhalter 1988; Pearl 1988). Since our network is not extremely large the Join Tree Algorithm was used.

2.4 Model training

Model was trained using Hugin software, and, more specifically, the *learning wizard* toolbox (Andersen et al. 1990) with the PC algorithm (Spirtes et al. 1993). It is a set of intuitive steps to perform the model training process from the data and, if it is necessary, with the inclusion of expert knowledge. This process can be summarized in Figs. 3 and 4. Once data were included into the *learning wizard* toolbox, all variables were displayed and a set of topological rules can be included (*i.e.* variable X should or should not be linked with variable Y). In our case, based on literature and expert knowledge (described in the introduction), general rules were proposed. Accordingly, variables are classed into three types: (i) socioeconomic variables that summarize the behavior of the social component, (ii) landscape represented by land use variables and, (iii) water flows. There is a wide recognition of the relationship between socio-economy

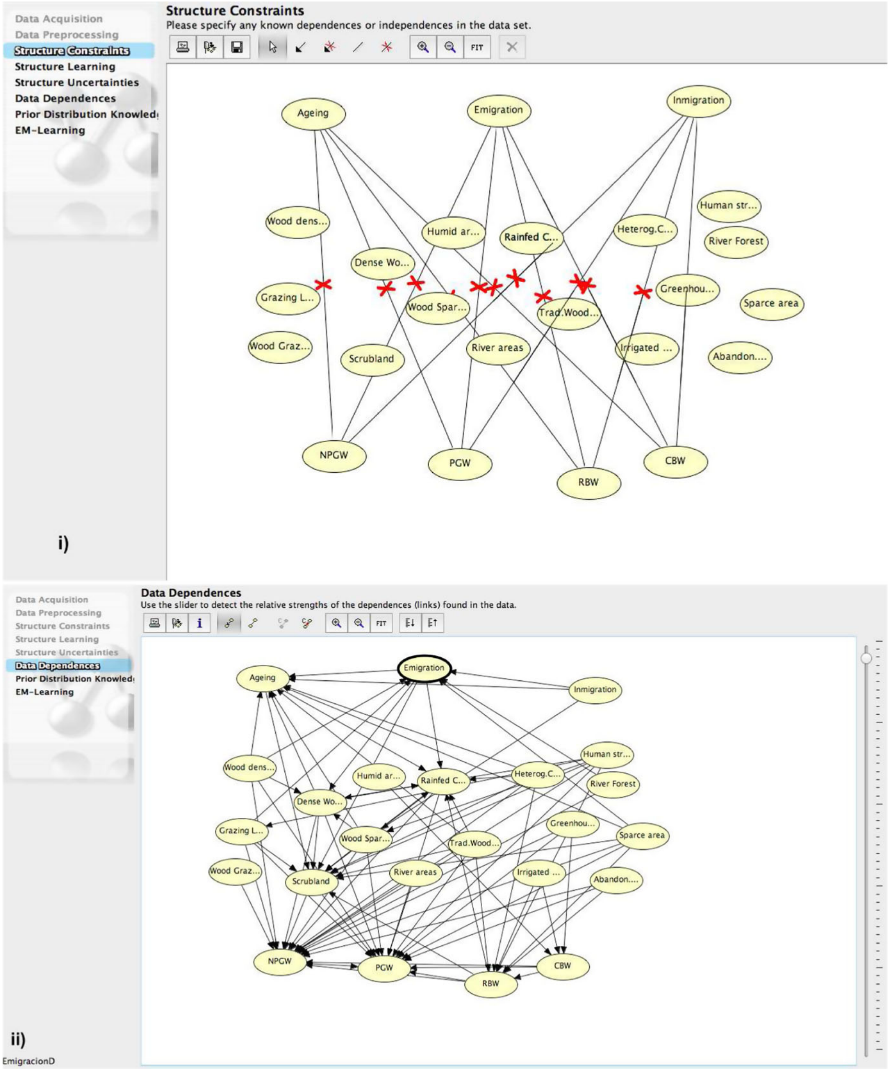

Fig. 4 Screenshot of Hugin *learning wizard* toolbox that shows the model training steps

and landscape and the scientific literature shows that socioeconomic changes affect the structure of the landscape, and changes in the structure of the landscape would affect water flows (Aranzabal et al. 2008; Schmitz et al. 2005). Therefore, it is well documented that there are no direct links between socioeconomic variables and water flows. So, the network should show links from socioeconomy to landscape and from landscape to water flows, but no direct links between socioeconomic variables and water flows Figs. 3i and 4i.

The second step consists of an initial model training. Hugin allows to estimate both the structure and parameters of the model using the PC algorithm (Spirtes et al. 1993).

It is based on χ^2—conditional independence tests for discovering the (in)dependence relationships between the variables, using the cross entropy statistic measured in the sample:

$$G^2 = 2 \sum_{i \ in \ cells} O_i ln \left\{ \frac{O_i}{E_i} \right\}, \tag{4}$$

where *cells* represents all the combinations of the categories of all the variables involved in the test, O stands for the *observed values* and E for the *expected values*. Note that, in the case of conditional independence, the expected values are computed dividing by the restricted-to-condition sample size, and in the case of simple (0-level conditional) independence, dividing by the total sample size.

This statistic is theoretically distributed as a χ^2 distribution, with degrees of freedom, when testing the conditional independence of X_i and X_j conditional on $\mathbf{W}$:

$$df = (Cat(X_i) - 1) \times (Cat(X_j) - 1) \times \prod_{i=1}^{n} Cat(W_i), \tag{5}$$

where $W_i \in \mathbf{W}$ and $Cat(X)$ returns the number of categories of variable X. However, each zero-observed cells decreases df in one.

When the sample size available to carry out the test is less than ten times the number of cells, the procedure assumes that the variables are conditionally dependent, and the link is not removed (Spirtes et al. 1993).

The PC algorithm begins with a undirected complete graph, in which all variables are assumed to be dependent on each other; then it is reduced by removing first, those links joining nodes 0-level conditionally independent (marginally independent); then, those links joining nodes verifying a 1-level conditionally independence, and so on. Given a pair of variables X_i, X_j, the set of variables to condition on is extracted from the set of adjacent nodes to both X_i, X_j. Through these tests, some variables were considered totally independent and no link between them and the rest of variables was added, so these variables were removed. It is important to note that, in the PC algorithm, only the direction of certain links is set according to the independencies found in the network. The direction of the rest of the nodes is set randomly, avoiding the inclusion of cycles in the graph. Some other independence tests can be used (not in the Hugin software) in the PC algorithm, such as the proposed by Zhang et al. (2012), or even Fisher exact independence test, however they are not implemented in Hugin and they would probably output marginal differences in the model (only some links would change, those which don not show a clear dependence) which would be minimized afterwards in the refining step by the experts.

So, once the structure was obtained according to the independence test carried out, it is displayed by the *learning wizard* toolbox (Fig. 4ii) in order to decide if it will be the final model structure or not. It gives the option to change the degrees of freedom and show the changes in the network of relationships, in such a way that experts can decide the more adequate level for each problem. Besides, since the PC determines the direction of some links randomly, in this step modeler and experts can fix this direction.

So that, before parameter estimation, expert can partially modified the structure of the network, but according to the independence structure previously estimated. When all these changes have been made, the parameters were finally estimated. from the data. Note that the structural learning and the parameter estimation process are carried out using the same training data; this is a standard procedure in Bayesian networks, and no overfitting is produced, since, even though they are not independent, the structure and the parameters of the network are considered two components of the network. The parameters are estimated through the corresponding frequencies, which in this case correspond to the maximum likelihood estimators. The form of the likelihood is

$$\prod_{i} \prod_{j} P(x_{ji} | pa(x_j)_i) \tag{6}$$

with $i = 1$ to n, $j = 1$ to k (i iterates over the cases, j iterates over the variables in the network).

In the upper formula, x_{ji} is the value of variable j in case i, $pa(x_j)$ are the parents of variable X_j in the graph and $pa(x_j)_i$ are the values of the parents of variable X_j in the graph in case i.

The parameters estimations are then computed by

$$P(X_j = x_{jk} | pa(x_j) = v) = \frac{n(x_{jk}, pa(x_j) = v)}{n(pa(x_j) = v)} \tag{7}$$

where $P(X_j = x_{jk})$ represents the probability that variable X_j takes the value x_{jk}, $pa(x_j) = v$ represents that the parent variables of x_j takes the values v, $n(x_{jk}, pa(x_j)$ represents the number of cases in the database in which variable $X_j = x_{jk}$ and $pa(x_j) = v$ simultaneously, and $n(pa(x_j) = v)$ represents the number of cases in the database in which $pa(x_j) = v$.

Note that, even when literature and expert knowledge were included into the structural model training stage, the network of relationships were partially trained from the data, whilst parameters were directly estimated from them.

2.5 Landscape change scenario

One of the main advantages of BNs is their ability to perform a prediction once new information is received. This process provides an interesting tool for water managers since information in terms of probability about the change in some variables given the value of others can be observed and evaluated.

In this paper, the aim is to evaluate the impact of landscape change on both socioeconomic structure and generation of water flows. Thus, evidences are included into land uses variables by setting the maximum probability into one of the three possibles states. It means that, *for example* if the evidence is set in the State 2, (Appendix A, Table 3) of the variable *Rainfed crops* then rather than find scattered small patches of rainfed crops, these observations would occupy a larger area and provoke changes in the most probable extension of other land uses as well.

Agricultural intensification is one of the main trends of landscape change in Mediterranean agroecosystems. The opposite trend is the maintenance of traditional agricultural techniques. In this paper, two simple scenarios of land use planning were proposed:

1. Maintenance of traditional croplands. The upper catchment of both *Sierra Nevada* and *Filabres*, but mainly *Sierra de Gádor*, belies its cultural heritage with extensive agriculture of olive, grapes and almonds, mixed with herbaceous crops and small patches of scrub. Nowadays, this traditional activity is being abandoned and substituted by intensive agriculture. This first scenario involves the protection and promotion of traditional areas. Evidences are introduced into the following variables: *Traditional Woody Crops* and *Heterogeneous Crops* are set to have maximum probability in State 2, while *Abandonment Woody crops* variable is set in State 0. At the same time, State 1 of the variable *Rainfed Crops* is set to maximum probability. This means that small and scattered patches of heterogeneous and traditional crops change to larger, more continuous areas, with the disappearance of abandoned croplands (State 0 includes all zero values) and rainfed areas present in moderate extent.
2. Agricultural intensification through greenhouses. The lower reaches of the catchment include a significant proportion of agricultural intensification with greenhouses and some areas of irrigated crops. This second scenario assumes this intensification to extend to the foothills of the three mountain ranges. Evidence was introduced into the *Greenhouse* and *Abandonment Woody crops* variables, which have the maximum probability in the highest interval, whilst *Traditional Woody crops*, *Rainfed Crops* and *Heterogeneous Crops* variables are set in the State 0. Finally, *Irrigated crops* is set to the maximum probability in State 1.

2.6 Model validation

Once evidences are included in the model and the remaining variables updated, the *Kullback–Leibler divergence* was calculated to evaluate the difference between the a priori and a posteriori situations for water flow variables. This measure is defined as a non-symmetric measure of the difference between two probability distributions over one variable. Let $p(x)$ and $q(x)$ are two probability distributions of a discrete random variable x, the *Kullback–Leibler divergence* is defined as:

$$D_{KL}(p(x)||q(x)) = \sum p(x) ln \frac{p(x)}{q(x)} \tag{8}$$

This metric gives us information about the global impact of both scenarios on the water flow variables, our variables of interest, in terms of divergence between the resulting distributions. If there is a low-impact on the interest variables, the KL-divergence will be small. On the contrary, if the impact is high and meaningful, there should be a great difference between the a priori and a posteriori distributions, leading to a higher KL-divergence. Since we will use a exact inference algorithm implemented in Hugin, we can always obtain the exact a posterior probability distributions, and then compute the corresponding DKL exactly.

An alternative way to validate results obtained by the BN when performing the probability propagation is a Sensitivity Analysis to Evidence (Jensen and Nielsen 2007), which, given some hypothesis h over the variables of interest, provides detailed information about which piece of evidence is really significant for h. Since the scenarios proposed in the paper are defined by the instantiation of several variables of the networks, it is interesting to devise which of the evidences (or combination of them) that compose the scenario are the most important. This is carried out by computing in a systematic way the so-called *normalized likelihood* for every possible subset $e^{'} \subseteq e$, given by

$$\frac{P(h|e^{'})}{P(h)} \tag{9}$$

as a way to measure the change in probability for h given different evidence combinations. In the case of binary variables, just one value of Eq. (9) is provided; in the case of variables with more than two possible states, a *normalized likelihood* value is provided for each state.

A different question can be posed, which is to decide which evidence in a given scenario supports a given hypothesis h_1 against a *contrary* h_2. In a similar way as in Eq. (9), the ratio of the *normalized likelihoods* can be computed for each subset $e^{'} \subseteq e$, and then decide which evidences are good discriminators between both hypothesis.

$$\frac{P(e^{'}|h_1)}{P(e^{'}|h_2)} \tag{10}$$

A third approach to validate the model is using so-called Sensitivity to Parameters, which give local measures of how robust the model is to *small* changes in the parameters, *e.g.* probability values in the CPTs, of the model. Since the number of possible combinations make the problem intractable, it is solved in a local way for each CPT-entry, and a goal variable to observe. Given an evidence e, a goal variable state h, and a parameter s of a BN, $P(h|e)$ is stablished as a function of s, which assuming *proportional scaling* (Jensen and Nielsen 2007), has the form

$$\frac{P(h,e)}{P(e)} = \frac{\alpha s + \beta}{as + b} \tag{11}$$

and so, the variation of the h-value according to variations in s can easily be tracked.

α, β, a and b values are determined by entering two different values for s and propagating, which yields a linear equation system whose solutions are α, β, a and b.

3 Results and discussion

3.1 Model structure and sensitivity to parameters

Before the model training step, a correlation matrix was calculated from the original continuous data (Table 12), to identify linear relationships between the variables. It was computed just as a pre-processing step in order to check that the relations observed in the matrix were incorporated in the model through the automatic training process. Only socioeconomic and water variables yielded coefficients above |0.4|, the remaining giving coefficients of less than |0.1|.

However, the correlation matrix provides only information about linear relationships, which is not the only type of relationship between variables. In this case, model training using the PC algorithm allows (in)dependent relationships to be discovered which, also, serves a process of variable selection. During training process, a set of land use variables were completely independent with respect to the rest of variables, and no links were added between them. As a results, six land use variables were eliminated, so that the definitive model was composed by 3 socioeconomic, 11 land use and 4 water flow variables.

Taking advantage of the versatility of BNs, a prior expert knowledge was used to train the structure of the model, considering general topological restrictions, by following the ecological structure of the socioecosystems. By this means, direct relationships between socioeconomic variables and water flows were avoided. However, the structure and parameters were directly estimated from the data.

Figure 5 shows the structure of the model with three levels of variables. In the first level, social variables are related to each other and to land uses, *i.e. Emigration* acts

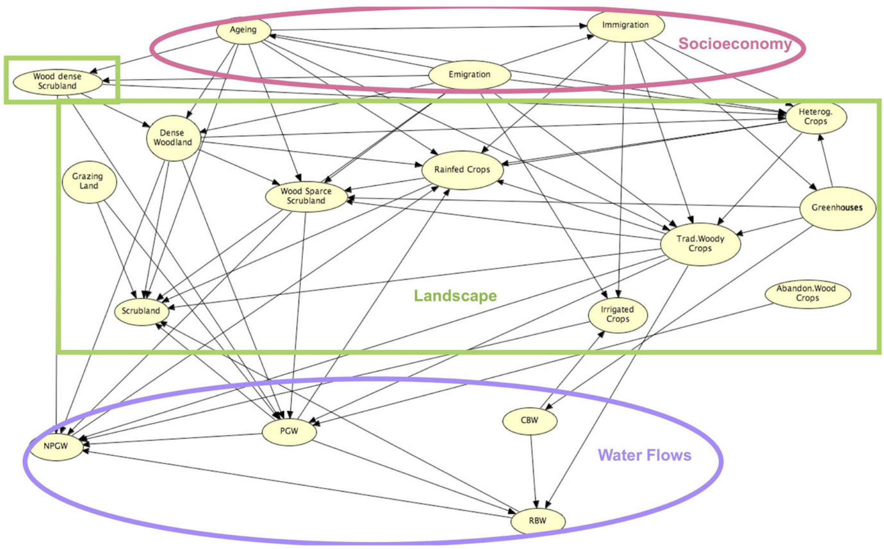

Fig. 5 Structure of the Bayesian network trained from the data using the PC algorithm, implemented in HUGIN, and refined by experts

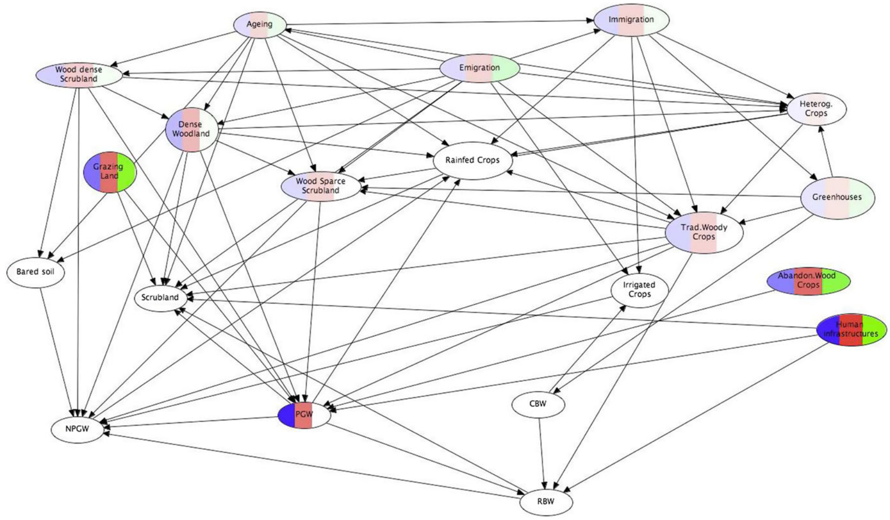

Fig. 6 *Sensitivity Analysis to Parameters* for PGW variable. Blue color represents the maximum sensitivity value, Red color the minimum sensitivity value and Green color, the average sensitivity value. When the node is color in white, sensitivity values are equal to 0

as a parent of both *Ageing* and *Immigration*, *Scrubland* is linked to both *Ageing* and *Emigration*, and *Immigration* variable is directly linked to *Greenhouse*. In a second level, land uses are related with the four water flow variables. Here, there is a difference, since blue water variables are linked to *Irrigated crops* and *Greenhouses*, land uses are related to agriculture intensification, whilst green water are linked to natural vegetation and extensive crops.

These relationships are even clearer from the Sensitivity to Parameters analysis. This measures how sensitive the result of a propagation is against changes to the parameters in other variables of the network. The analysis for each of our variables of interest are shown in Figs. 6, 7, 8 and 9. Figures 6 and 7 show results for both kinds of green water. These two variables are related to almost all land use and social variables, whilst blue water (Figs. 8 and 9) is only related to a small subset of land uses and the *Immigration* variable.

Each node is divided into three and shown in a different color. Each variable may affect in a different way depending in the change in its parameters, so for each variable there is a maximum, minimum and average impact on the inference process. If a variable is dis-connected (via the d-separation rules) to the goal variable, then it will be painted completely in white. The greatest the maximum and average values with respect to the other variables, the brightest the color (the closest to zero the minimum value, the lighter the color, being white if it is zero). This way we can easily observe the most influential variables to the goal variable (Andersen et al. 1990). This value is computed as the derivative of the so-called sensitivity function, which represents $P(A = a \mid E)$ as a function of the (perturbation) values of other variables, see Van Deer Gag and Renooij (2001) for more information.

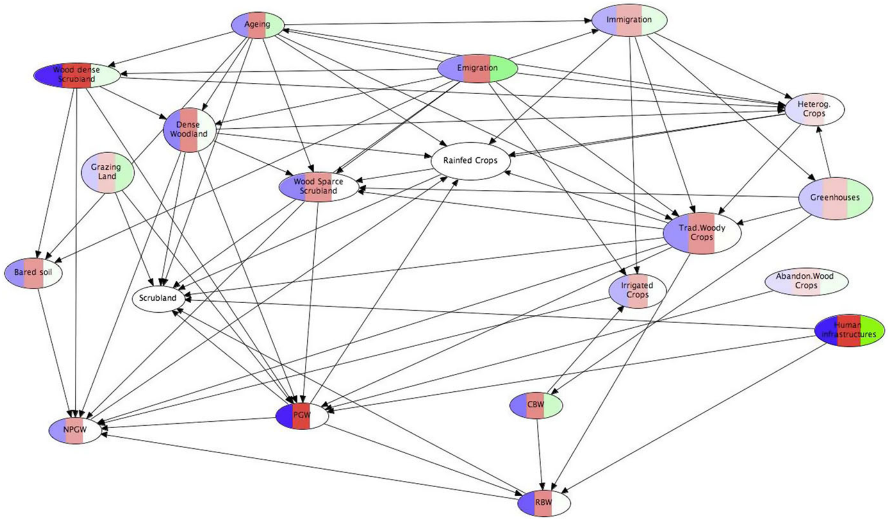

Fig. 7 *Sensitivity Analysis to Parameters* for NPGW variable. Blue color represents the maximum sensitivity value, Red color the minimum sensitivity value and Green color, the average sensitivity value. When the node is color in white, sensitivity values are equal to 0

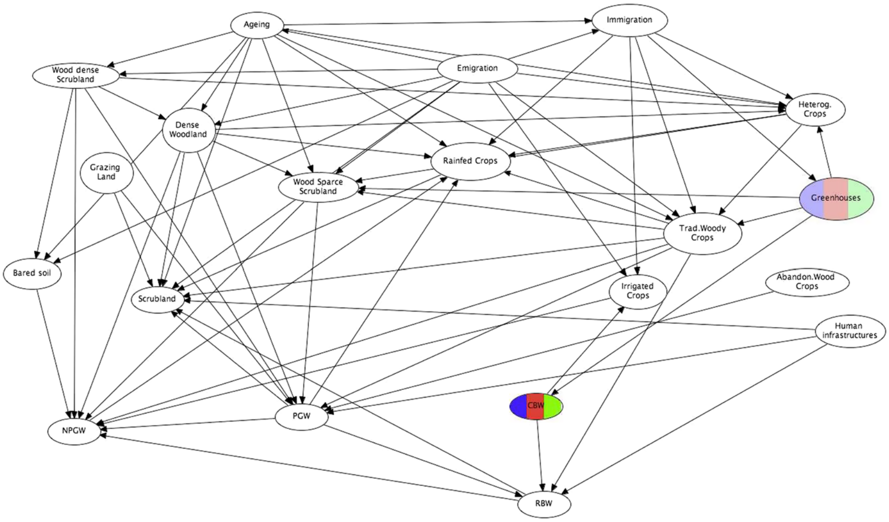

Fig. 8 *Sensitivity Analysis to Parameters* for CBW variable. Blue color represents the maximum sensitivity value, Red color the minimum sensitivity value and Green color, the average sensitivity value. When the node is color in white, sensitivity values are equal to 0

This metric gives us information about how robust the model is. In general, for all water flow variables, values of parameter sensitivity are not so high, with the majority of variables painted as white nodes, in the case of blue water (Figs. 8 and 9), or with really clear colors (Figs. 6 and 7) for green water. This means our model is quite robust

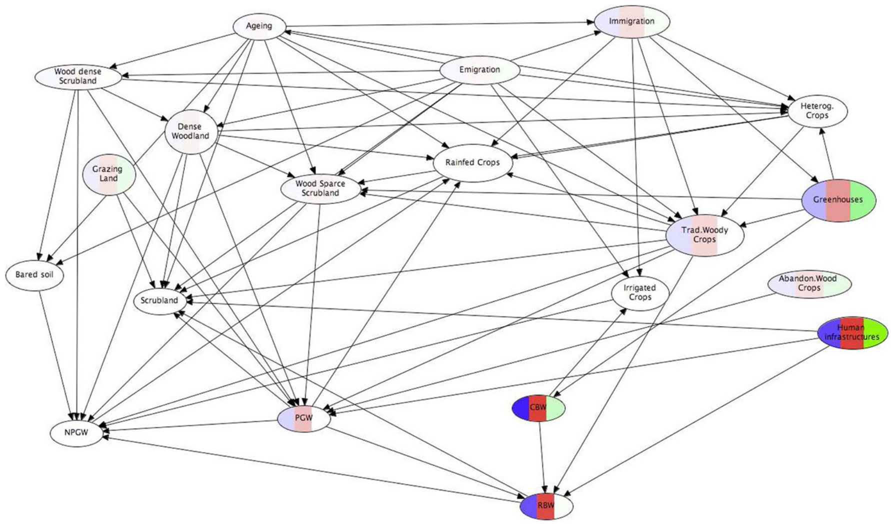

Fig. 9 *Sensitivity Analysis to Parameters* for RBW variable. Blue color represents the maximum sensitivity value, Red color the minimum sensitivity value and Green color, the average sensitivity value. When the node is color in white, sensitivity values are equal to 0

against various changes, so that, if small changes are included into the parameters, the model will provide similar results.

3.2 Scenarios of change and sensitivity to evidence

The relationships between variables in the model are clearly observed for both scenarios of change. Figure 10 shows the values for the *Kullback–Leibler divergence* which gives information about the change to variables once the evidence was included. Under the scenario of Maintenance of Traditional croplands, *Emigration* and *Immigration* took values below 0.01, whilst *Ageing* value was 0.09. By contrast, for the scenario of Agricultural Intensification through Greenhouses, the *Immigration* variable gave the lowest value in comparison to social variables.

In terms of probabilities, in the socioeconomic variables, the first scenario assumes an increase in the probabilities of the State 0 for *Immigration*, whilst *Emigration* remains similar. *Ageing* variable shows higher probabilities in the State 1 compared to the a priori situation (Fig. 11). However, the scenario of agricultural intensification leads to a higher probability for all social variables in State 2 (Fig. 11).

An increase of intensive agriculture with greenhouses assumes a larger workforce is needed and that new people are welcome into the area; *Immigration* takes higher values of probability (García-Álvarez-Coque 2002). At the same time, since jobs are taken up by outsiders, local people emigrate looking for a career unrelated to the agricultural sector. In both scenarios, the *Ageing* variable takes a higher value of probability. This tendency is more evident in the first scenario (Maintenance of traditional agriculture). This is related to the fact that this type of landscape structure is usually maintained by

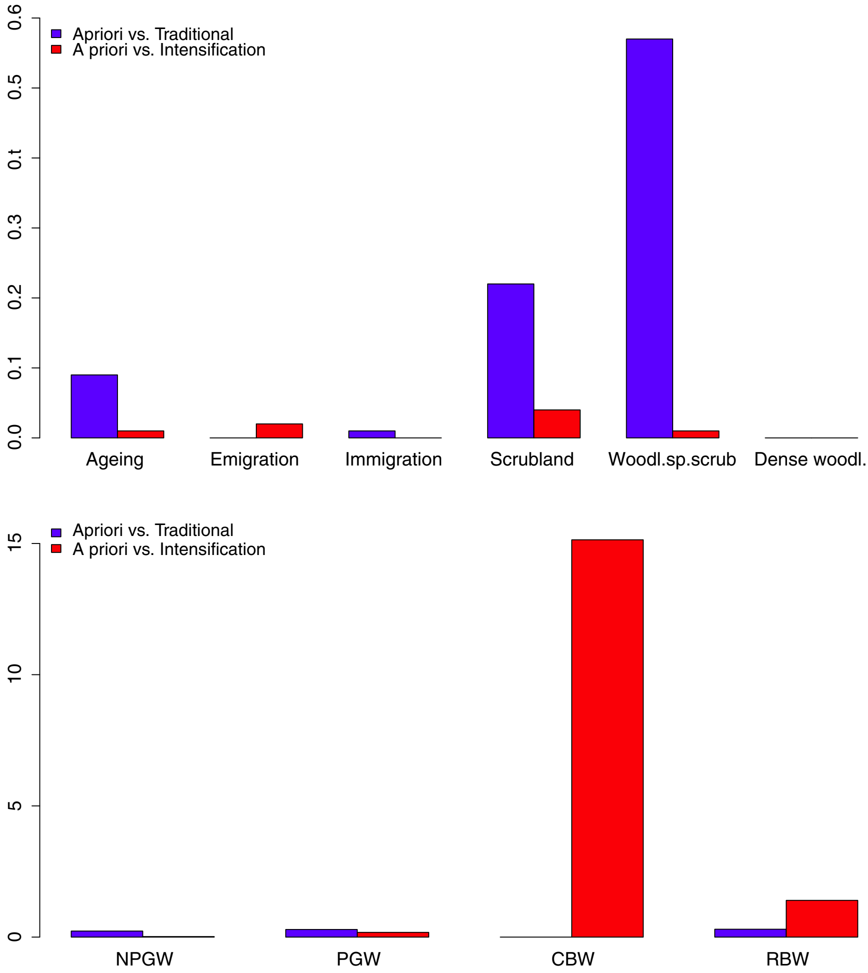

Fig. 10 Kullback–Leibler values for the social and water flow variables, and a selection of land use variables in the comparison between a priori and both scenarios. NPGW, Non-Productive Green Water; PGW, Productive Green Water; CBW, Consumptive Blue Water; RBW, Runoff Blue Water; Woodl.sp.scrub, Woodland and sparse scrubland; woodl., Woodland

an old population through the maintenance of a subsistence agriculture. At the same time, this landscape often attracts the interest of older people from North Europe: the warmer weather and the lower cost of living encourages them to settle in places like the Adra catchment (Fig. 11).

In the second level, relationships between landscape and water flows were established. Figure 12 shows the changes under the first scenario: the maintenance of traditional croplands. Evidences were included into the variables *Traditional Woody Crops*, *Abandonment Woody Crops*, *Rainfed Crops* and *Heterogeneous Crops*. In the case of *Kullback–Leibler* values (Fig. 10), land uses related to scrubland take values of 0.22 (*Scrubland*) and 0.57 (*Woodland and sparse and scrubland*). For the water flow variables, values are close to 0.20 except for *Consumptive Blue Water*, which is

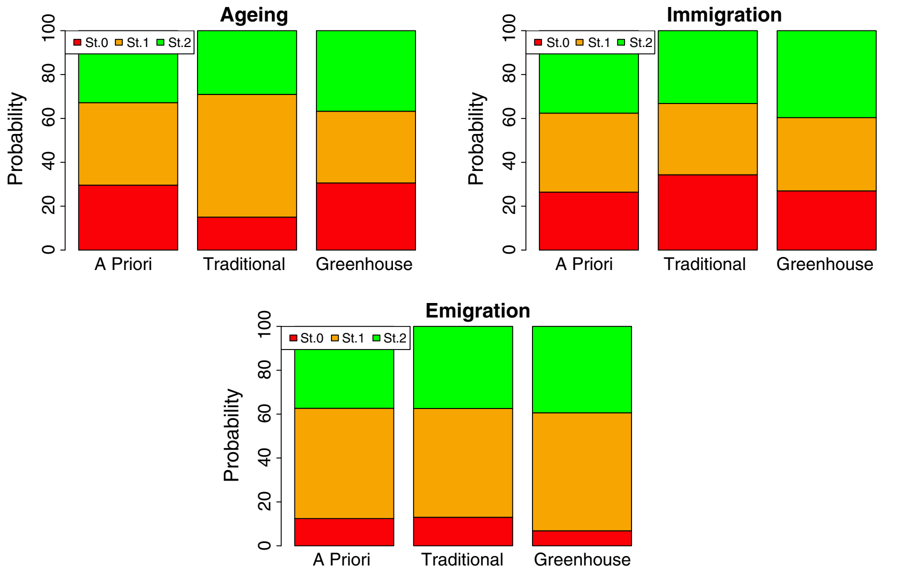

Fig. 11 Values of the probabilities of the socioeconomic variables both A priori and under the scenarios of changes: maintenance of traditional croplands (Traditional) and Agricultural Intensification through greenhouses (Greenhouse) . St. state

less than 0.001. Tables 4, 5, 6 and 7 (Appendix A) show the *normalized likelihood* for all water flow variables under this first scenario of change, while Table 2 shows the mean value of the ratio of *normalized likelihood*.

The Maintenance of Traditional Croplands scenario will lead to an increase of several agricultural areas, mainly those related with herbaceous crops (*Heterogeneous crops*), in which vegetation are usually separated by tracts of bare soil from which water could directly evaporate provoking an increase in *Non-Productive Green water* which becomes more probable in State 2 (Fig. 13). Also, the increase of vegetation cover causes State 1 of *Productive Green Water* to become more probable (Fig. 13). Values for the *normalized likelihood* are in accordance with the change in the probabilities, with higher values for State 2 of NPGW, and State 1 of PGW. However, there are some differences depending on the combination of variables evidenced. The highest values are found when all, or even just three, variables are evidenced. However, if only *Heterogeneous crops* with *Rainfed* or *Traditional crops* are evidenced, the impact shows that PGW would become more probable in State 2, and State 0 for the NPGW.

CBW is similar to a priori. The Sensitivity Analysis to Evidence shows *normalized likelihood* for CBW close to 1, which means no significant difference is found between a priori and a posteriori distributions.

Special mentions is required for the relationship between *Traditional Woody Crops* and RBW. Under the scenario of maintenance of the traditional agriculture, RBW shows a great increase of the probability of State 0 (Fig. 13). This change is mirrored by the *Kullback–Leibler* value (Fig. 10) and the *normalized likelihood* (Table 5). These values are distant from 1, both above and below, and always higher for State 0 compared to the other states for all possible combinations of evidenced variables. However, if

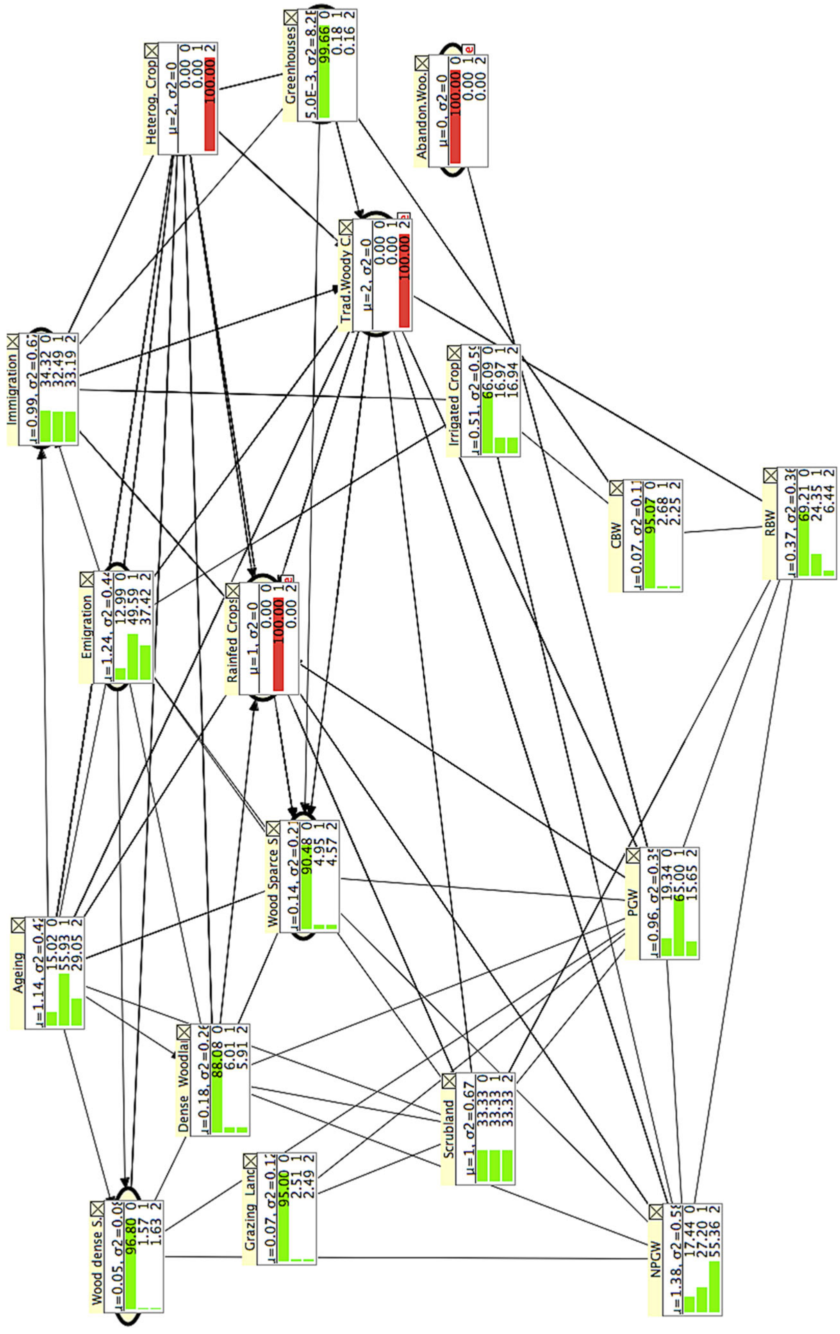

Fig. 12 Values of the marginal in the Traditional agricultural activity scenario. Evidenced variables are marked in red

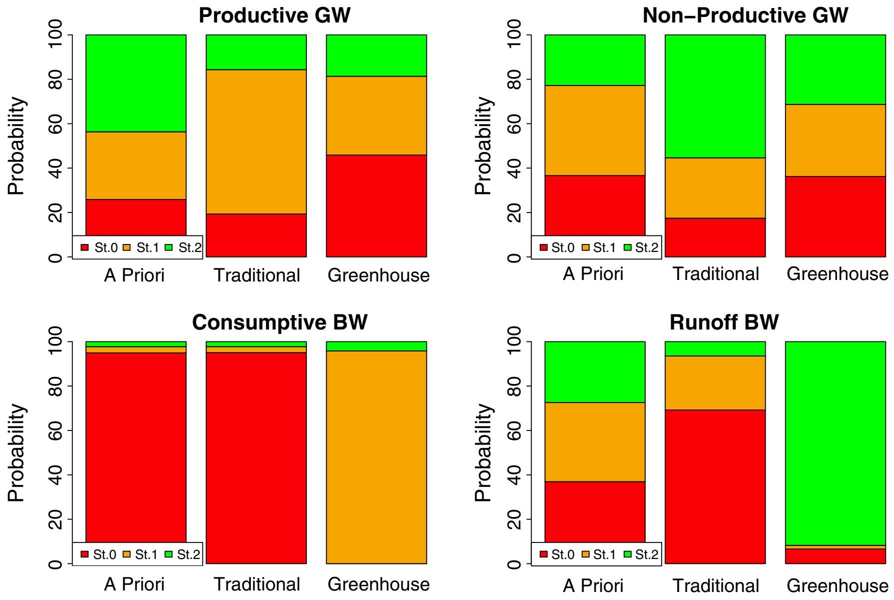

Fig. 13 Values of the probabilities of the water flows variables both A priori and under the scenarios of changes: maintenance of traditional croplands (Traditional) and Agricultural Intensification through greenhouses (Greenhouse). St. state

Table 2 Values for the ratio of *normalized likelihood* (N.L.) for pair of variables in their most probable state (expressed as RBW_0, state 0 of RBW variable) in both scenarios: maintenance of traditional croplands (traditional) and agricultural intensification through greenhouses (intensification)

Scenario	Pair	N.L.
Traditional	RBW_0 versus $NPGW_2$	1.045
	RBW_0 versus PGW_1	1.068
Intensification	CBW_1 versus PGW_0	18.34
	CBW_1 versus RBW_2	8.83

only the*Heterogeneous Crops* variable was evidenced, RBW is not affected. These traditional Mediterranean systems are characterized by an heterogeneous landscape as a result of the climate and traditional human activities (Castro-Nogueira et al. 2002; González-Bernáldez 1981) which confers a series of advantages: control over runoff water, soil development and nutrient retention (De-Lucio-Fernández et al. 2002). The promotion of this heterogeneous and traditional structure provokes an increased control of runoff water.

In order to evaluate if the scenario encourages *Green Water* or *Blue Water*, the ratio of *normalized likelihood* was calculated between two pairs of variables: RBW State 0 versus NPGW State 2, and RBW State 0 versus PGW State 1. Table 2 shows that the ratio of *normalized likelihood* in both cases are close to 1, so that, maintaining traditional Mediterranean agriculture has a similar influence over both water flows.

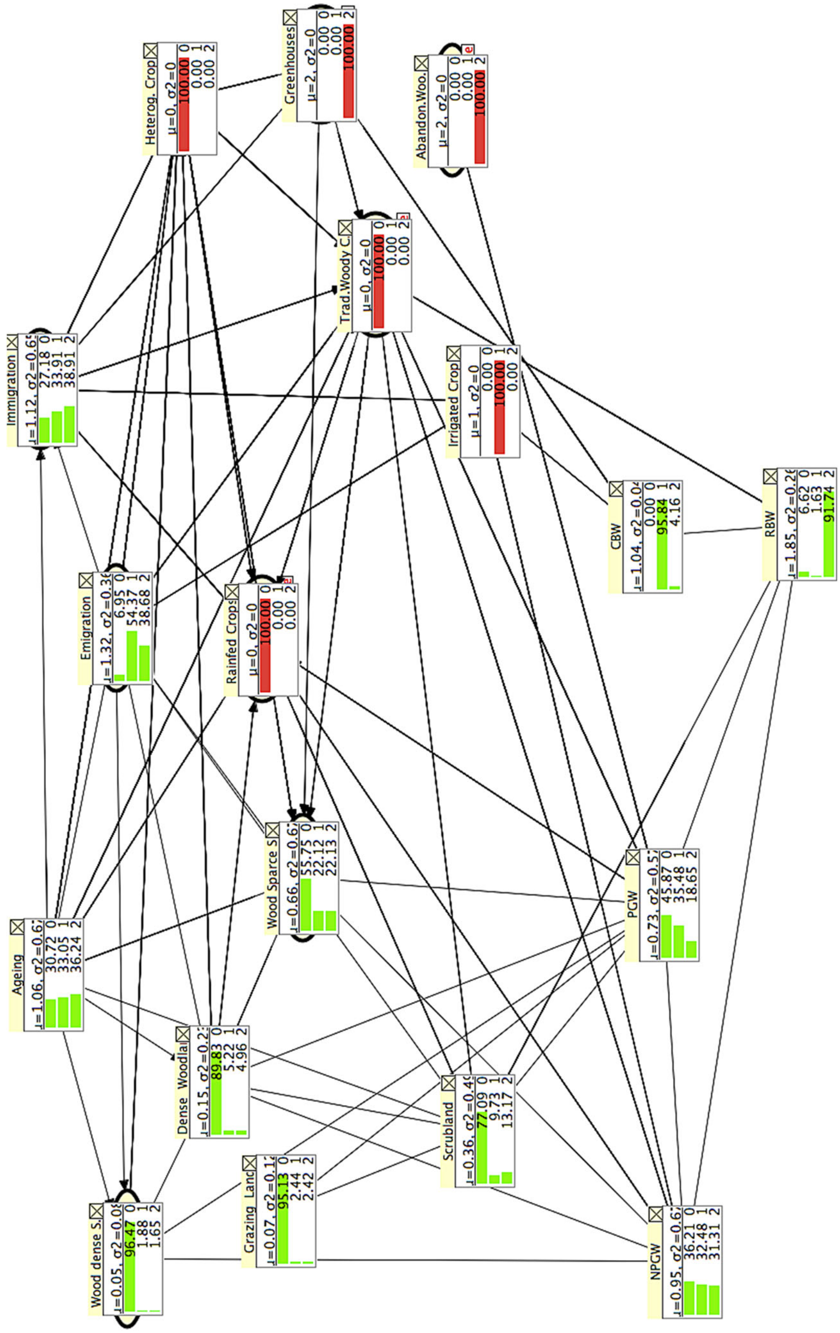

Fig. 14 Values of the marginal in the Intensive agriculture scenario. Evidenced variables are marked in red

By contrast, the promotion of intensive agriculture with greenhouses implies a more profound change in *Blue Water* (Fig. 14). In this case, the ratio of *normalized likelihood* was evaluated between CBW State 1 versus PGW State 0, and CBW State 1 versus RBW State 2, giving values of 18.34 and 8.83 respectively (Table 2). These values indicate that the scenario encourages CBW in favor of the other kinds of water flow, with a greater difference with green water flow (PGW) than against the other blue water (RBW). Figure 10 shows the values for the *Kullback–Leibler divergence* in which land uses metrics are lower than 0.1 in all cases whilst blue water presents values higher than green water flows, with a remarkable value of 15.14 for *Consumptive Blue Water*.

Looking at the probabilities CBW changes dramatically from the a priori situation in which State 0 carries a probability of 94.95, dropping to 0.00 as State 1 becomes more probable. An increase of greenhouses will cover a proportion of the catchment with a waterproof plastic surface, so increasing this CBW. Results for the impact support this change with values in some combinations of variables over 20 (Table 8). However, if the evidenced are only included into 2–3 variables including only *Heterogeneous crops* and *Traditional crops*, impact value is equal to 1, which means no change is found between a priori and a posteriori.

RBW change is also more evident with an increase from 27.43 to 91.75 of the probability in State 2 (Fig. 13). The increase in evaporative losses reduces the water available for human and agricultural supply (thus, in semiarid regions such as this, efforts need to focus on optimising water use and minimising water losses). Even when this change is significant, values for the *Kullback–Leibler* and *normalized likelihood* (Tables 9) are less deep than CBW values.

Changes in *Green Water* are less evident. Natural areas, PGW and NPGW become more probable in State 0. The promotion of greenhouses in the second scenario implies the substitution of extensive traditional crops and makes both *Productive* and *Non Productive Green Water* more probable in the lowest values. These results can be seen in Fig. 10 where values for *Green Water* are higher in the first scenario than in the promotion of greenhouses scenario. Besides, *normalized likelihood* values are close to 1 in both cases (Tables 10 and 11).

4 Conclusions

A model based on BNs was trained to study the relationships between socioeconomy, land uses and water flows in a Mediterranean catchment. The impact of two land use change trends on both social structure and water flows were also evaluated. Taking advantage of the probabilistic nature of BNs, several metrics were obtained from the results in such a way that the interpretation is richer. In particular, Sensitivity Analyses to Evidence and to Parameters were carried out by evaluating *normalized likelihood* values and the ratio of *normalized likelihood*, and the *Kullback–Leibler divergence* between a priori and a posteriori distributions were calculated.

Once the relationships between the three components of the socioecosystems had been analysed, two scenarios of land use change trend were evaluated. The maintenance of traditional mediterranean croplands implies a more elderly population who establish their livelihood in the area, attracted by the heterogeneous landscape and warm climate. At the same time, this heterogeneous landscape brings with it greater control of runoff flows and an increase in primary productivity, measured as green water flows. By contrast, the scenario of agricultural intensification through greenhouses provokes a profound change in how water flows in the catchment are propagated, with a significant increase in evaporative flows (both *Consumptive* and *Runoff Blue Water*) which, in a semiarid area, implies an important water losses.

The ability of BNs to face some of the challenges identified in the water management field has been demonstrated. Firstly, the uncertainty inherent in natural systems can be included into the model by means of probability values. In this paper, the changes studied are not only identified in terms of directions (will the water flow increase or decrease?), but also in terms of the probability of this change (this water flow will become 20% more or less probable) and the impact.

The model shown in this paper was trained using both literature and expert knowledge combined with an automatic training algorithm. Taken together with the visual component of BNs, it allows complex problems to be modeled, taking account of multiple relationships between variables, and making them more easily understood by non-mathematical experts.

However, some considerations have been made and future works can be identified. In one hand, Balance MED model is an appropriate model for green and blue water calculation in Mediterranean areas, but act, in certain way, as a black box. So, new hydrological model able to calculate these flows but in a stochastic or probabilistic way is needed. By this way, sensitivity analysis can be carried out in this step of the problem. Variables collected were continuous, but they were discretized for structural reasons. Another point would be develop this kind of model using continuous variables. Besides, water flows and territory are not stationary, and evolve along time. For that reason, a step forward would be the develop of a temporal model able to study the evolution of these water flows according to real and dynamic changes into the territory and socioeconomic structure.

Acknowledgements This work has been supported by the Spanish Ministry of Economy and Competitiveness through projects TIN2013-46638-C3-1-P and TIN2016-77902-C3-3-P; by the Junta de Andalucía through project P12-TIC-2541, and from ERDF funds.

Appendix A

Definition and thresholds for intervals/state for the discretization in the variables included in the model.

See Table 3.

Table 3 Variables, definition and thresholds for the intervals/states of the discretization

Variable	Definition	Thresholds
Ageing	Percentage of the total population older than 65 years old	22.6–26.1
Emigration rate	Percentage of the total population that migrate out of the study area	2.2–2.8
Immigration rate	Percentage of the total population that migrate into the study area	1.6–2.4
Scrubland	Areas where the main vegetation cover is scrub	0.001, 3.26
Woodland and sparse scrubland	Areas of woodland with less than 50% of scrubland	0.001, 2.7
Woodland and dense scrubland	Areas of woodland with more than 50% of scrubland	0.001, 3.04
Dense woodland	Areas of woodland with more than 50% of trees	0.001, 2.9
Riverbed vegetation	Areas surrounded the riverbed with natural vegetation	0.001, 0.8
Abandonment woody crops	Areas of abandonment olive, grapes and almond crops	0.001, 2.1
Traditional woody crops	Areas with extensive agriculture of olive, grapes and almonds, mixed with herbaceous crops and small patches of scrub	0.001, 2.1
Sparse vegetation	Areas of soil and a low density of vegetation cover	0.001, 4.1
Rainfed crops	Areas of crops not irrigated	0.001, 1.9
Grazing land	Areas of pasture	0.001, 2.2
Riverbed areas	Areas surrounded the riverbed without vegetation	0.001, 1.4
Heterogeneous crops	Areas with a mixture of different crops, both rainfed and irrigated	0.001, 2.6
Wood and grazing land	Areas of woodland with grazing	0.001, 3.2
Irrigated crops	Areas of irrigated crops	0.001,1.5
Greenhouses	Intensive agriculture under greenhouses	0.001, 0.9
Human infrastructures	Those infrastructures made by human	0.001, 1.1
Water areas	Rivers, wetlands and estuary	0.001, 0.5
Non-productive green water	fraction of the total precipitation intercepted by vegetation or soil and evaporates back to the atmosphere	27.5–83.7
Productive green water	fraction of the total precipitation reaches the soil and taken up by plants and transpired	203.8–242.9
Consumptive blue water	fraction of the total precipitation intercepted on impermeable surfaces and then returned to the atmosphere	10.7–468.1
Runoff blue water	fraction of the total precipitation that can can either percolate or drain when infiltrated water exceeds the soil storage capacity	201.1–280.2

Appendix B

This appendix shows the tables of *Normalized likelihood* values for all water flows variables in both scenarios.

See Tables 4, 5, 6, 7, 8, 9, 10, 11.

Table 4 *Normalized likelihood* values for CBW in the scenario of maintenance of tradicional croplands

Het.crops	Aban.W.crops	Trad.W.crops	Rainfed	St.0	St.1	St.2
False	False	False	False	1	1	1
False	False	False	true	1	1.04	1.02
False	False	True	False	1	0.97	1
False	False	True	True	0.99	1.08	1.11
False	True	False	False	1	1	1
False	True	False	True	1	1.04	1.02
False	True	True	False	1	0.97	1
False	True	True	True	0.99	1.08	1.11
True	False	False	False	1	0.97	1
True	False	False	True	1	1.01	1.06
True	False	True	False	1	0.96	1
True	False	True	True	1	0.96	1
True	True	False	False	1	0.97	1
True	True	False	True	1	1.01	1.06
True	True	True	False	1	0.96	1
True	True	True	True	1	0.96	1

St. state

Table 5 *Normalized likelihood* values for RBW in the scenario of maintenance of tradicional croplands

Het.crops	Aban.W.crops	Trad.W.crops	Rainfed	St.0	St.1	St.2
False	False	False	False	1	1	1
False	False	False	True	1.11	1.05	0.81
False	False	True	False	1.74	0.81	0.34
False	False	True	True	1.75	0.77	0.38
False	True	False	False	1	1	1
False	True	False	True	1.11	1.05	0.81
False	True	True	False	1.74	0.8	0.34
False	True	True	True	1.75	0.77	0.38
True	False	False	False	1.06	1.01	0.92
True	False	False	True	1.21	0.96	0.79
True	False	True	False	1.88	0.72	0.27
True	False	True	True	1.88	0.72	0.27
True	True	False	False	1.06	1.01	0.91
True	True	False	True	1.22	0.96	0.78
True	True	True	False	1.89	0.71	0.27
True	True	True	True	1.89	0.71	0.27

St. state

Table 6 *Normalized likelihood* values for PGW in the scenario of maintenance of tradicional croplands

Het.crops	Aban.W.crops	Trad.W.crops	Rainfed	St.0	St.1	St.2
False	False	False	False	1	1	1
False	False	False	True	0.59	1.47	0.93
False	False	True	False	1.02	1.49	0.64
False	False	True	True	1.06	1.11	0.88
False	True	False	False	0.99	1	1.01
False	True	False	True	0.59	1.47	0.94
False	True	True	False	1.02	1.49	0.64
False	True	True	True	1.06	1.12	0.88
True	False	False	False	0.85	0.89	1.18
True	False	False	True	0.62	1.14	1.15
True	False	True	False	0.74	2.09	0.4
True	False	True	True	0.74	2.09	0.4
True	True	False	False	0.84	0.89	1.19
True	True	False	True	0.62	1.14	1.15
True	True	True	False	0.73	2.11	0.4
True	True	True	True	0.73	2.11	0.4

St. state

Table 7 *Normalized likelihood* values for NPGW in the scenario of maintenance of tradicional croplands

Het.crops	Aban.W.crops	Trad.W.crops	Rainfed	St.0	St.1	St.2
False	False	False	False	1	1	1
False	False	False	True	0.49	1.19	1.52
False	False	True	False	0.69	0.77	1.94
False	False	True	True	0.96	0.74	1.53
False	True	False	False	1	1	1
False	True	False	True	0.49	1.19	1.52
False	True	True	False	0.69	0.77	1.94
False	True	True	True	0.96	0.74	1.53
True	False	False	False	1.23	0.99	0.64
True	False	False	True	1.45	0.64	0.9
True	False	True	False	0.49	0.7	2.38
True	False	True	True	0.49	0.7	2.38
True	True	False	False	1.23	0.99	0.64
True	True	False	True	1.45	0.63	0.89
True	True	True	False	0.49	0.7	2.38
True	True	True	True	0.49	0.7	2.38

St. state

Table 8 *Normalized likelihood* values for CBW in the scenario of agricultural intensification trough greenhouse

Het.crops	Aban.W.crops	Trad.W.crops	Greenh	Irrg.crops	Rainfed	St.0	St.1	St.2
False	False	False	True	False	True	0	32.62	4.04
False	False	False	True	True	False	0	34.48	1.75
False	False	False	True	True	True	0	34.44	1.8
False	False	True	False	False	False	1	1	1
False	False	True	False	False	True	1	1	1
False	False	True	False	True	False	0.09	24.62	10.05
False	False	True	False	True	True	0.07	25.15	10.28
False	False	True	True	False	False	0	32.64	4.02
False	False	True	True	False	True	0	32.62	4.03
False	False	True	True	True	False	0	34.45	1.79
False	False	True	True	True	True	0	34.43	1.81
False	True	False	False	False	False	1	1	1
False	True	False	False	False	True	1	0.99	1
False	True	False	False	True	False	0.16	22.85	9.43
False	True	False	False	True	True	0.11	24.21	9.95
False	True	False	True	False	False	0	32.64	4.02
False	True	False	True	False	True	0	32.63	4.02
False	True	False	True	True	False	0	34.48	1.75
False	True	False	True	True	True	0	34.45	1.78
False	True	True	False	False	False	1	1	1
False	True	True	False	False	True	1	1	1
False	True	True	False	True	False	0.09	24.62	10.05
False	True	True	False	True	True	0.07	25.2	10.28
False	True	True	True	False	False	0	32.64	4.02
False	True	True	True	False	True	0	32.64	4.02
False	True	True	True	True	False	0	34.45	1.79
False	True	True	True	True	True	0	34.44	1.8
True	False	False	False	False	False	1	1	1
True	False	False	False	False	True	1	0.99	1
True	False	False	False	True	False	0.09	24.68	10.21
True	False	False	False	True	True	0.07	25.18	10.46
True	False	False	True	False	False	0	32.64	4.02
True	False	False	True	False	True	0	32.62	4.04
True	False	False	True	True	False	0	34.44	1.8
True	False	False	True	True	True	0	34.41	1.84
True	False	True	False	False	False	1	1	1
True	False	True	False	False	True	1	1	1
True	False	True	False	True	False	0.04	26	10.66
True	False	True	False	True	True	0.04	25.99	10.71
True	False	True	True	False	False	0	32.64	4.02

Table 8 continued

Het.crops	Aban.W.crops	Trad.W.crops	Greenh	Irrg.crops	Rainfed	St.0	St.1	St.2
True	False	True	True	False	True	0	32.62	4.03
True	False	True	True	True	False	0	34.41	1.83
True	False	True	True	True	True	0	34.39	1.86
True	True	False	False	False	False	1	1	1
True	True	False	False	False	True	1	1	1
True	True	False	False	True	False	0.09	24.68	10.21
True	True	False	False	True	True	0.06	25.24	10.44
True	True	False	True	False	False	0	32.64	4.02
True	True	False	True	False	True	0	32.64	4.02
True	True	False	True	True	False	0	34.44	1.8
True	True	False	True	True	True	0	34.42	1.82
True	True	True	False	False	False	1	1	1
True	True	True	False	False	True	1	1.01	1
True	True	True	False	True	False	0.04	26	10.66
True	True	True	False	True	True	0.03	26.02	10.69
True	True	True	True	False	False	0	32.64	4.02
True	True	True	True	False	True	0	32.64	4.01
True	True	True	True	True	False	0	34.41	1.83
True	True	True	True	True	True	0	34.4	1.84

St. state

Table 9 *Normalized likelihood* values for RBW in the scenario of agricultural intensification trough greenhouse

Het.crops	Aban.W.crops	Trad.W.crops	Greenh	Irrg.crops	Rainfed	St.0	St.1	St.2
False	False	False	True	False	True	1	1	1
False	False	False	True	True	False	0.97	0.98	1.05
False	False	False	True	True	True	0.65	0.47	2.04
False	False	True	False	False	False	0.61	0.44	2.13
False	False	True	False	False	True	0.44	0.2	2.6
False	False	True	False	True	False	0.44	0.19	2.61
False	False	True	False	True	True	0.41	0.15	2.7
False	False	True	True	False	False	0.42	0.14	2.71
False	False	True	True	False	True	0.95	1.01	1.05
False	False	True	True	True	False	0.94	0.99	1.09
False	False	True	True	True	True	0.54	0.41	2.25
False	True	False	False	False	False	0.54	0.39	2.26
False	True	False	False	False	True	0.41	0.16	2.68
False	True	False	False	True	False	0.43	0.16	2.67
False	True	False	False	True	True	0.38	0.11	2.79
False	True	False	True	False	False	0.4	0.11	2.77

Table 9 continued

Het.crops	Aban.W.crops	Trad.W.crops	Greenh	Irrg.crops	Rainfed	St.0	St.1	St.2
False	True	False	True	False	True	0.84	1.02	1.16
False	True	False	True	True	False	0.8	1.01	1.23
False	True	False	True	True	True	0.58	0.47	2.13
False	True	True	False	False	False	0.54	0.42	2.24
False	True	True	False	False	True	0.31	0.2	2.77
False	True	True	False	True	False	0.31	0.18	2.79
False	True	True	False	True	True	0.26	0.15	2.88
False	True	True	True	False	False	0.25	0.13	2.92
False	True	True	True	False	True	0.79	1.03	1.22
False	True	True	True	True	False	0.76	1.01	1.28
False	True	True	True	True	True	0.47	0.4	2.34
True	False	False	False	False	False	0.48	0.37	2.37
True	False	False	False	False	True	0.28	0.16	2.85
True	False	False	False	True	False	0.29	0.15	2.85
True	False	False	False	True	True	0.22	0.1	2.98
True	False	False	True	False	False	0.23	0.1	2.98
True	False	False	True	False	True	1	1	1.01
True	False	False	True	True	False	0.97	0.98	1.05
True	False	False	True	True	True	0.59	0.43	2.15
True	False	True	False	False	False	0.58	0.41	2.19
True	False	True	False	False	True	0.44	0.2	2.6
True	False	True	False	True	False	0.45	0.19	2.61
True	False	True	False	True	True	0.42	0.15	2.69
True	False	True	True	False	False	0.42	0.14	2.7
True	False	True	True	False	True	0.95	1.01	1.05
True	False	True	True	True	False	0.94	0.98	1.09
True	False	True	True	True	True	0.51	0.37	2.32
True	True	False	False	False	False	0.53	0.37	2.3
True	True	False	False	False	True	0.42	0.16	2.68
True	True	False	False	True	False	0.43	0.16	2.66
True	True	False	False	True	True	0.39	0.11	2.78
True	True	False	True	False	False	0.4	0.11	2.76
True	True	False	True	False	True	0.84	1.02	1.17
True	True	False	True	True	False	0.79	1.01	1.24
True	True	False	True	True	True	0.54	0.42	2.23
True	True	True	False	False	False	0.52	0.39	2.3
True	True	True	False	False	True	0.31	0.2	2.76
True	True	True	False	True	False	0.31	0.18	2.79
True	True	True	False	True	True	0.26	0.15	2.88
True	True	True	True	False	False	0.25	0.13	2.91

Table 9 continued

Het.crops	Aban.W.crops	Trad.W.crops	Greenh	Irrg.crops	Rainfed	St.0	St.1	St.2
True	True	True	True	False	True	0.78	1.03	1.23
True	True	True	True	True	False	0.75	1.01	1.29
True	True	True	True	True	True	0.45	0.36	2.41
True	True	True	False	True	True	0.46	0.35	2.41
True	True	True	True	False	False	0.28	0.16	2.85
True	True	True	True	False	True	0.29	0.15	2.84
True	True	True	True	True	False	0.23	0.1	2.98
True	True	True	True	True	True	0.23	0.1	2.97

St. state

Table 10 *Normalized likelihood* values for PGW in the scenario of agricultural intensification trough greenhouse

Het.crops	Aban.W.crops	Trad.W.crops	Greenh	Irrg.crops	Rainfed	St.0	St.1	St.2
False	False	False	True	False	True	1	1	1
False	False	False	True	True	False	1.1	0.91	1
False	False	False	True	True	True	0.99	1.05	0.97
False	False	True	False	False	False	1.09	0.97	0.96
False	False	True	False	False	True	1.01	1.04	0.97
False	False	True	False	True	False	1.11	0.94	0.97
False	False	True	False	True	True	1.02	1.05	0.95
False	False	True	True	False	False	1.13	0.96	0.95
False	False	True	True	False	True	1	0.97	1.02
False	False	True	True	True	False	1.1	0.87	1.03
False	False	True	True	True	True	1	0.97	1.03
False	True	False	False	False	False	1.1	0.9	1.01
False	True	False	False	False	True	1.02	0.99	0.99
False	True	False	False	True	False	1.12	0.9	0.99
False	True	False	False	True	True	1.03	1	0.98
False	True	False	True	False	False	1.14	0.91	0.97
False	True	False	True	False	True	1.44	1.27	0.52
False	True	False	True	True	False	1.58	1.15	0.51
False	True	False	True	True	True	1.42	1.26	0.54
False	True	True	False	False	False	1.56	1.17	0.51
False	True	True	False	False	True	1.49	1.3	0.47
False	True	True	False	True	False	1.62	1.17	0.47
False	True	True	False	True	True	1.49	1.3	0.46
False	True	True	True	False	False	1.63	1.18	0.46
False	True	True	True	False	True	1.46	1.28	0.5
False	True	True	True	True	False	1.6	1.15	0.5
False	True	True	True	True	True	1.45	1.28	0.51

Table 10 continued

Het.crops	Aban.W.crops	Trad.W.crops	Greenh	Irrg.crops	Rainfed	St.0	St.1	St.2
True	False	False	False	False	False	1.59	1.17	0.5
True	False	False	False	False	True	1.5	1.31	0.46
True	False	False	False	True	False	1.64	1.17	0.46
True	False	False	False	True	True	1.51	1.31	0.44
True	False	False	True	False	False	1.65	1.18	0.45
True	False	False	True	False	True	1.01	1.01	0.99
True	False	False	True	True	False	1.11	0.92	0.99
True	False	False	True	True	True	1.01	1.04	0.97
True	False	True	False	False	False	1.11	0.96	0.95
True	False	True	False	False	True	1	1.03	0.98
True	False	True	False	True	False	1.11	0.93	0.98
True	False	True	False	True	True	1.01	1.04	0.96
True	False	True	True	False	False	1.12	0.95	0.96
True	False	True	True	False	True	1.01	0.98	1.01
True	False	True	True	True	False	1.11	0.88	1.02
True	False	True	True	True	True	1.01	0.98	1.01
True	True	False	False	False	False	1.12	0.91	0.99
True	True	False	False	False	True	1.01	0.98	1
True	True	False	False	True	False	1.12	0.89	1
True	True	False	False	True	True	1.02	0.99	0.99
True	True	False	True	False	False	1.13	0.91	0.98
True	True	False	True	False	True	1.45	1.28	0.51
True	True	False	True	True	False	1.6	1.16	0.5
True	True	False	True	True	True	1.44	1.27	0.52
True	True	True	False	False	False	1.59	1.18	0.49
True	True	True	False	False	True	1.48	1.29	0.48
True	True	True	False	True	False	1.61	1.16	0.48
True	True	True	False	True	True	1.49	1.3	0.47
True	True	True	True	False	False	1.62	1.18	0.47
True	True	True	True	False	True	1.47	1.29	0.49
True	True	True	True	True	False	1.61	1.15	0.49
True	True	True	True	True	True	1.47	1.29	0.49
True	True	True	False	True	True	1.61	1.17	0.48
True	True	True	True	False	False	1.49	1.3	0.47
True	True	True	True	False	True	1.63	1.16	0.47
True	True	True	True	True	False	1.5	1.31	0.45
True	True	True	True	True	True	1.63	1.18	0.46

St. state

Table 11 *Normalized likelihood* values for NPGW in the scenario of agricultural intensification trough greenhouse

Het.crops	Aban.W.crops	Trad.W.crops	Greenh	Irrg.crops	Rainfed	St.0	St.1	St.2
False	False	False	True	False	True	1	1	1
False	False	False	True	True	False	1.13	0.94	0.89
False	False	False	True	True	True	0.83	0.93	1.4
False	False	True	False	False	False	0.9	0.89	1.35
False	False	True	False	False	True	0.88	0.94	1.31
False	False	True	False	True	False	0.98	0.88	1.25
False	False	True	False	True	True	0.81	0.98	1.36
False	False	True	True	False	False	0.89	0.92	1.32
False	False	True	True	False	True	1.02	1.02	0.94
False	False	True	True	True	False	1.16	0.95	0.82
False	False	True	True	True	True	0.83	0.94	1.39
False	True	False	False	False	False	0.92	0.89	1.33
False	True	False	False	False	True	0.88	0.94	1.29
False	True	False	False	True	False	0.99	0.88	1.23
False	True	False	False	True	True	0.8	0.99	1.35
False	True	False	True	False	False	0.9	0.92	1.31
False	True	False	True	False	True	1.07	0.93	1.01
False	True	False	True	True	False	1.18	0.91	0.86
False	True	False	True	True	True	0.88	0.85	1.48
False	True	True	False	False	False	0.93	0.83	1.42
False	True	True	False	False	True	0.94	0.85	1.37
False	True	True	False	True	False	1.01	0.82	1.3
False	True	True	False	True	True	0.88	0.85	1.47
False	True	True	True	False	False	0.94	0.83	1.41
False	True	True	True	False	True	1.09	0.94	0.95
False	True	True	True	True	False	1.21	0.91	0.8
False	True	True	True	True	True	0.88	0.85	1.47
True	False	False	False	False	False	0.95	0.82	1.4
True	False	False	False	False	True	0.94	0.85	1.37
True	False	False	False	True	False	1.02	0.82	1.28
True	False	False	False	True	True	0.88	0.85	1.47
True	False	False	True	False	False	0.95	0.82	1.4
True	False	False	True	False	True	0.99	1	1.02
True	False	False	True	True	False	1.12	0.94	0.91
True	False	False	True	True	True	0.83	0.94	1.39
True	False	True	False	False	False	0.91	0.89	1.34
True	False	True	False	False	True	0.89	0.93	1.3
True	False	True	False	True	False	0.99	0.87	1.24

Table 11 continued

Het.crops	Aban.W.crops	Trad.W.crops	Greenh	Irrg.crops	Rainfed	St.0	St.1	St.2
True	False	True	False	True	True	0.81	0.97	1.36
True	False	True	True	False	False	0.91	0.91	1.32
True	False	True	True	False	True	1	1.02	0.96
True	False	True	True	True	False	1.15	0.95	0.84
True	False	True	True	True	True	0.82	0.95	1.38
True	True	False	False	False	False	0.93	0.89	1.31
True	True	False	False	False	True	0.89	0.94	1.29
True	True	False	False	True	False	1.01	0.87	1.22
True	True	False	False	True	True	0.81	0.98	1.36
True	True	False	True	False	False	0.92	0.91	1.3
True	True	False	True	False	True	1.05	0.93	1.03
True	True	False	True	True	False	1.17	0.91	0.88
True	True	False	True	True	True	0.88	0.85	1.48
True	True	True	False	False	False	0.94	0.83	1.4
True	True	True	False	False	True	0.94	0.85	1.37
True	True	True	False	True	False	1.03	0.82	1.28
True	True	True	False	True	True	0.88	0.85	1.47
True	True	True	True	False	False	0.95	0.82	1.4
True	True	True	True	False	True	1.08	0.94	0.98
True	True	True	True	True	False	1.2	0.91	0.82
True	True	True	True	True	True	0.88	0.85	1.47
True	True	True	False	True	True	0.96	0.82	1.39
True	True	True	True	False	False	0.94	0.85	1.36
True	True	True	True	False	True	1.04	0.82	1.27
True	True	True	True	True	False	0.88	0.85	1.47
True	True	True	True	True	True	0.96	0.82	1.39

St. state

Appendix C

See Table 12.

Table 12 Correlation matrix

	Ag.	Im.	Em.	NPGW	PGW	CBW	RBW	Het. C	Den. W.	R. Veg.	Aban. C.	Tradi. C.
Ag.		*0.57*	*0.52*	−0.06	−0.12	0.03	−0.15	−0.02	−0.09	−0.02	−0.01	0.02
Im.	*0.57*		*0.81*	0.00	−0.11	0.09	−0.14	−0.05	−0.08	0.01	0.01	0.03
Em.	*0.52*	*0.81*		0.05	−0.12	0.12	−0.12	−0.05	−0.07	−0.01	0.00	0.02
NPGW	−0.06	0.00	0.05		0.16	*0.44*	−0.33	0.02	−0.10	−0.04	0.00	0.05
PGW	−0.12	−0.11	−0.12	0.16		−0.03	*−0.54*	0.04	0.09	0.06	−0.01	0.01
CBW	0.03	0.09	0.12	*0.44*	−0.03		0.07	−0.03	−0.03	−0.01	−0.01	−0.02
RBW	−0.15	−0.14	−0.12	−0.33	*−0.54*	0.07		−0.02	0.03	−0.01	0.00	−0.07
Het. C.	−0.02	−0.05	−0.05	0.02	0.04	−0.03	−0.02		−0.02	−0.01	0.00	−0.01
Den. W.	−0.09	−0.08	−0.07	−0.10	0.09	−0.03	0.03	−0.02		−0.01	0.00	−0.01
R. Veg.	−0.02	0.01	−0.01	−0.04	0.06	−0.01	−0.01	−0.01	−0.01		0.00	−0.01
Aban. C.	−0.01	0.01	0.00	0.00	−0.01	−0.01	0.00	0.00	0.00	0.00		0.00
Tradi. C	0.02	0.03	0.02	0.05	0.01	−0.02	−0.07	−0.01	−0.01	−0.01	0.00	
S. Veg.	0.00	0.02	0.03	−0.03	−0.03	−0.01	0.01	−0.01	−0.01	0.00	0.00	−0.01
Green	0.01	0.03	0.02	0.21	0.02	0.12	−0.01	−0.01	−0.01	0.00	0.00	0.00
W. d Sc.	0.03	0.01	0.01	−0.04	0.04	−0.02	−0.02	−0.01	−0.01	0.00	0.00	−0.01
W. dis Sc.	0.00	−0.05	−0.05	−0.07	0.07	−0.03	−0.01	−0.02	−0.02	−0.01	0.00	−0.01
Sc.	0.10	0.10	0.08	−0.02	−0.06	−0.04	−0.02	−0.02	−0.03	−0.01	−0.01	−0.02
R. areas	0.01	−0.01	−0.01	0.00	0.05	−0.01	−0.05	−0.01	−0.01	0.00	0.00	−0.01
Graz.	0.03	−0.01	0.00	−0.01	−0.04	−0.01	0.01	−0.01	−0.01	0.00	0.00	−0.01
W.-Graz.	−0.02	−0.03	−0.03	−0.04	0.01	−0.01	0.03	−0.01	−0.01	0.00	0.00	−0.01
Irrig. C	−0.01	0.01	0.01	0.23	0.02	0.33	0.05	−0.01	−0.01	0.00	0.00	−0.01
Rai. C.	−0.02	−0.04	−0.03	0.02	0.03	−0.02	−0.03	−0.01	−0.01	−0.01	0.00	−0.01

Table 12 continued

	Ag.	Im.	Em.	NPGW	PGW	CBW	RBW	Het. C	Den. W.	R. Veg.	Aban. C.	Tradi. C.
H. infras.	0.03	0.01	0.02	−0.19	−0.35	−0.03	0.25	−0.02	−0.02	−0.01	0.00	−0.01
Wat. areas	−0.01	0.03	0.01	−0.05	−0.10	0.25	−0.02	−0.01	−0.01	0.00	0.00	0.00
	S.Veg.	**Green.**	**W. d Sc.**	**W. dis Sc.**	**Sc.**	**R areas**	**Graz.**	**W.-Graz.**	**Irrig C.**	**Rai. C.**	**H. infras.**	**Wa areas**
Ag.	0.00	0.01	0.03	0.00	0.10	0.01	0.03	−0.02	−0.01	−0.02	0.03	−0.01
Im.	0.02	0.03	0.01	−0.05	0.10	−0.01	−0.01	−0.03	0.01	−0.04	0.01	0.03
Em.	0.03	0.02	0.01	−0.05	0.08	−0.01	0.00	−0.03	0.01	−0.03	0.02	0.01
NPGW	−0.03	0.21	−0.04	−0.07	−0.02	0.00	−0.01	−0.04	0.23	0.02	−0.19	−0.05
PGW	−0.03	0.02	0.04	0.07	−0.06	0.05	−0.04	0.01	0.02	0.03	−0.35	−0.10
CBW	−0.01	0.12	−0.02	−0.03	−0.04	−0.01	−0.01	−0.01	0.33	−0.02	−0.03	0.25
RBW	0.01	−0.01	−0.02	−0.01	−0.02	−0.05	0.01	0.03	0.05	−0.03	0.25	−0.02
Het. C.	−0.01	−0.01	−0.01	−0.02	−0.02	−0.01	−0.01	−0.01	−0.01	−0.01	−0.02	−0.01
Den. W.	−0.01	−0.01	−0.01	−0.02	−0.03	−0.01	−0.01	−0.01	−0.01	−0.01	−0.02	−0.01

Table 12 continued

	S.Veg.	Green.	W. d Sc.	W. dis Sc.	Sc.	R areas	Graz.	W.-Graz.	Irrig C.	Rai. C.	H. infras.	Wa areas
R. Veg.	0.00	0.00	0.00	– 0.01	– 0.01	0.00	0.00	0.00	0.00	– 0.01	– 0.01	0.00
Aban. C.	0.00	0.00	0.00	0.00	– 0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Tradi. C	– 0.01	0.00	– 0.01	– 0.01	– 0.02	– 0.01	– 0.01	– 0.01	– 0.01	– 0.01	– 0.01	0.00
S. Veg.		0.00	0.00	– 0.01	– 0.01	0.00	0.00	0.00	0.00	– 0.01	– 0.01	0.00
Green	0.00		0.00	– 0.01	– 0.01	0.00	0.00	0.00	0.00	0.00	– 0.01	0.00
W. d Sc.	0.00	0.00		– 0.01	– 0.01	– 0.01	– 0.01	0.00	– 0.01	– 0.01	– 0.01	0.00
W. dis Sc.	– 0.01	– 0.01	– 0.01		– 0.02	– 0.01	– 0.01	– 0.01	– 0.01	– 0.01	– 0.02	– 0.01
Sc.	– 0.01	– 0.01	– 0.01	– 0.02		– 0.01	– 0.01	– 0.01	– 0.01	– 0.02	– 0.03	– 0.01
R. areas	0.00	0.00	– 0.01	– 0.01	– 0.01		0.00	0.00	– 0.01	– 0.01	– 0.01	0.00
Graz.	0.00	0.00	– 0.01	– 0.01	– 0.01	0.00		0.00	– 0.01	– 0.01	– 0.01	0.00
W.-Graz.	0.00	0.00	0.00	– 0.01	– 0.01	0.00	0.00		0.00	– 0.01	– 0.01	0.00
Irrig. C	0.00	0.00	– 0.01	– 0.01	– 0.01	– 0.01	– 0.01	0.00		– 0.01	– 0.01	0.00
Rai. C.	– 0.01	0.00	– 0.01	– 0.01	– 0.02	– 0.01	– 0.01	– 0.01	– 0.01		– 0.01	0.00
H. infras.	– 0.01	– 0.01	– 0.01	– 0.02	– 0.03	– 0.01	– 0.01	– 0.01	– 0.01	– 0.01		– 0.01
Wat. areas	0.00	0.00	0.00	– 0.01	– 0.01	0.00	0.00	0.00	0.00	0.00	– 0.01	

Values over |0.4| are shown in italic *Ag.* Ageing, *Im.* Immigration, *Em.* Emigration, *NPGW* non-productive green water, *PGW* Productive Green Water, *CBW* Consumptive blue water, *RBW* Runoff Blue Water, *Het. C.* heterogeneous crops, *Den. W.* dense woodland, *R. Veg.* riverbed vegetation, *Aban. C.* abandonment crops, *Tradi. C.* traditional crops, *S.Veg.* sparse vegetation, *Green* greenhouses, *W. d Sc.* Woodland and dense scrubland, *W. dis Sc.* woodland and disperse scrubland, *Sc.* scrubland, *R areas* riverbed areas, *Graz.* grazing land, *W.-Graz.* woodland and grazing land, *Irrig C.* irrigated crops, *Rai. C.* rainfed crops, *H. infras.* human infrastructures, *Wa areas* water areas

References

Aguilera PA, Fernández A, Fernández R, Rumí R, Salmerón A (2011) Bayesian networks in environmental modelling. Environ Model Softw 26:1376–1388

Andersen SK, Olesen KG, Jensen FV, Jensen F (1990) HUGIN: a shell for building Bayesian belief universes for expert systems. In: Shafer G, Pearl J (eds) Readings in uncertain reasoning. Kaufmann, San Mateo, pp 332–337

Aranzabal ID, Schmitz MF, Aguilera PA, Pineda FD (2008) Modelling of landscape changes derived from the dynamics of socio-ecological systems. A case in a semiarid Mediterraneam landscape. Ecol Indic 8:672–685

Baynes J, Herbohn J, Russell I, Smith C (2011) Bringing agroforestry technology to farmers in the philippines: identifying constraints to the success of extension activities using systems modelling. Small Scale For 10:357–376

Bonneau M, Peyrard N, Gaba S, Sabbadin R (2016) Sampling for weed spatial distribution mapping need not be adaptive. Environ Ecol Stat 23:233–255

Bromley J, Jackson NA, Clymer OJ, Giacomello AM, Jensen FV (2005) The use of Hugin® to develop Bayesian networks as aid to integrated water resource planning. Environ Model Softw 20:231–242

Caillault S, Mialhe F, Vannier C, Delmotte S, Kedowidé C, Amblard F, Etienne M, Bécu N, Gautreau P, Houet T (2013) Influence of incentive networks on landscape changes: a simple agent-based simulation approach. Environ Model Softw 45:64–73

Casadei S, Pierleoni A, Bellezza M (2016) Integrated water resources management in lake system: a case study in central italy. Water 8(12):1–18

Castelletti A, Soncini-Sessa R (2007a) Bayesian networks and participatory modelling in water resource management. Environ Model Softw 22:1075–1088

Castelletti A, Soncini-Sessa R (2007b) Coupling real-time control and socio-economic issues in participatory river basin planning. Environ Model Softw 22:1114–1128

Castro-Nogueira H, de la Guerra MM, de Lucio-Fernández J, Alandi C, Sastre-Olmos P, Atauuri-Mezquida J, Montes C, Molina-Vázquez F, Rosarío-García-Mora M (2002) Integración territorial de espacios naturales protegidos y conectividad ecológica en paisajes mediterráneos. ISBN 84-95785-21-8

De-Lucio-Fernández J, Atauri-Mezquida J, Sastre-Olmos P, Martínez-Alandi C (2002) Conectividad y redes de espacios naturales protegidos: Del modelo teórico a la visión práctica de la gestión. In: Environmental connectivity: protected areas the mediterranean context. 26–28 September. Málaga, Spain

Falk MG, OLeary R, Nayak M, Collins P, Low-Choy S (2015) A bayesian hyurdle model for analysis of an insect resistence monitoring database. Environ Ecol Stat 22:207–226

Falkenmark M (1997) Society interaction with the water cycle: a conceptual framework for a more holistic approach. Hydrol Sci 42(4):451–466

Falkenmark M, Folke C (2002) The ethics of socio-ecohydrological catchment management: towards hydrosolidarity. Hydrol Earth Syst Sci 6(1):1–9

Foley JA, DeFries R, Asner GP, Barford C, Bonan C, Carpenter SR, CHapin FS, Coe MT, Daily GC, Gibss HK, Helkowski JH, Holloway T, Howard EA, Kucharik CJ, Monfreda C, Patz JA, Prentice IC, Ramankutty N, Snyder P (2005) Global consequences of land use. Science 309:50–574

Frayer J, Sun Z, Muller D, Munroe D, Xu J (2014) Analyzing the drivers of tree planting in Yunnan, china, with Bayesian networks. Land Use Policy 36:248–258

Fung R, Chang KC (1990) Weighting and integrating evidence for stochastic simulation in Bayesian networks. In: Uncertainty in artificial intelligence, pp 209–220

García-Álvarez-Coque JM (2002) La agricultura mediterránea en el siglo XXI. Caja Rural Intermediterránea, Cajamar, Almería. Spain, pp 7–312

García-Latorre J, Sánchez-Picón A (2001) Dealing with aridity: socio-economic structures and environmental changes in an arid Mediterranean region. Land Use Policy 18:53–64

Gitelman A, Herlihy A (2007) Isomorphic chain graphs for modeling spatial dependence in ecological data. Environ Ecol Stat 14:27–40

González-Bernáldez F (1981) Ecología y Paisaje

Gordon L, Finlayson C, Falkenmark M (2010) Managing water in agriculture for food production and other ecosystem services. Agric Water Manag 97:512–519

Grau HR, Aide TM, Zimmerman JK, Thomlinson JR, Helmer E, Zou X (2003) The ecological consequences of socioeconomic and land-use changes in postagriculture puerto rico. Bioscience 53:1159–1168

Henriksen HJ, Barlebo HC (2008) Reflections on the use of Bayesian belief networks for adaptive management. J Environ Manag 88:1025–1036

Henriksen HJ, Rasmussen P, Brandt G, von Bülow D, Jensen FV (2007) Public participation modelling using Bayesian networks in management of groundwater contamination. Environ Model Softw 22:1101–1113

Hui X, Lei C, Zehenyao S (2015) Assessment of agricultural best management practices using models: current issues and future perspectives. Water 73(3):1–21

Irvine KM, Gitelman A (2011) Graphical spatial models: a new view on interpreting spatial pattern. Environ Ecol Stat 18:447–469

Jensen FV, Nielsen TD (2007) Bayesian networks and decision graphs. Springer, Berlin

Joshi L, Wibawa G, Sinclair F (2001) Local ecological knowledge and socio-economic factors influencing farmersÕ management decisions in jungle rubber agroforestry systems in Jambi, Indonesia. DFID Project R7264 Forestry Research Programme

Kelly R, Jakeman AJ, Barreteau O, Borsuk M, ElSawah S, Hamilton S, Henriksen HJ, Kuikka S, Maier H, Rizzoli E, Delden H, Voinov A (2013) Selecting among five common approaches for integrated environmental assessment and management. Environ Model Softw 47:159–181

Kersebaum KC, Kroes J, Gobin A, Takac J, Hlavinka P, Trnka M, Ventrella D, Giglio L, Ferrise R, Moriondo M, Marta AD, Luo Q, Eitzinger J, Mirschel W, Weigel HJ, Manderscheid R, Hoffmann M, Nejedlik P, Iqbal MA, Hosch J (2016) Assessing uncertainties of water footprints using an ensemble of crop growth models on winter wheat. Water 8(12):1–20

Lambin EF, Meyfroidt P (2010) Land use transitions: socio-ecological feedback versus socio-economic change. Land Use Policy 27:108–118

Lauritzen SL, Spiegelhalter DJ (1988) Local computations with probabilities on graphical structures and their application to expert systems. J R Stat Soc Ser B 50:157–224

Maes WH, Hueuvelmans G, Muys B (2009) Assessment of land use impact on water-related ecosystem services capturing the integrated terrestrial—aquatic system. Environ Sci Technol 43:7324–7330

Mantyka-Pringle CS, Martin TG, Moffatt DB, Linke S, Rhodes JR (2014) Understanding and predicting the combined effects of climate change and land-use change on freshwater macroinvertebrates and fish. J Appl Ecol 51:572–581

Pal C, Swayne D, Frey B (2001) The automated extraction of environmentally relevant features from digital imagery using Bayesian multi-resolution analysis. Adv Environ Res 5:435–444

Pearl J (1988) Probabilistic reasoning in intelligent systems: network of plausible inference. Morgan Kaufmann, San Mateo, California

Phan TD, Smart JC, Capon SJ, Hadwen W (2016) Applications of Bayesian belief networks in water resource management: a systemic review. Environ Model Softw 85:98–111

Rockstroem J (2000) Water resources management in smallholder farms in eastern and southern Africa: an overview. Phys Chem Earth 25:275–283

Rockstroem J, Karlberg L, Wani S, Barron J, Hatibu N, Oweis T, Bruggeman A, Farahani J, Quiang Z (2010) Managing water in rainfed agriculture—the need for a paradigm shift. Agric Water Manag 97:543–550

Ropero R, Renooij S, van der Gaag L (2018) Discretizing environmental data for learning bayesian-network classifiers. Ecol Model 368:391–403

Rudel TK, Schneider L, Uriarte M, Turner B, DeFries R, Lawrence D, Geoghegan J, Hecht S, Ickowitz A, Lambin EF, Birkenholtz T, Baptista S, Grau R (2009) Agricultural intensification and changes in cultivated areas, 1970-2005. PNAS 106:20,675–20,680

Sadoddin A, Letcher RA, Jakeman A, Newham L (2005) A bayesian decision network approach for assessing the ecological impacts of salinity management. Math Comput Simul 69:162–176

Sal AG, García AG (2007) A comprehensive assessment of multifunctional agricultural land-use systems in spain using a multi-dimensional evaluative model. Agric Ecosyst Environ 120:82–91

Scanlon BR, Reedy R, Stonestrom D, Prudic D, Dennehys K (2005) Impact of land use and land cover change on groundwater recharge and quality in the southwestern US. Global Change Biol 11:1577–1593

Schmitz M, Pineda F, Castro H, Aranzabal ID, Aguilera P (2005) Cultural landscape and socioeconomic structure. Environmental value and demand for tourism in a Mediterranean territory. Consejería de Medio Ambiente. Junta de Andalucía. Sevilla

Spirtes P, Glymour C, Scheines R (1993) Causation, prediction and search, Lecture Notes in Statistics, vol 81. Springer

Stafford R, Clitherow TJ, Howlett SJ, Spiers EK, Williams RL, Yaselga B, Valarezo SZ, Vera-Izutieta DF, Cornejo M (2016) An integrated evaluation of potential management processes on marine reserves in continental Ecuador based on a Bayesian belief network model. Ocean Coast Manag 121:60–69
Teegavarapu RSV (2010) Modeling climate change uncertainties in water resources management models. Environ Model Softw 25:1261–1265
Uusitalo L (2007) Advantages and challenges of Bayesian networks in environmental modelling. Ecol Model 203:312–318
Van Deer Gag L, Renooij S (2001) Analysing sensitivity data from probabilistic networks. UAI pp 530–537
Varis O, Kuikka S (1997) BENE-EIA: a Bayesian approach to expert judgment elicitation with case studies on climate change impacts on surface waters. Clim Change 37:539–563
Willaarts BA (2009) Dinámica del paisaje en la Sierra Norte de Sevilla. Cambios funcionales e implicaciones en el suministro de servicios de los ecosistemas. PhD thesis, Facultad de Ciencias Experimentales. Departamento de Biología Vegetal y Ecología. Universidad de Almería
Willaarts BA, Volk M, Aguilera PA (2012) Assessing the ecosystem services supplied by freshwater flows in Mediterranean agroecosystems. Agric Water Manag 105:21–31
Zhang K, Peters J, Janzing D, Scholkopf B (eds) (2012) Kernel-based conditional independence test and application in causal discovery. In: UAI'11 proceedings of the twenty-seventh conference on uncertainty in artificial intelligence, pp 804–813

Rosa F. Ropero is a Ph.D. student in the Department of Ecology of the University of Almería (Spain). Her research insterests include the application of Bayesian Networks in Environmental Science and Ecology.

Rafael Rumí is an assistant professor in the Department of Mathematics of the University of Almería (Spain). He carries out research in inference and learning about Hybrid Bayesian Networks with mixtures of truncated exponentials (MTE), with applications to finance and environmental problems.

Pedro A. Aguilera is an associate professor of Ecology at the University of Almería (Spain). His research interests include Numerical Ecology and the application of Intelligent Systems in Ecology.

Affiliations

Rosa F. Ropero[1] · Rafael Rumí[2] · Pedro A. Aguilera[1]

✉ Rosa F. Ropero
rosa.ropero@ual.es

Rafael Rumí
rrumi@ual.es

Pedro A. Aguilera
aguilera@ual.es

1 Informatics and Environmental Research Group, Department of Biology and Geology, University of Almería, Almería, Spain

2 Department of Mathematics, University of Almería, Almería, Spain