Possibilities and limits of prospective GIS land cover modelling–a compared case study: Garrotxes (France) and Alta Alpujarra Granadina (Spain)

Martin Paegelow, Maria T. Camacho Olmedo

Research Article

Possibilities and limits of prospective GIS land cover modelling—a compared case study: Garrotxes (France) and Alta Alpujarra Granadina (Spain)

MARTIN PAEGELOW*† and MARÍA TERESA CAMACHO OLMEDO‡

†GEODE UMR 5602 CNRS, Maison de la Recherche, 5, all. A. Machado, F-31058 Toulouse cedex, France

‡Instituto de Desarrollo Regional, Universidad de Granada C/Rector López Argüeta s/n, 18071 Granada, Spain

This study focuses on the possibilities and the limits of a prospective GIS land cover modelling applied to two case studies (France and Spain). The methodology, based on available GIS tools, consists of using earlier land cover maps and relevant environmental factors (calibration data) to model actual, known land cover to validate the model. The model aggregates Markov chain analysis for time prediction and multi-critera evaluation, multi-objective evaluation and cellular automata to perform spatial contiguity of modelled land cover scores. The first results give an accurate, pixel by pixel prediction rate of approximately 75%. An important issue of this study consists of analysing prediction residues to improve the model.

1. Introduction

This paper focuses on prospective, fine resolution, land cover modelling applied to two Mediterranean mountain regions (19 and 25 m grid cells, respectively). This means extrapolation over space and time in the context of complex social and environmental systems. In this frame the principal problems are non linear system behaviour, a high number of relevant criteria but a low number of available data and variable inertia of land cover to human activities and natural factors spread out in space.

Basically land cover modelling means time interpolation or extrapolation when the modelling exceeds the known period. Prospective modelling is the prediction of a future state. Applied to only spatial problems, the geostatistics such as variogramme analysis and kriging offer improved tools which are commonly implemented in GIS. Tools for time modelling only appeared in GIS during recent years and should be considered as interesting and experimental algorithms rather than improved and operational instruments for decision support. At the same time the social supply for decision support and modelling tools is increasing quickly so as to be able to assist different management tasks like risk prevention, land planning and environmental management.

---

*Corresponding author. Email: paegelow@univ-tlse2.fr

Among methodological approaches for prospective modelling of high resolution land cover data we can make a distinction between automatic and supervised models. We call *automatic* a model of which the algorithm analyses the relationship among calibration data, i.e. land cover training maps and land cover relevant criteria also coming from time t0 of before (training period), to carry out a reasonable time projection with spatial localization. In the case of a *supervised* model this relationship has to be performed by the geographer to get information about the suitability or the probability of spatial location and time transition. The results of this thematic analysis assist the modelling process so that this approach is similar to decision support. In practice many models, like our approach, mix automatic and supervised aspects. The presented model uses an automatic (markovian) procedure to compute time transition probabilities but a geographic supervised analysis to establish land cover suitability used for spatial location of predicted time transitions. In this context we can mention various methods based on fuzzy logic(Zadeh 1988) often used in the context of GIS (Mezzadri-Centeno 1998) and based on remote sensing data (Selleron and Mezzadri-Centeno 2002). Other models are predominantly automatic with manual supervising of algorithms optimization. This is the case of statistical approaches by neural networks (Bishop 1995, Parlitz and Merkwirth 2000), particularly networks with one or more hidden layers. During the training phase of the model the weights of the hidden layer(s) are chosen to minimize the quadratic error of the data set. These neural multi-layer networks (perceptrons) are able to approach with a chosen precision any function (universal approximation) (Hornik 1993). Non linear parametric models are another interesting approach based on a special case of logistic regression: polychotomous regression where qualitative response may have more than two modalities (Kooperberg *et al.* 1997). The polychotomous regression also needs a training phase to optimize the generally used Newton-Raphson algorithm.

The authors of this article intend to contribute to advances in GIS based on prospective modelling. We use Markov chain analysis to perform land cover time-transition probabilities. The model also calls on multi-criteria evaluation to calculate land cover suitability useful for spatial allocation of predicted probability rates. The spatial integration of land cover categorical specific results is obtained by multi-objective evaluation assisted by cellular automata. This approach is easy to apply because all algorithms are implemented in available GIS software. The model may be considered as semi-automatic: Markov chain analysis automatically performs time transition probabilities. However, the spatial allocation of predicted probabilities is supervised by multi-criteria evaluation.

2. Test areas and data sets: common characteristics but regional specificities

The Mediterranean mountains are going through a deep social and economic reorganization visible by major landscape changes. In the French test area reform that began around 1850 with the decline of the traditional agro pastoral economy and an important rural exodus. Cropland was transformed into pastures and later it often became colonized back by forests. The Spanish site shows, with a century gap, the same decrease but also a more substantial economic reorganization. To minimize the influence of a specific data set, we analysed two test areas: Garrotxes (France) and Alta Alpujarra Granadina (Spain) (Paegelow and Camacho Olmedo 2003).

Garrotxes is a 8 750 ha catchment located in the western part of the department of Eastern Pyrenees (France). The difference of height between the major summit at

the north extremity (Madrès, 2 469m) and the confluence of the Cabrils, draining the basin, with the valley of Têt river (650 m) on the SE limit is important. On the right side, characterized by a ponderous geomorphologic relief based on granite, are situated almost all earlier terrace cultivations and coniferous forests (Pinus uncinata)—a space characterized by a quick speed of vegetation changes. The left bank forms a large, steep and south facing side on schist used as pasture. The demographic maximum (1820/30) meant intensive use of all natural resources. In 1826 25% of the whole area were developed as crop terraces (Napoleon cadastre of 1826). The population pulled down from 1832 inhabitants in 1826 to 92 inhabitants in 1999. Crop terraces were transformed to pastures and later became bushes or forests. Actually, the crops are completely marginal and the near future probably depends on the intensity of pastoral activity and management which will determine the extent of spontaneous forest spreading.

Alta Alpujarra Granadina (34 500 ha) forms the western part of the south side of Sierra Nevada (Spain); a region with a characteristic landscape (Camacho Olmedo *et al.* 2002a, 2002b, 2002c, Camacho Olmedo 2003) which historical evolution is documented up from the 15th century (Carcía Martínez 1999). The important difference of height (2 900 m) partially compensates the latitudinal shift to the Pyrenees. The southern limit (600 m) is close to Guadelfeo, a river separating the test area from Contraviesa. The northern limit is formed by the highest summits of the Iberian Peninsula: Mulhacen (3 479 m) and Veleta (3 396 m). In 40 years, the population pulled down from 4 200 to 1 200 today. The proximity to urban centres (Granada, Alméria or Malaga) explains an important development of rural tourism.

The maximal use of natural resources occurred at the end of the 19th century followed by a progressive decrease in agriculture. It only stopped during the 1940s (period of economic autarky). Up from the beginning of the 1960s the rural exodus became general in Spanish mountains with a successive surrender of not irrigated land and irrigated land located on high altitude. This process is followed by surrender or a semi surrender of irrigated land which is located on a lower level. The transformation of cropland into fallow land contributed to a landscape homogenizing. In the regional context, Alta Alpujarra Granadina is a significant example for this intensive process of surrender.

The GIS database includes available land cover layers and maps which deal with relevant environmental and social factors: elevation, slope, orientation, accessibility cost maps, hydrographs, geomorphology, administrative limits and information about management status of particular areas (public forests, pasture management, etc.).

Land cover layers have different origins (Napoleon cadastre, aerial photographs and terrain on the French site; aerial photographs, thematic maps and terrain in Spain) and mark out the historical evolution as far as possible: 6 land cover maps in Garrotxes (1826, 1942, 1962, 1980, 1989, 2000), 4 in Alta Alpujarra (1957, 1974, 1987, 2001).

Land cover changes in considered periods are dramatic as shown in fig. 2 and 3. In Garrotxes we can notice that crops (terrace cultivating), occupying 25% in 1826, almost disappeared in the early 1980s. They first became pastures then also scrubs and forests. In Alta Alpujarra Granadina the cultivated land pulled down from 21.9% in 1957 to 7.1% today. Irrigated land decreased from 14.6% to 5.8% but proportionally increased from 66.7% to 81.7% of the overall agricultural activities. Because the vegetal recolonization is very slow, the fallows became a constitutional
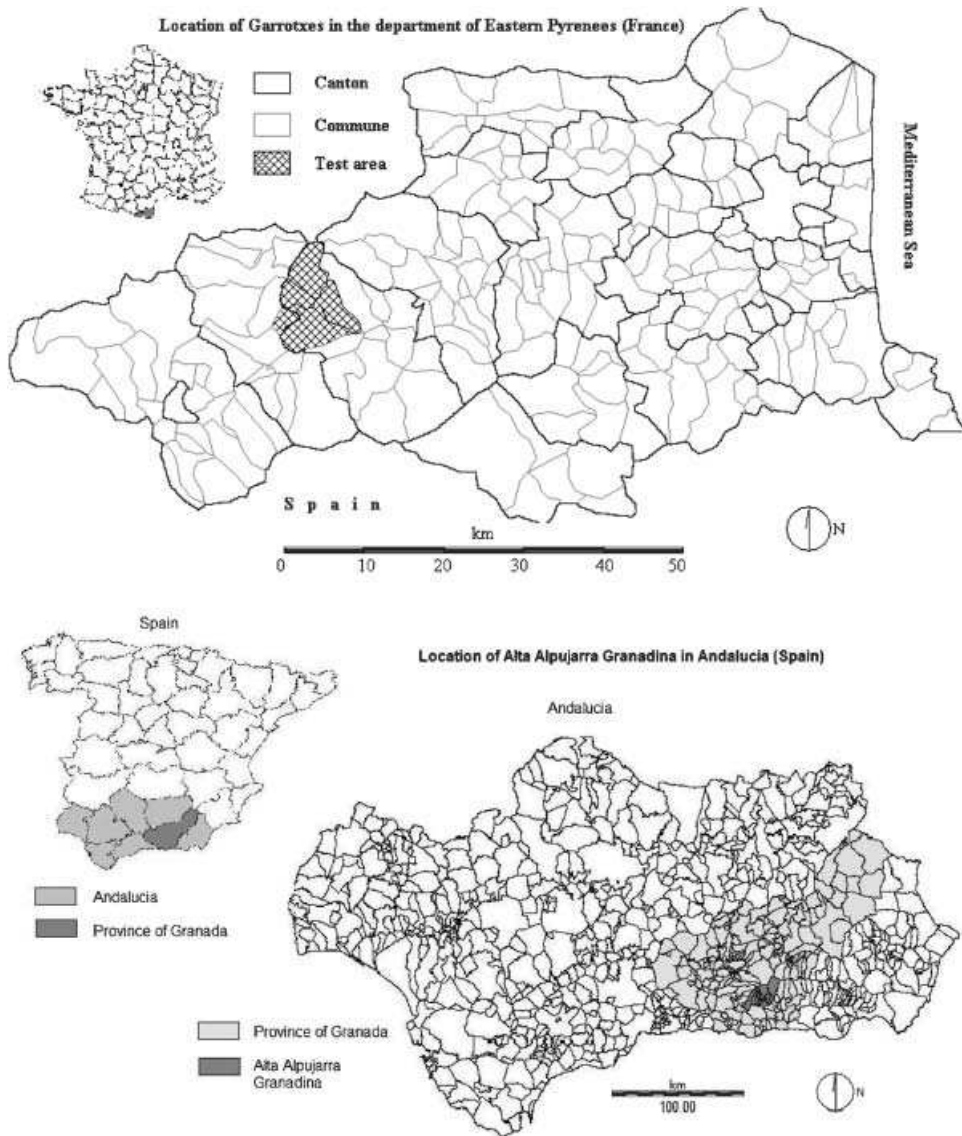
Figure 1.　Location of the test areas: Garrotxes (at the top) and Alta Alpujarra Granadina.

part of the landscape. Another important evolution concerns coniferous reforestation which increased twofold on high elevation sides.

3.　Methodology

The principal aim of this work is to establish and validate a GIS based model for prospective land cover changes. To do so we intend to use available GIS software components and a restrictive list of criteria so that the methodology would be easy to apply to other terrains (Paegelow *et al.* 2002). The validation will be performed by modelling a known land cover state, the last available date (t_1). For model calibration we use as training data the two earlier land cover layers (t_{-1} and t_0) and known and relevant environmental and social criteria coming from this training
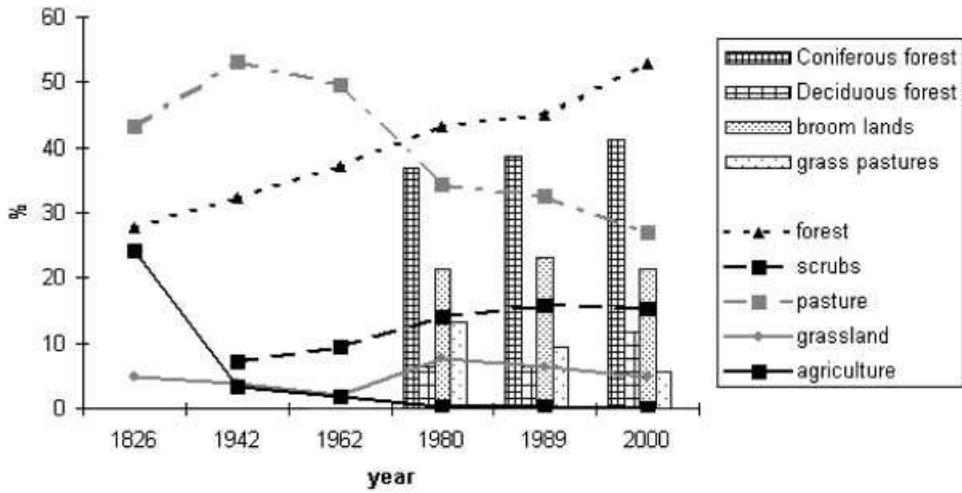
Figure 2.   Land cover net quantity changes in Garrotxes (France): 1826–2000.

period. (To intend prospective modelling we consider only the last 3 dates—the earlier dates clearly do not match contemporary economic and social conditions, particularly in Garrotxes where the first land cover layers reflect the situation at the beginning of the 19th and 20th century.) For the French test area calibration data come from 1980 and 1989; the validation date is 2000. For the Spanish test area we used 1957 and 1987 as calibration data and 2001 as the validation date. In 1974 a fine resolution cartography allowed us to distinguish between cropland and fallows
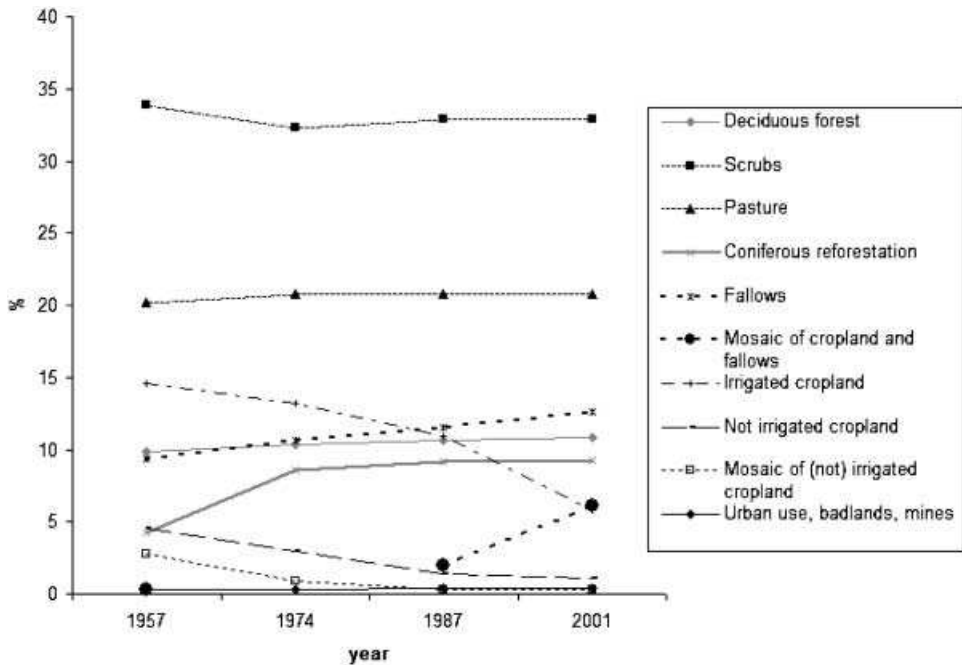

Figure 3.   Land cover net quantity changes in Alta Alpujarra Granadina (Spain): 1957–2001 (no data available for Mosaic of cropland and fallows in 1974).

also for micro parcels. For all other dates, original data (aerial photographs) do not permit correct interpretation of micro parcels. They were regrouped in a special type called 'mosaic of cropland and fallows'.

The chosen approach may be considered as a 'supervised' model with manual establishment of a knowledge base in comparison with the two other 'automatic' approaches actually performed. (Work in progress: statistical approaches by non linear parametric model and neural networks.) The land cover model calls on a chain of different tools implemented in Idrisi 32: Markovian chain analysis (MCA) to compute time transition probabilities, multi-criteria evaluation (MCE) to assist spatial implementation of predicted time transitions by suitability maps, multi-objective evaluation (MOLA) to integrate predicted land cover specific probability maps and a cellular automata (CA) introducing the principle of spatial contiguity.

3.1 *Construction of a knowledge base—the 'rules' of land cover dynamics in time and space*

The knowledge about former dynamics is essential to attempt the prediction of the future evolution or to build prospective scenarios. Therefore, any model has to be supplied with values of initial conditions (calibration). The used values are the two training land cover maps (depending variable) used to perform time transition probabilities (cf. 3.2) and land cover relevant criteria (independent variable) correlated to land cover. The statistical tests (logistic regression, PCA) helped us to choose the criteria. On the other hand we only use easily available data to facilitate model generalization.

The criteria were split up into Boolean constraints (a specific land cover is possible or not) which simply mask space, and factors (weighted and allowed to trade off each other) which express a land cover specific degree of suitability, variable in space.

In the French data set one of the used constraints was applied regardless of land cover type. Using 19 m grid cells the extension and localization of developed areas (villages) nether changed in the last two centuries. We consider these 0.23% of area stable also for the next decades. All the following model validation statistics exclude this category. The other constraints are land cover specific (e.g. elevation limits for forests, distance to roads for crops, accessibility level for grassland or masking of public woods for pasture suitability).

Among high resolution available data a set of six factors express the degree of suitability for each land cover (cf. table 1). Elevation, slope and aspect are important physical factors. The accessibility to roads and villages is a cost factor expressing

Table 1. Used factors and involved techniques to process suitability.

Factor	Technique to process suitability
Elevation	Manual recoding based on significant (99% and
Slope	99.9% level) differences between real and
Aspect	theoretic distribution
Accessibility to roads and villages (cost distance)	
Probability for land cover change	Stretching of observed transitions during t_{-1}–t_0
Proximity to existing land cover features (distance)	Fuzzy function based on observed distance parameters for border and spontaneous appearances

time used to have access to any place. In test areas transport is exclusively road based and people live in grouped residential areas. The accessibility map results from a cost distance analysis where Euclidean distance is weighted by the quality of roads. The proximity to existing land cover features is the measured distance (average and standard variation) of each land cover to borders of the same type into the training period. It takes into account border dynamics and spontaneous, not border induced appearances. This factor is important in rural and mountainous regions where spontaneous forest spreading is a widespread process. The probability for land change cover is worked out by cross tabulation of observed land cover dynamics into the training period.

Because each factor is expressed in proper units they have to be standardized to become comparable. Standardization signifies the recoding of original values (degrees, meters, per cent) to suitability values on a common scale reaching from 0 (lowest suitability) to 255 (best suitability). The factor standardization is processed in different ways. The suitability of elevation, slope, aspect and accessibility is carried out by analysing the statistical significance of the spatial distribution of a real land cover type related to one factor versus a null hypothesis (cf. table 2). Not significant differences (less than 99% level) means average suitability (128). Significantly lesser land cover as expected pulls down suitability and vice versa.

Probability for land cover change means observed net quantity change between the two calibration land cover maps. The proportional values are stretched on suitability range (linear stretch of matrix values on 0–255 range). The last factor, proximity to existing land covers features, takes into account the gain of each land cover category during t_{-1} to t_0 related to the distance (average and standard deviation) to the nearest land cover feature of the same category at t_0.

The factors, once standardized, are weighted by pairs using Saaty matrix (Saaty 1977). This matrix gives a relative importance to factors by pair comparison. Up from the set of pair comparisons Saaty matrix calculates the eigenvector of each factor, his final weight.

Used weights for elevation, slope, aspect and accessibility come from the number and strength of significant differences between real land cover specific area by factor classes and null hypothesis (cf. table 2). For most land cover categories probability for land cover change and proximity are considered as dominant factors (about 50% of overall weight). The weight of proximity depends essentially on the balance between border dynamics and spontaneous appearance.

Multi-criteria evaluation (MCE) with ordered weighted averaging (OWA) (Eastman *et al.* 1993) allows choice of risk and trade off levels. 'Trade off' means the possibility to compensate a low suitability score of one factor by a high suitability score of another factor. The allowed amount of trade off is related to assumed risk level (cf. fig. 4) of which extremities are Boolean And (risk adverse) and Boolean OR (maximum risk). The axis of risk level is called *andness*. The number of order weights is equal to the number of factors; the weights sum to 1. Order weights are calculated for each pixel. The first order weight is assigned to the factor with the lowest weighted suitability. The last order weight is assigned to the highest suitability among the weighted factors for the considered pixel.

To give whole weight to the first order weight means a risk adverse decision without any factor trade off, while the opposite means a high risk strategy because we only consider the factor the highest suitability score; an approach that also does not allow trade off (Yager 1988). We choose a strategy which may be considered as

Table 2. Significantly differences between real land cover specific area by factor classes and null hypothesis (land cover extent equal to extent of the factor class)[a].

Elevation (m)	Null hypothesis: extent of coniferous forest equal to extent of elevation classes (%)	Real coniferous forest area (%)	R–N	Suitability score
600–700	0.283	0.000	−0.283	128
700–800	0.961	0.000	−0.961	128
800–900	1.712	0.000	−1.712	128
900–1000	2.259	0.000	−2.259	128
1000–1100	2.970	0.001	−2.969	85
1100–1200	4.053	0.516	**−3.537**	42
1200–1300	5.350	1.769	**−3.581**	42
1300–1400	6.434	2.499	**−3.935**	42
1400–1500	7.460	3.556	**−3.904**	42
1500–1600	9.283	6.413	<u>−2.869</u>	85
1600–1700	12.070	12.112	0.042	128
1700–1800	13.261	21.983	**8.721**	255
1800–1900	14.197	26.478	**12.281**	255
1900–2000	9.391	16.872	**7.481**	255
2000–2100	3.566	4.094	0.528	128
2100–2200	2.192	1.709	−0.483	128
2200–2300	2.382	1.826	−0.556	128
2300–2400	2.017	0.172	−1.845	128
2400–2500	0.160	0.000	−0.160	128
Total	100.000	100.000		
Standard deviation			4.529	
99% level			<u>2.676</u>	
99.9% level			**3.419**	

[a]The table shows an example for Garrotxes: the elevation related distribution of coniferous forest. The fourth column from the left indicates the statistical significance of the difference between the real land cover distribution and the null hypothesis. Numbers in bold are significant at the 99.9% level. Underlined numbers are significant at the 99% level. All other numbers are considered as not differing significantly. The rightmost column shows the resulting suitability score: 128 (average) if the difference is not significant; 85/171 at 99% level; 42/213 at 99.9% level and 0/255 if the difference is twice the 99.9% level.

low risk taking with some trade off as shown in figure 4. Results of multi-criteria evaluation with ordered weighted averaging (MCE−OWA) are expressed as land cover specific suitability maps: one for each caption.
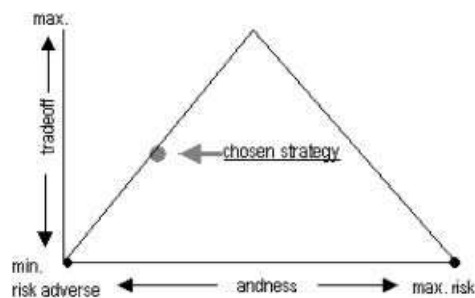

Figure 4.   Decision strategy space and chosen approach in MCE-OWA. Used order weights: 0.45, 0.20, 0.15, 0.10, 0.07, 0.03.

3.2 *Computing time transition probabilities*

To perform land cover extrapolation, we use Markov chain analysis (MCA), a discrete process with discrete time for which values at instance t_1 depend on values at instances t_0 and t_{-1} (Markov order 2). The prediction is given as an estimation of transition probabilities.

MCA produces a transition matrix recording the probability that each land cover category will change to every other category and the number of pixels expected to change. The algorithm also generates conditional probability maps for each land cover showing the probability with which it would be found at each pixel after a specified number of time units. The set of Markovian probability maps may be integrated in only one map by stochastic choice (Flamm and Turner 1994). To do so, the algorithm evaluates the conditional probabilities of each land cover for each pixel against a random probability distribution; the land cover category exceeding random threshold is assigned to the evaluated pixel (Eastman 2001). Stochastic based integration generally gives a rather poor idea of predicted land cover because it neither includes suitability knowledge nor spatial contiguity.

3.3 *Spatial allocation of predicted land cover probabilities using suitability maps*

The integrating step, combining knowledge about likely spatial distribution (MCE), time transition probabilities (MCA) as well as multi-objective land allocation, is performed in the chosen software by CA_Markov, an aggregated module. It performs a spatial distribution of transition probabilities, computed by MCA, for each land cover using MCE results. A multi-objective evaluation (MOLA) then integrates the set of predicted land cover maps. The land cover prediction procedure finally adds an element of spatial contiguity. The applied cellular automaton is based on a standard contiguity filter (5×5) to down-weight the suitability of pixels that are distant from existing instances of the land cover type under consideration. The algorithm is iterative so as to match with time distances between t_{-1}–t_0 and between t_0–t_{-1}.

The inputs are: land cover maps at instances t_{-1} and t_0, land cover suitability maps resulting from multi-criteria evaluation and area transition probabilities performed by MCA. The output is a prospective, modelled, land cover map for instance t_1. It thus allows the testing of the model by comparing it to reality. All modelling steps are resumed in figure 5.

4. Results

To validate the land cover prospective model, we applied it to model known land cover (2000—Garrotxes; 2001—Alta Alpujarra Granadina) based on information about two earlier land cover layers.

4.1 *Suitability maps performed with MCE-OWA*

The resulting probability maps (MCE-OWA) express our calibration knowledge of specific suitability for each land cover. This knowledge includes land cover distribution at each training date, land cover dynamics into this training period and land cover relevant criteria as described in section 3.1.

The suitability of the majority of land cover types fit to reality at t_0 (last calibration date, cf. table 3 and figures 6 and 7). It turns out that modelling spatial

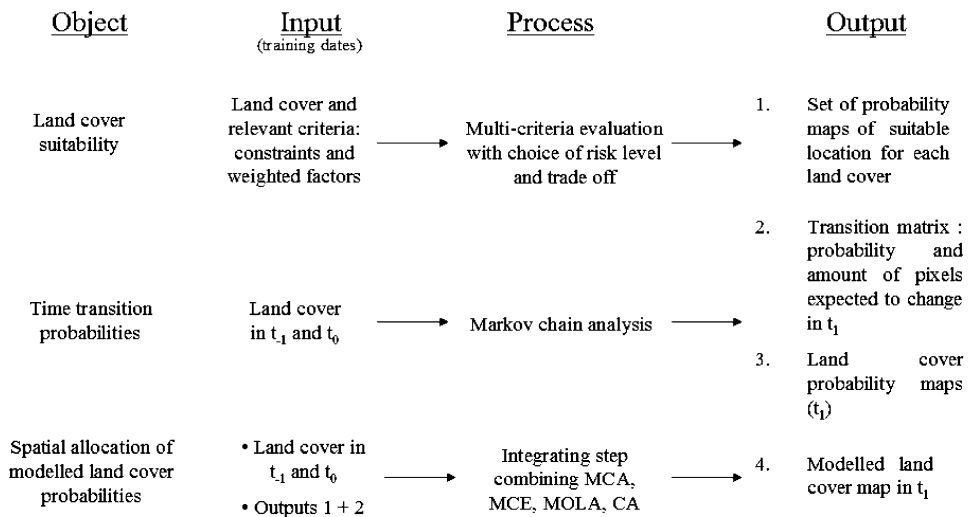| Object	Input (training dates)	Process	Output
Land cover suitability	Land cover and relevant criteria: constraints and weighted factors	Multi-criteria evaluation with choice of risk level and trade off	1. Set of probability maps of suitable location for each land cover
Time transition probabilities	Land cover in t_{-1} and t_0	Markov chain analysis	2. Transition matrix : probability and amount of pixels expected to change in t_1 3. Land cover probability maps (t_1)
Spatial allocation of modelled land cover probabilities	• Land cover in t_{-1} and t_0 • Outputs 1 + 2	Integrating step combining MCA, MCE, MOLA, CA	4. Modelled land cover map in t_1

Figure 5.   GIS-based prospective land cover modelling: methodological overview.

distribution and their dynamics fit less accurately with reality for categories having a little area (cultivated land in Garrotxes) or being very unsteady (like grassland). Alta Alpujarra Granadina even gives better suitability scores because the amount of stable area without any changes is higher into the training period (90.4%) than in Garrotxes (75.5%) where land cover dynamics is significantly faster. Table 3 shows the average and standard deviation of suitability scores for each land cover (pixels really covered at t_0) in comparison with all other categories. In this way we test the goodness of fit of calibration of the suitability maps used for spatial allocation of time transition probabilities.

To evaluate the goodness of fit of calibration for MCE-OWA performed probability maps, authors used ROC (Relative Operating Characteristic) statistics (Pontius and Schneider 2001), a measurement of agreement in terms of location. Therefore, we compare each MCE-OWA probability map expressing our calibration knowledge to a Boolean image showing the correspondent land cover distribution at the validation date t_1 (date of modelled land cover). In Garrotxes (table 4, figure 8) the relationship is very strong for coniferous forest, broom land and grass pasture. The lowest scores concern scrubs (a very dynamic category, cf. LUCC budget in section 4.3) and grassland for which the used criteria seem to be insufficient to dress a valuable suitability model. The ROC score for agriculture has to be interpreted carefully: in 2000 only 16 pixels (less than 0.01%) were crops. On the Spanish test area ROC is always near to 100% (table 5, figure 9)—a score level which may be explained by sufficient criteria and land cover stability (cf. LUCC budget, section 4.3).

4.2   *Markovian time transition probabilities and land cover caption specific probability maps*

Markovian probability scores are computed for the validation date t_1 using the calibration dates t_{-1} and t_0. Table 6 shows that predicted probability scores are always distinctly higher (average) on areas where prediction matches real land cover

Table 3. MCE performed average and standard deviation suitability scores for each land cover at t_0 compared to all other categories (data in parentheses)[a].

Garrotxes				Alta Alpujarra Granadina				
Land cover	Average		Standard deviation		Land cover	Average		Standard deviation	
Coniferous forest	191	(68)	26	(42)	Deciduous forest	255	(82)	7	(74)
Deciduous forest	125	(57)	71	(81)	Scrubs	244	(137)	16	(47)
Scrubs	163	(94)	30	(35)	Pasture	248	(65)	10	(75)
Broom land	173	(77)	29	(63)	Coniferous reforestation	251	(93)	10	(70)
Grass pasture	150	(87)	33	(50)	Fallows	223	(19)	4	(44)
Grassland	117	(84)	73	(85)	Mosaic cropland/fallows	217	(10)	2	(28)
Agriculture	148	(63)	61	(82)	Irrigated cropland	242	(33)	13	(63)
					Non-irrigated cropland	249	(7)	12	(29)
					Mosaic of (non-) irrigated cropland	217	(10)	2	(10)

[a]The average suitability scores are higher and more homogeneous on areas really occupied by the respective land cover. Figure 6 shows the probability map for land cover of being broom land in 1989 in Garrotxes. Figure 7 shows an equivalent probability map for land cover of being fallow in 1987 in Alta Alpujarra Granadina.
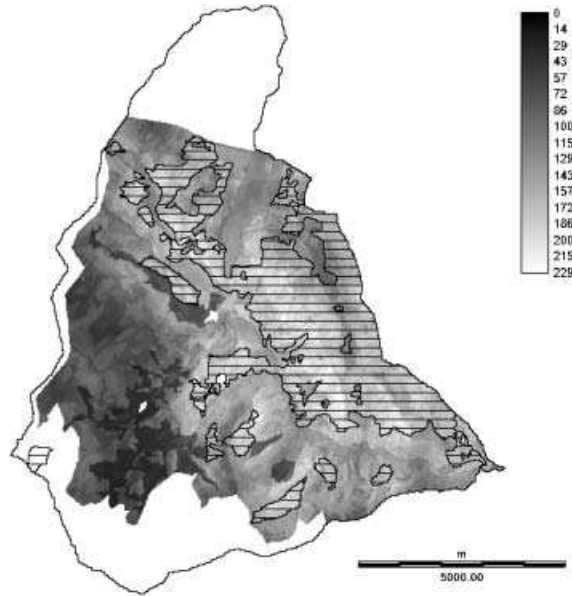
Figure 6. MCE-OWA probability map for land cover to be broom land in 1989 (Garrotxes). Horizontal stripes show the real broom land distribution in 1989. Uniform white areas (north, southwest and inside) correspond to masked land by constraints (villages inside and public forests).
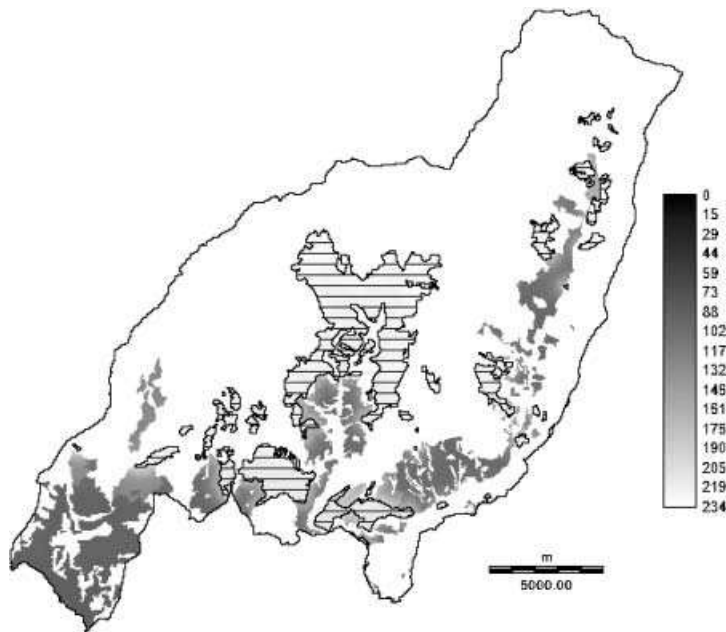

Figure 7. MCE-OWA probability map for land cover to be fallow in 1987 (Alta Alpujarra Granadina). Horizontal stripes show real fallow distribution in 1987. Uniform white areas correspond to masked land by constraints.

Table 4. Garrotxes ROC statistics based on MCE probability maps expressing the suitability of land cover specific occurrence (calibration data) and reality (validation date).

Land cover	ROC
Coniferous forest	0.949
Deciduous forest	0.768
Scrubs	0.647
Broom land	0.866
Grass pasture	0.833
Grassland	0.643
Agriculture	0.713

in 2000 (Garrotxes) or 2001 (Alta Alpujarra Granadina) but also variable (standard deviation).

A Garrotxes example of Markovian probability for land cover to be broom land is shown in figure 10 while figure 11 shows probability scores of fallows in Alta Alpujarra Granadina.

4.3 *Modelled land cover*

The modelled land cover over space and time is performed by coupling Markovian time transition probabilities with suitability maps performed by multi-criteria
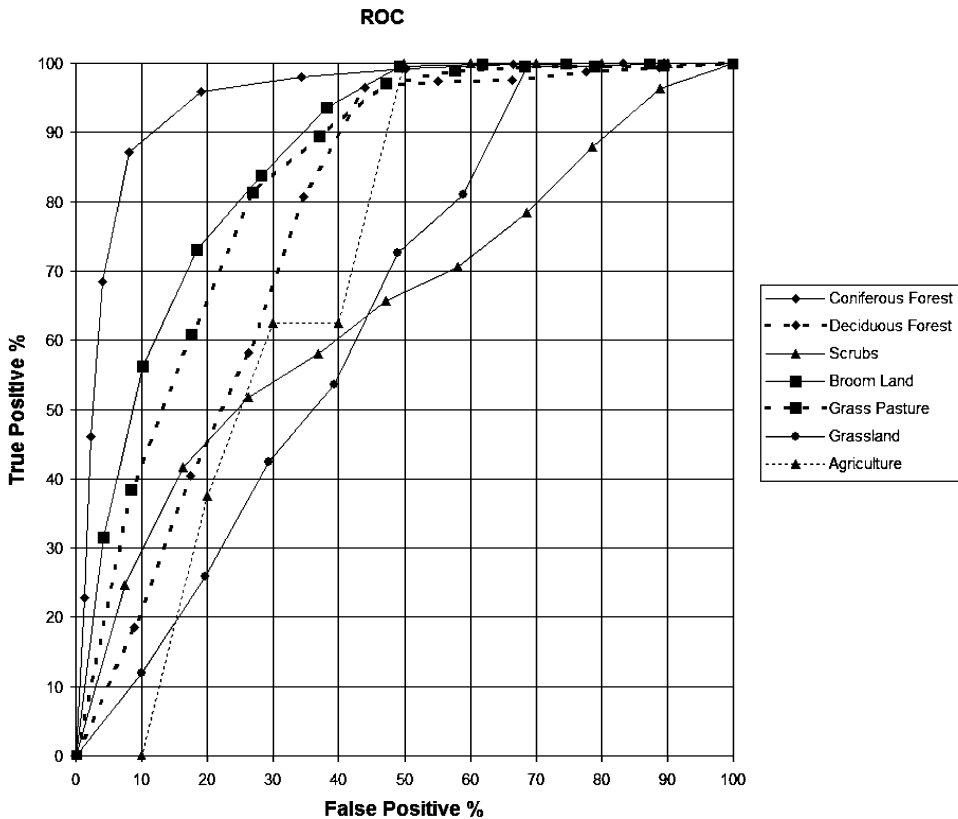

Figure 8. ROC curves, Garrotxes.

Table 5. Alta Alpujarra Granadina ROC statistics based on MCE probability maps expressing the suitability of land cover specific occurrence (calibration data) and reality (validation date).

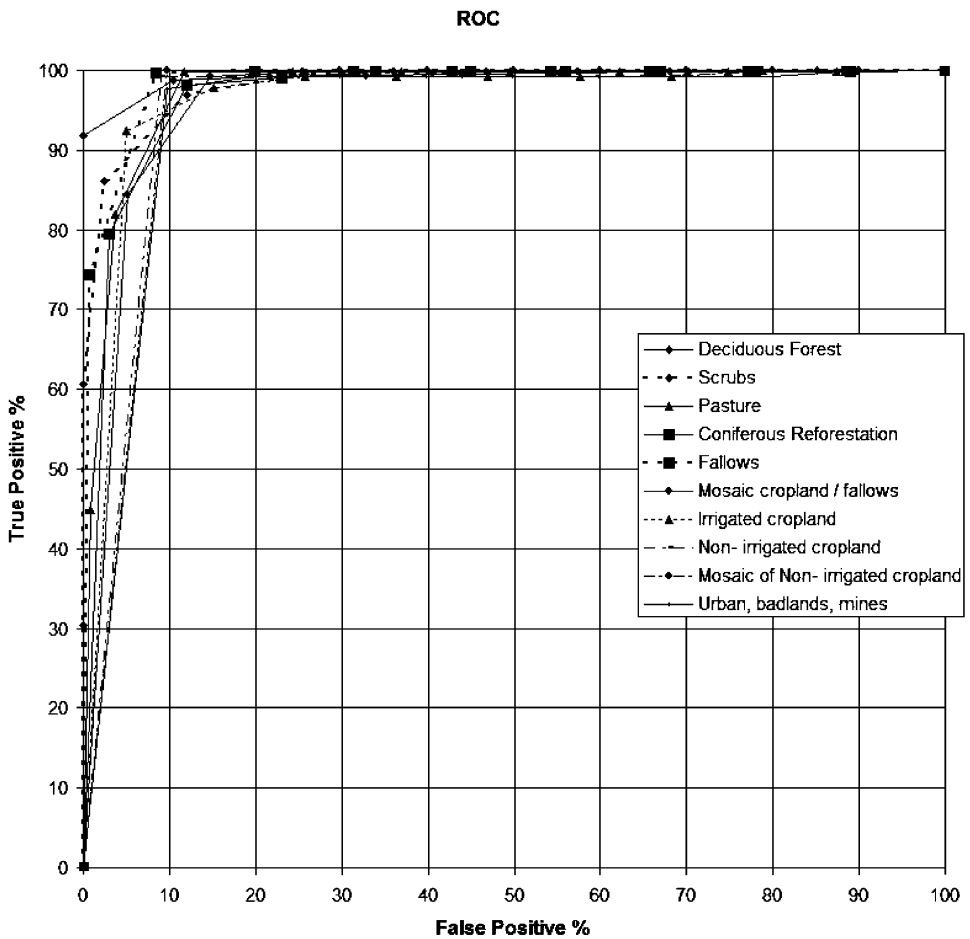| Land cover	ROC
Deciduous forest	0.991
Scrubs	0.982
Pasture	0.974
Coniferous reforestation	0.969
Fallows	0.985
Mosaic cropland/fallows	0.961
Irrigated cropland	0.963
Non-irrigated cropland	0.950
Mosaic of (non-) irrigated cropland	0.951
Urban use, badlands, mines	0.944

ROC

Figure 9.   ROC curves, Alta Alpujarra Granadina.

Table 6. Markov chain analysis predicted average probability scores (%) and standard deviation for each real land cover in t_1 in Garrotxes and Alta Alpujarra Granadina compared to all other categories (data in parentheses).

| Garrotxes					Alta Alpujarra Granadina				
Land cover	Average		Standard deviation		Land cover	Average		Standard deviation	
Coniferous forest	70	(13)	22	(19)	Deciduous forest	82	(2)	13	(2)
Deciduous forest	31	(3)	30	(5)	Scrubs	79	(7)	8	(13)
Scrubs	30	(19)	15	(10)	Pasture	85	(1)	0	(1)
Broom land	55	(13)	11	(19)	Coniferous reforestation	79	(13)	10	(7)
Grass pasture	23	(8)	9	(8)	Fallows	64	(2)	19	(9)
Grassland	47	(3)	20	(9)	Mosaic cropland/ fallows	14	(2)	8	(5)
Agriculture	1	(0)	0	(0)	Irrigated cropland	68	(5)	19	(17)
					Non-irrigated cropland	41	(0)	1	(3)
					Mosaic of (non-) irrigated cropland	17	(0)	5	(1)
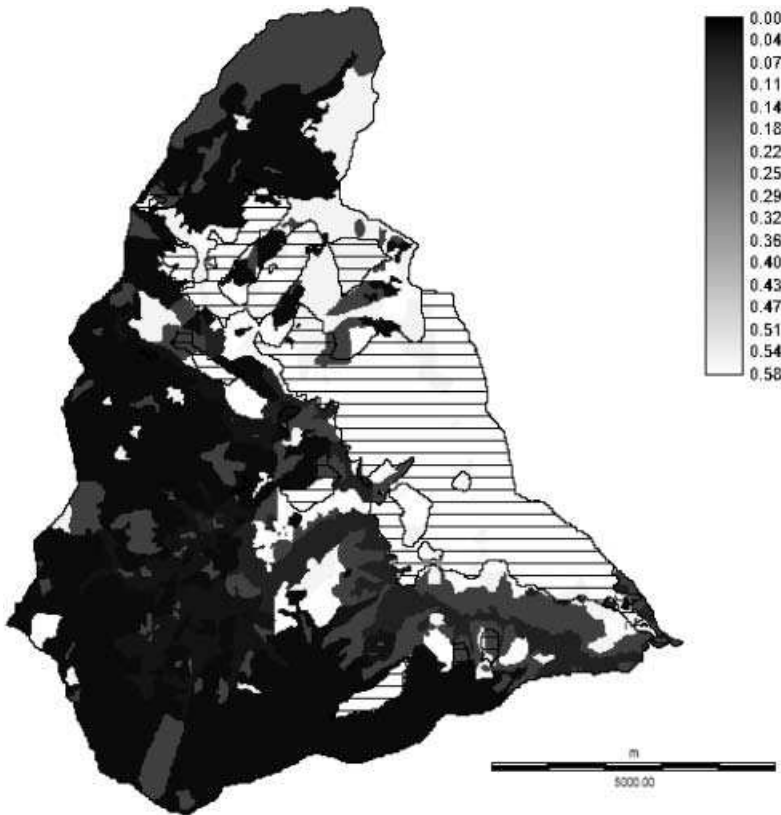
Figure 10. Markovian probability of being covered by broom land in 2000 and real broom land distribution (horizontal stripes): Garrotxes.
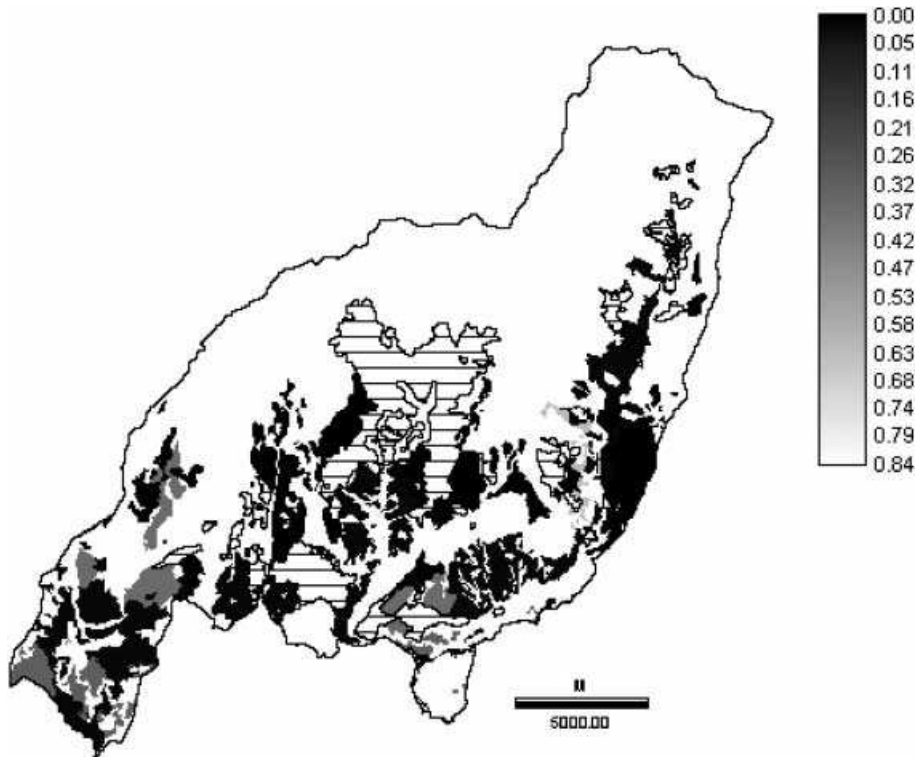
Figure 11. Markovian probability of being covered by fallows in 2001 and real fallow distribution (horizontal stripes, always on high probability): Alta Alpujarra Granadina.

evaluation. To improve land cover distribution we added cellular automata to increase spatial contiguity.

For both study areas modelled land cover is very close to reality (table 7, 8): 73% in Garrotxes, 77.6% in Alta Alpujarra Granadina but the correctness is due for a major part to persistence. In Garrotxes 75.5% of areas were stable during the training period (1980–1989), 72% between t_0–t_1 (1989–2000). The Spanish area persistence is about 90.4% during the training period (1957–1987) and concerns 83.9% of areas between the last calibration date and validation date (1987–2001).

The Garrotxes cross-matrix (table 7) confirms that the model fits better to reality for categories covering large areas like coniferous forests and broom lands. The prediction rate is lower for categories covering little areas and gets close to zero for crops covering only 0.007% of the area (16 pixels) while developed areas (urban) corresponds, by definition (constrained area), to reality. The Alta Alpujarra Granadina cross matrix (table 8) shows that the amount of modelled land cover is also close to reality except for coniferous reforestation which means land planning difficult to model. Two other categories (confusion between mosaic of cropland and fallows and irrigated cropland) do not match the model.

Figures 12 and 13 are maps of modelled land cover at t_1 compared with real land cover in t_1 and t_0. Adding the last calibration date (t_0) is helpful to evaluate the correctness of the model for land cover dynamics. Because land cover persistence is high, a null model of pure land cover persistence between t_0 and t_1 gives better results for the Spanish area (land cover persistence: 83.9%, correct modelled land

Table 7. Matrix comparing real land cover in 2000 (rows) to predicted land cover in 2000 (columns) in Garrotxes[a].

Garrotxes		Predicted land cover in 2000%								
		Coniferous forest	Deciduous forest	Scrubs	Broom lands	Grass pastures	Grassland	Agriculture	Urban	Sum of real land cover
Real land cover 2000 (%)	Coniferous forest	35.81	0.12	3.47	0.63	0.43	0.47	0.00	0.00	40.93
	Deciduous forest	1.19	6.45	2.92	0.58	0.32	0.23	0.00	0.00	11.69
	Scrubs	2.28	1.26	6.87	1.70	1.82	1.13	0.01	0.00	15.08
	Broom lands	0.32	0.40	1.68	18.00	1.12	0.09	0.00	0.00	21.61
	Grass pastures	0.21	0.08	0.44	2.64	2.29	0.00	0.00	0.00	5.66
	Grassland	0.55	0.02	0.75	0.01	0.27	3.19	0.01	0.00	4.80
	Agriculture	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.01
	Urban	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.230	0.23
Sum predicted land cover		40.37	8.33	16.13	23.56	6.24	5.12	0.03	0.23	100.00

[a]Data are expressed as a percentage of the total area. Crosswise data mean correct prediction. Data in the top-right part of the matrix indicate a predicted landscape 'opening', while data in the bottom-left part mean a predicted landscape 'closing'.

Table 8. Matrix comparing real land cover in 2001 (rows) to predicted land cover in 2001 (columns) in Alta Alpujarra Granadina[a].

| Alta Alpujarra Granadina		Predicted land cover in 2001%									
		Deciduous forest	Scrubs	Pasture	Coniferous reforestation	Fallows	Mosaic crop-land/fallows	Irrigated cropland	Non-irrigated cropland	Mosaic of (not) irrigated cropland	Urban use, badlands, mines	Sum of real land cover
Real land cover 2001 (%)	Deciduous for-est	9.31	0.04	0.00	1.27	0.10	0.00	0.14	0.00	0.00	0.00	10.86
	Scrubs	0.35	27.94	0.09	4.30	0.00	0.00	0.05	0.01	0.00	0.04	32.78
	Pasture	0.22	0.01	18.15	2.36	0.00	0.00	0.00	0.00	0.00	0.00	20.74
	Coniferous reforestation	0.01	0.16	0.00	9.05	0.00	0.00	0.00	0.00	0.00	0.00	9.22
	Fallows	0.11	1.49	0.14	1.49	8.33	0.35	0.69	0.00	0.03	0.00	12.63
	Mosaic cropland/ fallows	0.09	0.67	0.00	0.23	0.92	0.39	3.85	0.00	0.00	0.01	6.16
	Irrigated cropland	0.11	0.40	0.00	0.01	0.21	1.44	3.53	0.07	0.00	0.01	5.78
	Non-irrigated cropland	0.01	0.05	0.00	0.00	0.36	0.10	0.00	0.53	0.00	0.00	1.05
	Mosaic of (not) irrigated cropland	0.01	0.00	0.00	0.03	0.14	0.00	0.04	0.00	0.03	0.00	0.25
	Urban use, badlands, mines	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.34	0.35
Sum predicted land cover		10.21	30.76	18.38	18.74	10.22	2.28	8.31	0.61	0.06	0.40	100.00
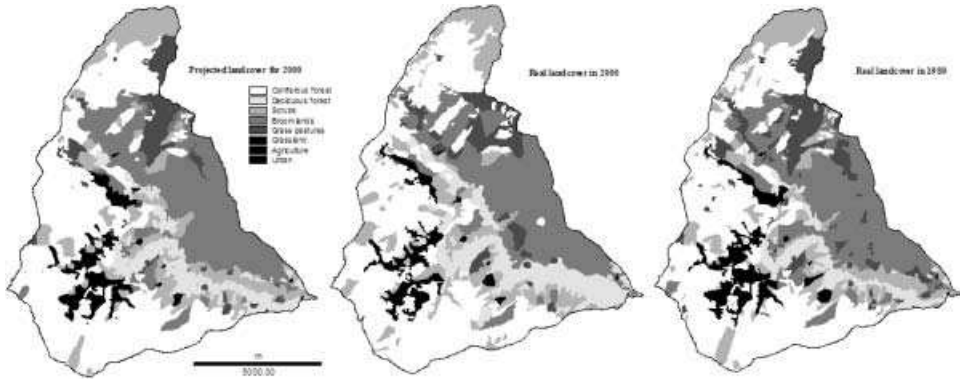
Figure 12. Modelled (2000) and real land cover (1989, 2000) in Garrotxes.

cover: 77.6%) as the used model. In French Pyrenees modelled land cover (73% of pixel correct modelled) is barely better than null hypothesis (72%).

ROC statistics (cf. section 4.1) indicate that suitability maps used for spatial allocation of time transition probabilities match with reality in terms of location. Another criterion is the amount of modelled land cover change. To do so, we compare LUCC (land use/land cover change) budgets (Pontius *et al.* 2004) of real to modelled land cover change. The following LUCC budget tables give answers to major land cover change questions: for a given period, what are the gain and loss of each category? What is the amount of total change and what is the proportion of swap (changing of location) and net change?

LUCC budgets of real land cover change (number in bold) between t_0 and t_1 show that land cover change is faster in Garrotxes within a shorter period. If net change is dominant on both study areas, the swap is very reduced in Sierra Nevada (only a quarter of total change). In other words, land cover succession (e.g. from pasture via scrubs to forest) with spatial moving of categories is active in French Pyrenees while it is nearly locked in the south of Spain (vegetal recolonization is very slow in semiarid mountains). LUCC budgets of modelled land cover change differ from reality in opposite directions. While modelled total change is half reality in Garrotxes it is twice reality in Alta Alpujarra Granadina (AAG). The responsible model step, Markov chain analysis, underestimates the amount of land cover change in Garrotxes and overestimates it on the second test area. The difference between real and modelled total change is principally imputed to swap. Real swap in Garrotxes is three times more important than modelled swap. In AAG real swap is

Figure 13. Modelled (2001) and real land cover (1987, 2001) in Alta Alpujarra Granadina.

Table 9. LUCC budget of real (number in bold) and modelled land cover changes in Garrotxes (1989–2000).

	Gain		Loss		Total change		Swap		Absolute value of net change	
Coniferous forest	**5.92**	2.43	**3.60**	0.68	**9.53**	3.11	**7.21**	1.35	**2.32**	1.76
Deciduous forest	**6.20**	2.40	**0.51**	0.08	**6.71**	2.47	**1.01**	0.15	**5.70**	2.32
Scrubs	**8.25**	3.87	**8.92**	3.49	**17.18**	7.36	**16.51**	6.98	**0.67**	0.38
Broom land	**4.34**	3.35	**6.17**	3.21	**10.51**	6.56	**8.68**	6.43	**1.83**	0.13
Grass pasture	**2.39**	0.62	**6.18**	3.81	**8.57**	4.43	**4.79**	1.23	**3.79**	3.20
Grassland	**0.83**	0.27	**2.44**	1.56	**3.27**	1.82	**1.66**	0.53	**1.61**	1.29
Agriculture	**0.01**	0.00	**0.13**	0.11	**0.14**	0.11	**0.01**	0.00	**0.12**	0.11
Total	**27.95**	12.93	**27.95**	12.93	**27.95**	12.93	**11.92**	3.75	**16.03**	9.18

only the third part of observed swap. Going into the details tables 9 and 10 show that modelled swap is underestimated particularly for coniferous forest, scrubs and grass pasture while the amount of net change is clearly below reality for deciduous forest and broom land. In AAG swap is particularly overestimated for scrubs, fallows but also for deciduous forest while net change is clearly overestimated for coniferous reforestation and, in minor proportions, for scrubs and pasture. On the contrary, net change in mosaic of cropland and fallows and irrigated cropland is underestimated by the model.

Comparing real and modelled LUCC budgets to ROC statistics (cf. tables 4 and 5) about suitability, one can notice that there is no clear relationship. First, used suitability maps produced with training data match generally well with reality in t_1

Table 10. LUCC budget of real (numbers in bold) and modelled land cover changes in Alta Alpujarra Granadina (1987–2001).

	Gain		Loss		Total change		Swap		Absolute value of net change	
Deciduous forest	**0.28**	0.88	**0.04**	1.27	**0.32**	2.15	**0.08**	1.76	**0.24**	0.39
Scrubs	**0.46**	2.60	**0.42**	4.75	**0.88**	7.35	**0.83**	5.21	**0.05**	2.15
Pasture	**0.00**	0.23	**0.00**	2.60	**0.00**	2.82	**0.00**	0.45	**0.00**	2.37
Coniferous reforestation	**0.24**	9.62	**0.15**	0.00	**0.40**	9.62	**0.31**	0.00	**0.09**	9.62
Fallows	**1.79**	1.83	**0.66**	3.10	**2.45**	4.93	**1.32**	3.65	**1.13**	1.28
Mosaic cropland/fallows	**4.85**	1.65	**0.69**	1.37	**5.54**	3.02	**1.38**	2.74	**4.16**	0.28
Irrigated cropland	**0.40**	0.08	**5.61**	2.77	**6.01**	2.85	**0.80**	0.16	**5.21**	2.70
Non-irrigated cropland	**0.00**	0.01	**0.37**	0.82	**0.37**	0.83	**0.00**	0.03	**0.37**	0.80
Mosaic of (not) irrigated cropland	**0.02**	0.00	**0.10**	0.27	**0.11**	0.27	**0.03**	0.00	**0.08**	0.27
Urban use, badlands, mines	**0.01**	0.06	**0.01**	0.00	**0.01**	0.06	**0.01**	0.00	**0.00**	0.05
Total	**8.05**	16.95	**8.05**	16.95	**16.09**	33.90	**4.76**	13.99	**11.33**	19.91

in terms of location while Markov time transition probabilities differ in opposite directions from real-time transitions. Then, we notice excellent ROC values for land cover categories causing major model failures in terms of amount of space. However, ROC statistics include land cover persistence, the preponderant land cover aspect, while LUCC budget focuses, by definition, on land cover change.

5. Discussion of results

The performed results do not intend to predict future reality but they can help us to better understand environmental and social complex changes in time and space. Therefore, the interpretation must be done carefully; the land cover modelling means simulation showing what could be reality. It is a scenario in frame of a decision support. Nevertheless, an accurate interpretation may be useful to improve the prediction rate. On both test areas the accurate prediction is around ¾ of the whole area. Prediction rate is better on the Spanish area which is also more stable particularly on the upper sides. This may appear poor compared with Null model (land cover persistence), but it is interesting to notice that the annual land cover change rate during the training period (t_{-1}—t_0; model calibration) is different from the rate during t_0—t_1 (validation date). This average annual rate decreased moderately in Garrotxes (from 2.7% to 2.5%) but increased in AAG (from 0.3% to 0.95%).

Before analysing the prediction rate it may be helpful to discuss the results of the two preliminary steps: the suitability maps and time transition probabilities.

The suitability maps performed with MCE-OWA reveal the conditions during the training period. The statistical results in table 3 show that it is easier to model land cover suitability on the Spanish test area (higher average suitability and lower standard deviation). This may be related to greater land cover persistence in Alta Alpujara Granadina. This general noting is also valid regarding specific land cover suitability scores. Table 3 shows that it is always easier to model the suitability for land cover categories of which the size is important and stable in time although they are less important to model land cover change. On the contrary, we notice lower suitability scores for land cover categories of which location is difficult to explain with used criteria (deciduous forest and grassland in Garrotxes or mosaic of cropland and fallows and mosaic of irrigated and not irrigated cropland in Alta Alpujarra Granadina). ROC statistics values generally indicate goodness of fit of calibration.

The time transition probabilities are computed automatically and based on only two input land cover maps and do not include land cover relevant criteria. The quality of results is expressed in table 6 by comparing predicted probability scores obtained by MCA to reality. Projected on real land cover at instance t_1 the average probability to be covered by a specific land cover category is even higher on locations really covered than on the rest of the study area. Generally higher and more homogeneous scores on the Spanish test area are also related to greater land cover stability. This is also true for specific categorical results. Only for scrubs (Garrotxes) the probability to be covered is almost equal on the area really covered by scrubs in 2000 than on the rest of the study area. The mentioned high speed of land cover change in Garrotxes results basically from spontaneous forest spreading. In this process the scrubs form an only transitional category, difficult to model when we use only two land cover states contrary to manual suitability modelling using also land cover relevant criteria. We notice the opposite effect for grassland which suitability is difficult to model while MCA score shows a high correspondence to

real distribution. Performed LUCC budgets indicate that MCA performed time transition probabilities explain the major part of prediction failure.

The computed land cover maps reflect previous considerations. In Garrotxes stable land cover categories with important extent like coniferous forest and broom are correctly predicted while less than half of the pixels covered by deciduous forest, scrubs and grass pasture (weak MCA probability) are correctly predicted (table 7). On the contrary, the weak suitability scores for grassland are compensated for stability in time. These considerations may also be applied to Alta Alpujarra Granadina (table 8): relative stable categories with large extent like deciduous forest, scrubs and pastures correspond in more than 85% (diagonal of table 8 compared with sum of real land cover) to reality. For fallows, characterized by a lower suitability score but stable in time we notice a correspondence of 66%. At the opposite, the modelling results for mosaic of cropland and fallows (6.3%) as well as mosaic of irrigated and not irrigated cropland (12%) do not match with real land cover. This may be explained by an inadequate, but nevertheless fine, resolution (25 m) compared with the size of concerned micro parcels. Another interesting aspect is coniferous reforestation. While suitability and time transition probabilities are elevated, the integrating step ends in an extent that is twice as reality. It is the only category of which the sum of predicted area is very different from real area; a process difficult to predict because it is a matter of land planning programmes changing abruptly, from time to time, land cover.

Another approach to interpret the quality of performed data is, depending on map legend logic, an interpretation of residues for each category (prediction failures). The Garrotxes land cover is described as qualitative data but may be put on a landscape rank scale of which the extremities are closed formations (1: coniferous and 2: deciduous forests) and open landscape corresponding to more intensive use (7: agriculture). By this way the matrix data of table 4 are mapped as the difference, expressed in parametric distance of land cover categories, as shown in fig. 14. The negative balance means a predicted landscape opening while positive values express a predicted landscape closing that did not happen. This map visualises the geographic distribution of prediction failures and their parametric importance. The prediction failures occur outside stable areas which are concentrated along the western and southern limits of the study area (public forests) and on the left bank forming large pastures. The map also shows that almost the half of false predicted area (12.9% of 27.1%) has a prediction failure of only one category.

In Alta Alpujarra it is clearly impossible to rank up land cover categories—the vegetation dynamics is slower and, consequently the land cover categories more static. So, fig. 15 focuses on specific prediction residues. Among the residues (22.39%), the principal prediction failure concerns coniferous reforestation (9.71%) of which the predicted area is twice as reality. The overestimated areas are close to earlier reforestation areas (contiguous filter) and also form islands on locations affected by forest fire during the training period. The second most important prediction failure (4.77%) is about irrigated cropland, also overestimated. It concerns especially the land near villages where, in reality, we notice a mosaic of cropland and fallows. On the other side, our model does not predict some new irrigated cropland (olive plantations and herbaceous cultures). A 2.82% overestimation of scrubs (like deciduous forest) shows that regeneration of natural vegetation, mainly on fallows, is slower than predicted. The summed rest of prediction failure (0.41%) concerns a variety of different but less significant situations.
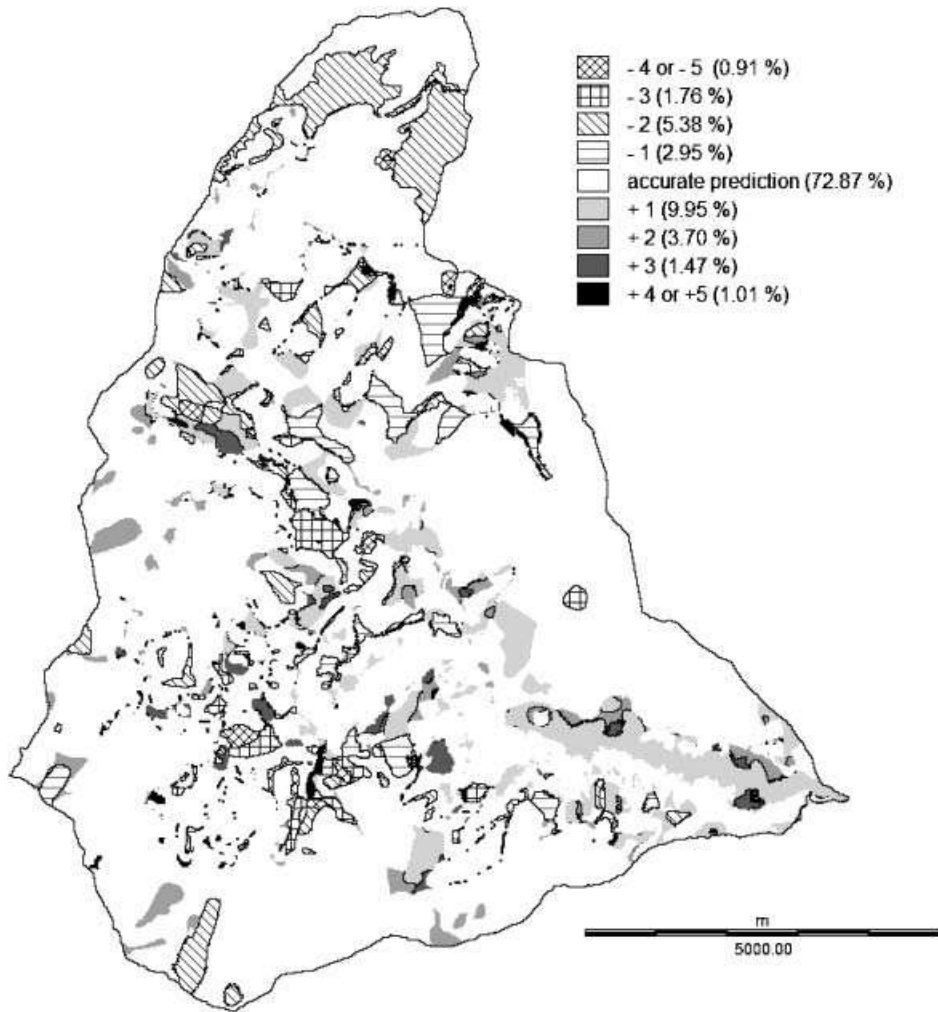
Figure 14.   Mapped difference (real minus modelled land cover at t_1) of land cover ranks in Garrotxes.

6.   Conclusion and outlook

The first results can lead us to draw some conclusions concerning the data set and the methodology. The land cover is described as qualitative data by a small number of land cover categories. The disadvantage is certain variability inside of the categories. This may be illustrated by an example on Garrotxes. Figure 12 shows in the south eastern part a large oblong area predicted as scrubs (same land cover as in 1980 and 1989). In reality this area changed by natural vegetation dynamics and became a deciduous forest emphasizing a general problem while handling with qualitative data: intra-class changes. In this case scrubs became even higher and more compact. At the same time the floristic composition changed and Quercus ilex got the upper hand.

Another limiting factor resulting from used data is training time series. Only two land cover maps were used for calibration. On the other hand, social and
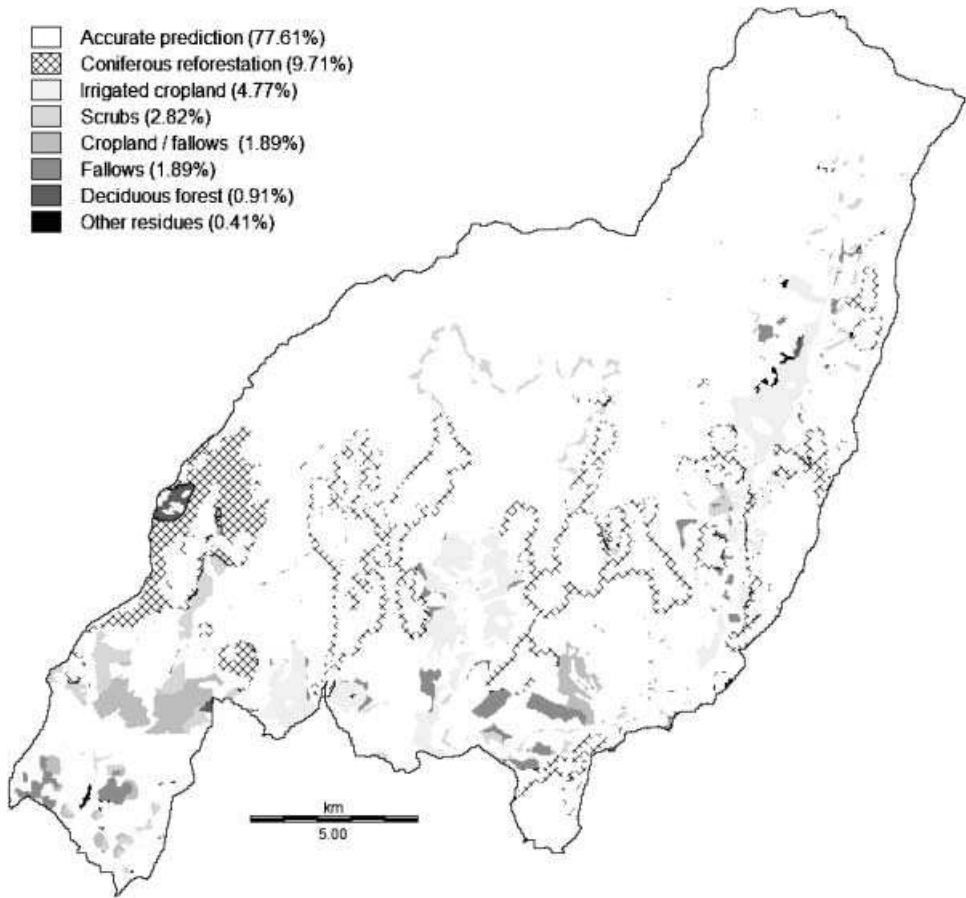
Figure 15.   Accurate prediction and typology of major prediction failures, Alta Alpujarra Granadina.

environmental conditions change quickly so that the rhythm of traditional, fine resolution, data supports like standard aerial photographs do not match the rhythm of land-use changes. The model calibration on two different test areas illustrates the difficulty to predict correctly land cover dynamics using fine resolution data. Data availability and uncertainty, random-side affects and the open status of complex systems are limiting factors to model land cover.

   Concerning the used methodology for prospective land cover modelling, it seems that complementing automatic modelling tools with manual supervised tools improves model results as shown for suitability and time transition probabilities for scrubs and grassland in eastern Pyrenees or fallows in the Spanish study area. A limit of the model is the fact that multi-criteria evaluation, multi-objective land allocation and cellular automata only influence the spatial distribution of predicted scores by Markovian chain analysis. LUCC budgets (cf. tables 9 and 10) emphasize that the principal prediction failure is the incorrect amount of predicted land cover change. Finally, each model is affected by random noise. Random side effects like forest fire and wind fall, even if they are not significant (less than 2% of the whole area), seem impossible to model correctly.

We insist that the model is based on a small number of easily available data so that the application to other areas is easy to perform by using existent and GIS implemented tools. To improve the model some steps are under work:

- Application to a third test area;
- Critical review of used Markov chain analysis;
- Comparing and combining the GIS based model with statistic approaches (nonlinear parametric model and neural networks) using the same data sets;
- Considering intra-class variance by using semi-quantitative land cover data (covering rates);
- Increasing use of validation tools.

References

BISHOP, C.M., 1995, *Neural Networks for pattern recognition* (New York: Oxford University Press).

CAMACHO OLMEDO, M.T., 2003, Cambios en el paisaje de la Alta Alpujarra granadina: determinación de los ritmos temporales y principales dinámicas con ayuda de un Sistema de Información Geográfica, in: *Geografías para una sociedad global: diversidad, identidad y exclusión social* AGE. Universidad Autónoma de Barcelona.

CAMACHO OLMEDO, M.T., GARCÍA MARTÍNEZ, P., JIMENÉZ OLIVENCIA, Y., MENOR TORIBIO, J. and PANIZA CABRERA, A., 2002a, Dinámica evolutiva del paisaje vegetal de la Alta Alpujarra granadina en la segunda mitad del siglo XX, *Cuadernos Geográficos de la University of Granada*, **32**, pp. 25–42.

CAMACHO OLMEDO, M.T., GARCÍA MARTÍNEZ, P., JIMENÉZ OLIVENCIA, Y., MENOR TORIBIO, J. and PANIZA CABRERA, A., 2002b, La transformación de un paisaje de montaña: el proceso de abandono de la agricultura en la Alta Alpujarra granadina. In: *Los espacios rurales entre el hoy y el mañana—XI Coloquio de Geografía Rural*. Santander, AGE. University of Cantabria, pp. 547–559.

CAMACHO OLMEDO, M.T., GARCÍA MARTÍNEZ, P., JIMENÉZ OLIVENCIA, Y., MENOR TORIBIO, J. and PANIZA CABRERA, A., 2002c, La Alta Alpujarra granadina en la segunda mitad del siglo XX a través de la cartografía evolutiva de su paisaje: Dinámica vegetal y repoblación forestal. In: *Los espacios rurales entre el hoy y el mañana—XI Coloquio de Geografía Rural*. Santander, AGE. University of Cantabria, pp. 535–547.

EASTMAN, J.R., 2001, *Idrisi32 release 2 Tutorial* (Worcester: Clarklabs).

EASTMAN, J.R., KYEM, P.A.K., TOLEDANO, J. and JIN, W., 1993, *Explorations in Geographic Systems Technology Volume 4: GIS and Decision Making* (Geneva: UNITAR).

FLAMM, R.O. and TURNER, M.G., 1994, Alternative model formulations for stochastic simulation of landscape change, *Landscape Ecology*, **9**(1), pp. 37–44.

GARCÍA MARTÍNEZ, P., 1999, *La transformación del paisaje y la economía rural en la Alta Alpujarra Occidental*, Granada: Editorial de la Universidad de Granada.

HORNIK, K., 1993, Some new results on neuronal network approximation, *Neural Networks*, **6**(8), pp. 1069–1072.

KOOPERBERG, C., BOSE, S. and STONE, J., 1997, Polychotomous Regression, *Journal of the American Statistical Association*, **92**, pp. 117–127.

MEZZADRI-CENTENO, T., 1998, *La modélisation et la projection spatio-temporelle dans les SIG* (Toulouse: Thèse de Doctorat. IRIT–UPS).

PAEGELOW, M. and CAMACHO OLMEDO, M.T., 2003, Le processus d'abandon des cultures et la dynamique de reconquête végétale en milieu montagnard méditerranéen: L'exemple des Garrotxes (P.O., France) et de la Alta Alpujarra Granadina (Sierra Nevada, Espagne), *Sud Ouest Européen*, **16**, pp. 113–130.

PAEGELOW, M., CAMACHO OLMEDO, M.T. and MENOR TORIBIO, J., 2002, Modelización prospectiva del paisaje mediante Sistemas de Información Geográfica, *X Congreso de Métodos Cuantitativos, Sistemas de Información Geográfica y Teledetección*. AGE. Universidad de Valladolid, CD-ROM.

PARLITZ, U. and MERKWIRTH, C., 2000, Nonlinear prediction of spatial-temporal time series, In Proceedings of ESANN'2000, Bruges, pp. 317–322.

PONTIUS, R.G. and SCHNEIDER, L.C., 2001, Land-cover change model validation by an ROC method for the Ipswich watershed, Massachusetts, USA, *Agriculture, Ecosystems & Environment*, **85**, pp. 239–248.

PONTIUS, R.G., SHUSAS, E. and MCEACHERN, M., 2004, Detecting important categorical land changes while accounting for persistence, *Agriculture, Ecosystems & Environment*, **101**, pp. 251–268.

SAATY, T.L., 1977, A Scaling Method for Priorities in Hierarchical Structures, *Journal of Mathematical Psychology*, **15**, pp. 234–28.

SELLERON, G. and MEZZADRI-CENTENO, T., 2002, Télédétection et logique floue: diagnostic et prospections temporelles de la déforestation sur un front pionnier tropical, *Societe. Française de Photogrammétrie et Télédétection*, **167**, pp. 4–15.

YAGER, R.R., 1988, On Ordered Weighted Averaging aggregation operators in multicriteria decision making, *IEEE Transactions on Systems, Man, and Cybernetics*, **8**(1), pp. 183–190.

ZADEH, L.A., 1988, Fuzzy Logic. *Computer*, **21**, pp. 83–93.