



Assessment of a Markov logic model of crop rotations for early crop mapping

Julien Osman, Jordi Inglada, Jean-François Dejoux

► **To cite this version:**

Julien Osman, Jordi Inglada, Jean-François Dejoux. Assessment of a Markov logic model of crop rotations for early crop mapping. *Computers and Electronics in Agriculture*, Elsevier, 2015, 113, pp.02.015. <10.1016/j.compag.2015.02.015>. <hal-01134729>

HAL Id: hal-01134729

<https://hal.archives-ouvertes.fr/hal-01134729>

Submitted on 30 Mar 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

1 Assessment of a Markov logic model of crop rotations
2 for early crop mapping

Email address: jordi.inglada@cesbio.eu (Jordi Inglada)

Preprint submitted to Elsevier

March 30, 2015

3 Assessment of a Markov logic model of crop rotations
4 for early crop mapping

5 Julien Osman^a, Jordi Inglada^{a,*}, Jean-François Dejoux^a

6 ^a*CESBIO - UMR 5126*
7 *18 avenue Edouard Belin*
8 *31401 TOULOUSE CEDEX 9 - France*

9 **Abstract**

Detailed and timely information on crop area, production and yield is important for the assessment of environmental impacts of agriculture, for the monitoring of the land use and management practices, and for food security early warning systems. A machine learning approach is proposed to model crop rotations which can predict with good accuracy, at the beginning of the agricultural season, the crops most likely to be present in a given field using the crop sequence of the previous 3 to 5 years. The approach is able to learn from data and to integrate expert knowledge represented as first-order logic rules. Its accuracy is assessed using the French Land Parcel Information System implemented in the frame of the EU's Common Agricultural Policy. This assessment is done using different settings in terms of temporal depth and spatial generalization coverage. The obtained results show that the proposed approach is able to predict the crop type of each field, before the beginning of the crop season, with an accuracy as high as 60%, which is better than the results obtained with current approaches based on remote sensing imagery.

10 *Keywords:* early crop type mapping, crop rotations, Markov Logic Networks

**Corresponding author*
Email address: jordi.inglada@cesbio.eu (Jordi Inglada)

11 **1. Introduction**

12 Detailed and timely information on crop area, production and yield is im-
13 portant for the assessment of environmental impacts of agriculture (Tilman,
14 1999), for the monitoring of the land use and management practices, and for
15 food security early warning systems (Gebbers and Adamchuk, 2010). Yield
16 production can be forecasted using models which need information about the
17 surface covered by each type of crop (Resop et al., 2012).

18 There are different ways of gathering this information, such as statistical
19 surveys or automatic mapping using Earth observation remote sensing imagery.
20 Statistical surveys are expensive to implement, since they need field work, which
21 is time consuming when large areas need to be covered. The use of remote sens-
22 ing imagery has been found to produce good quality maps when using high
23 resolution satellite image time series (Inglada and Garrigues, 2010). These ap-
24 proaches use supervised classification techniques which efficiently exploit satel-
25 lite image time series acquired during the agricultural season. Describing the
26 approach used for the supervised classification of satellite images is beyond the
27 scope of this paper and the details can be found in (Inglada and Garrigues,
28 2010), (Petitjean et al., 2012) or (Petitjean et al., 2014).

29 As an example of these approaches, figure 1 presents a 5-class crop map
30 obtained using a time series of 13 images acquired by the Formosat-2 satellite
31 during 2009 over a study site near Toulouse in Southern France. The data set
32 is described in Osman et al. (2012). The supervised classification is performed
33 using a Support Vector Machine as described in Inglada and Garrigues (2010).

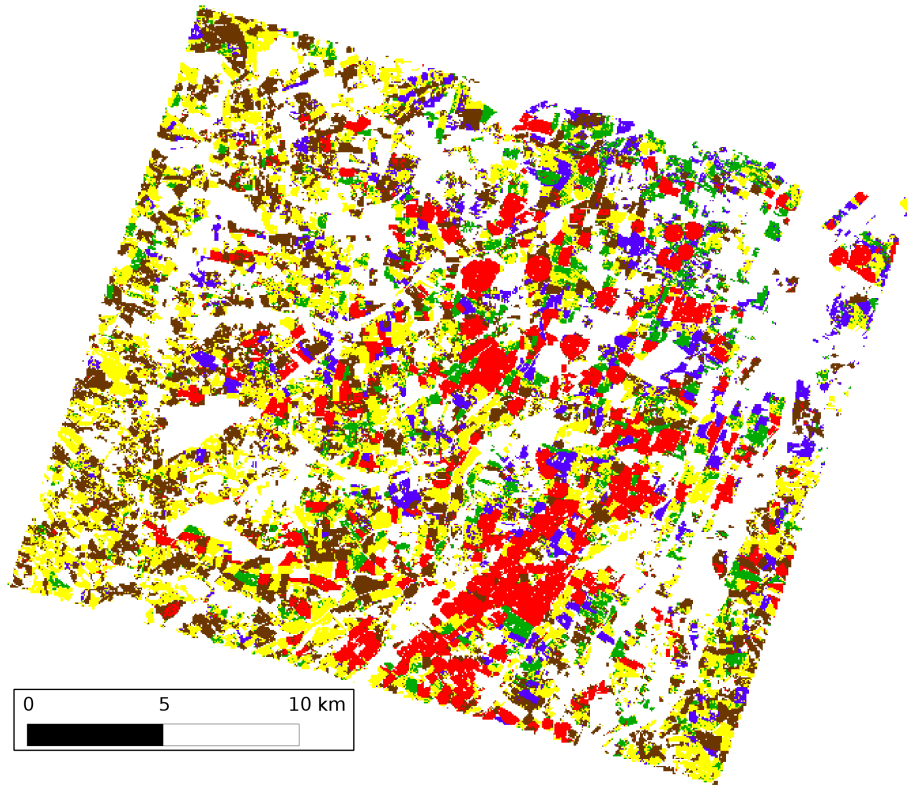


Figure 1: Example of crop map obtained by supervised classification of satellite image time series. Only croplands are classified. Corn (red), wheat (yellow), rapeseed (purple), barley (green), sunflower (brown). White areas represent non croplands.

34 The resulting classification has an accuracy close to 90%. However, this accuracy
35 can only be achieved at the end of the agricultural season when all images
36 are available. This delay in crop map production has led the remote sensing
37 community to develop near-real-time approaches, where the maps are updated
38 during the season every time a new image is available. Figure 2 shows the
39 evolution of the accuracy of each map produced during the season. A point in
40 the curve represents the accuracy obtained using all the images available up to
41 a given date. In this particular example, one can observe that a quality close

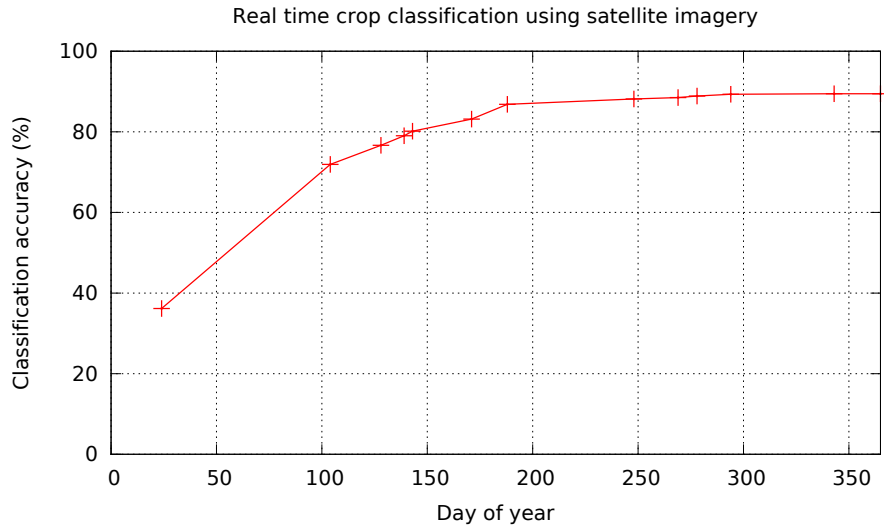


Figure 2: Classification accuracy obtained with satellite image time series. Each cross represents a new image acquisition. The accuracy increases when more images are available.

42 to the maximum can be obtained before 200 days into the year. However, no
 43 information is available before the first image is acquired at the end of January.
 44 For many crop systems, the beginning of the season coincides with the end of
 45 Autumn or the beginning of Winter. In this period, satellite images are very
 46 likely to be cloudy and therefore of little use for crop mapping. Furthermore,
 47 the accuracy of the land cover classification obtained with only one image is
 48 below 40%, which is not enough for most applications.

49 The goal of this paper is to introduce an approach which is able to produce
 50 land cover maps for agricultural areas at the beginning of the crop season with-
 51 out relying on remote sensing imagery. We propose to use the knowledge about
 52 the crop type which was present in every field the previous seasons to predict the
 53 crop grown the current year. The proposed approach uses a statistical model

54 for crop rotations.

55 Crop rotations – specific sequences of crops in successive years – improve or
56 maintain crop yield while reducing input demands for fertilizers and pesticides,
57 and therefore they are widely used by farmers. This regularity on the agricul-
58 tural practices allows predicting with some accuracy the type of crop present in
59 a given field at one point in time if the previous crop sequence is known.

60 Many crop rotation models exist, ranging from purely agronomic (crop-soil
61 simulation models ([Wechsung et al., 2000](#))), to approaches integrating expert
62 knowledge and field data ([Dogliotti et al., 2003](#)). The complexity of these models
63 makes them difficult to adapt to variable situations and evolving conditions.
64 Crop rotations may evolve in time, either slowly due to for instance climate
65 change impact in rain-fed crops, or very quickly due to environmental regulations
66 dealing with the use of pesticides or water management. Economic factors, as
67 for instance seed prices, can also introduce drastic changes. Hence, crop rotation
68 models which can be easily updated and which can exploit the history of the
69 different territories are needed.

70 Yearly cropland mapping can be obtained either using farmers administra-
71 tive declarations or maps produced using remote sensing data at the end of the
72 season (like the one of [figure 1](#)). Therefore, the history of the fields can be
73 known.

74 We propose a machine learning approach to model crop rotations which can
75 predict, at the beginning of a season, with good accuracy, the crops the most
76 likely to be present in a given field, using the crop sequence of the previous 3

77 to 5 years.

78 We assess its accuracy using the French Land Parcel Information System
79 RPG in different settings in terms of temporal depth and spatial generalization
80 coverage.

81 The paper is organized as follows. In section 2, we review several approaches
82 for crop rotation modeling in the literature. Section 3 presents the proposed
83 approach. In section 4, we present the type of data on which our approach
84 relies and we define the experimental setup used for this work; then, we present
85 the details of the assessment and analyze the results. The paper ends with a
86 conclusion and some perspectives.

87 **2. Modeling crop sequences**

88 The predictive model presented in this work (section 3) aims at providing
89 a first guess of crop type maps before the beginning of the crop season. Our
90 model uses knowledge about crop rotations.

91 Crop rotations have been intensively studied by both agronomists and economists
92 leading to farm management models in the economics and life sciences models
93 in agronomy. Some of them are presented in section 2.1.1. They often require
94 inputs of sequences of crops grown on a specific field over several years. In
95 recent years, there has been an increased focus on sustainable farming systems.
96 This has led to an increase in the use of farm models used to assess the environ-
97 mental impact of farming. In models of complete exploitations including crop
98 production, it is important to consider the rotation of crops, since this has a
99 major impact on the consequences of the crop production.

100 However, for the forecasting of crop type mapping, there are specific needs
101 which are not covered by existing modeling approaches. These specific needs
102 are:

- 103 1. Field level information: the crop type has to be predicted for every indi-
104 vidual field; aggregate data or regional trends are not enough.
- 105 2. Different landscapes and different climatic conditions lead to different
106 management practices. Therefore, regional information has to be com-
107 bined with field-level history.
- 108 3. The approach should be portable to different countries and regions of the
109 globe with minimum adaptations. Therefore, it should be able to both,
110 learn from data and to exploit *expert knowledge*. The approach should
111 also be able to use only one of these 2 types of information in case the
112 other one is not available.
- 113 4. To cover very large areas, the approach must not rely exclusively on field
114 surveys which are expensive in terms of time and manpower.
- 115 5. The model should be able to evolve in time to take into account changing
116 conditions which influence managing practices (climate change, regulatory
117 constraints).

118 To the best of our knowledge, no existing approach in the literature allows
119 fulfilling all these requirements.

120 *2.1. Existing approaches in the literature*

121 Crop rotation modeling has been addressed in different ways. We may clas-
122 sify these approaches in 2 groups:

- 123 1. the approaches using mainly theoretical knowledge, that is models from
124 life sciences, economics or using expert knowledge by agronomists;
- 125 2. the approaches which learn from data.

126 *2.1.1. Theoretical knowledge*

127 One simple example of theoretical knowledge is the ROTAT software tool
128 ([Dogliotti et al., 2003](#)) which generates all possible rotations of the crops present
129 in a particular area, and then applies a selection based on agronomic criteria
130 provided by experts. This approach allows producing accurate results at the
131 exploitation level, but not at the field level.

132 The creation of transition matrices adapted to the agricultural landscape
133 under study requires expert knowledge on the type of crop rotation to model
134 and an understanding of the internal dynamics of crop successions. Such knowl-
135 edge may be derived from research on decision-making by farmers about crop
136 succession ([Castellazzi et al., 2008](#)). Castellazzi et al. use Markov chains with
137 transition probabilities set by experts, but their values are limited to 0 and 1.

138 The specialization of the models to particular sites needs adequate tools.
139 For example [Detlefsen et al. \(Detlefsen and Jensen, 2007\)](#) propose the use of
140 network modeling to find an optimal crop rotation for a given selection of crops
141 on a given piece of land. This model can give advice about the appropriate
142 crop to be grown on a field, but it needs information about the farm (surface,
143 number of fields) and about the costs of farming operations (ploughing, etc.).
144 This kind of information may not be available when mapping very large areas.

145 Farm management models often produce average crop shares over a num-

146 ber of years, whereas models from the natural sciences often require inputs of
147 sequences of crops grown on a specific field over several years.

148 For instance, the SWIM model used by Wechsung et al. ([Wechsung et al.,](#)
149 [2000](#)) can not be applied efficiently over large areas at the individual field level,
150 since it needs very detailed information about specific parameters of the crops.
151 The works of Klöcking et al. ([Klöcking et al., 2003](#)) or Salmon-Monviola et al.
152 ([Salmon-Monviola et al., 2012](#)) fall in the category of models which perform
153 stochastic simulations for scenarios, but not for accurate mapping at the field
154 level.

155 In interdisciplinary modeling, this difference can be an obstacle. To bridge
156 this gap, an approach is presented in ([Aurbacher and Dabbert, 2011](#)) that allows
157 disaggregating results from farm management models to the level required by
158 many natural science models. This spatial disaggregation consists in deriving a
159 spatial distribution of some information which is only available as a summary for
160 a large area. Aurbacher et al. ([Aurbacher and Dabbert, 2011](#)) use Markov chains
161 for the disaggregation at the field level. This approach needs detailed knowledge
162 about the activity at the field and farm levels, as well as other economical
163 information as for instance gross margin. This level of detail is difficult to
164 obtain for large areas and therefore the approach is not suited to mapping.

165 The integration of many types of knowledge is challenging, and one of the
166 approaches for overcoming this difficulty is to use multi-agent systems, as for
167 instance in the Maelia platform ([Taillandier et al., 2011](#)). This approach suffers
168 from the same drawbacks as the previous ones: the need to access detailed

169 knowledge at the farm level.

170 The main drawback of models based on theoretical knowledge is their in-
171 ability to easily adapt to changing conditions, since these new conditions have
172 to be accounted for in the models, or adaptive decision rules have to be imple-
173 mented. However, some attempts have been made to take into account changes.
174 For instance, Supit et al. (Supit et al., 2012) model climate change impacts on
175 potential and rain-fed crop yields on the European continent using the outputs
176 of three General Circulation Models in combination with a weather generator.
177 However, this model is only able to evolve with respect to climate and not with
178 respect to other types of changes.

179 *2.1.2. Automatic learning from data*

180 One way to overcome the problem of adaptation to changing environments
181 or to specific areas, is to integrate field surveys or similar data in the models.

182 There are models which are used to describe existing data, as for instance
183 CarrotAge (Le Ber et al., 2006), which allows analyzing spatio-temporal data
184 to study the cropping patterns of a territory. The results of CarrotAge are
185 interpreted by agronomists and used in research works linking agricultural land
186 use and water management. The underlying algorithms use Markov models.
187 The main limitation of CarrotAge for our needs is that it does not perform crop
188 prediction at the field level.

189 Another example is the crop rotation model CropRota (Schönhart et al.,
190 2011), which integrates agronomic criteria and observed land use data to gen-
191 erate typical crop rotations for farms and regions. CropRota does not work at

192 the field level.

193 Similar to the previous one, ROTOR ([Bachinger and Zander, 2007](#)) is a tool
194 for generating and evaluating crop rotations for organic farming systems. It was
195 developed using data from field experiments, farm trials and surveys and expert
196 knowledge. Its originality is the integration of a soil–crop simulation model. As
197 the two previous approaches, ROTOR does not perform predictions at the field
198 level.

199 As our goal is to map the croplands, we need not only to model the transi-
200 tions of crops, but also to take into account the geospatial information available.

201 Usually, the data available for integration in models comes from census and
202 has no continuous spatial distribution. Many approaches for the spatialization
203 of this kind of information exist, as for instance krigging ([Flatman and Yfantis,](#)
204 [1984](#)). In the case of crop distribution, You et al. ([You et al., 2006](#)) proposed
205 an approach to go from census data to raster information, but their work is not
206 applied to the field level, which is needed in our case for crop mapping.

207 Although limited to 3 crops, Xiao et al. ([Xiao et al., 2014](#)) used field level
208 information to perform a regional scale analysis, but they did not perform fore-
209 casting of the selected crops in the individual fields.

210 Among the cited approaches, none of them fulfill the 5 constrains listed
211 at the beginning of this section. However, some of these works have shown
212 that statistical modeling of crop rotations in general, and Markovian models in
213 particular are appropriate tools for crop type prediction. The drawback of the
214 Markovian approaches used in the literature is that they are not easily updated

215 when expert knowledge complementary to existing data is available.

216 **3. Modeling with Markov Logic**

217 We start (section 3.1) by justifying the use of Markovian approaches for crop
218 rotation modeling and we point out their main limitation for our needs : the
219 impossibility of easily integrating expert knowledge. We then present in section
220 3.2 the Markov Logic approach which solves this issue. Finally, in section 3.3
221 we describe how to use Markov Logic Networks to model crop rotations and to
222 forecast future crops.

223 *3.1. Properties of the model*

224 At the beginning of section 2, the specific needs for the forecasting of crops
225 at the field level were listed. After the literature review on crop rotation mod-
226 els, the properties that a model for our application should possess can now be
227 precised.

228 1. Learning from past sequences, both at the field and at the regional scale.

229 This allows taking into account regional trends together with specific field
230 information.

231 2. Exploiting the past information for every particular field (either using
232 Land Parcel Information Systems or existing land-cover maps).

233 3. Incorporating changes in practices without needing the compilation of new
234 data bases containing examples of these evolutions. This allows the model
235 to quickly evolve without the need of a time lag before being able to exploit
236 information about changing conditions.

237 As we saw in section 2, existing approaches to assess agricultural practices
238 focus on the assessment of single crops or statistical data because spatially
239 explicit information on practically applied crop rotations was lacking (Lorenz
240 et al., 2013), but this is not the case anymore in the EU. For instance Leteinturier et al. (Leteinturier et al., 2006) used the land parcel management system
241 implemented in the frame of EU's Common Agricultural Policy to assess many
242 common rotation types from an agro-environmental perspective. Also, in the
243 USA, the USDA's Cropland Data Layer provides annual crop cover data at 30
244 m. resolution (Boryan et al., 2011).

246 When learning from data representing sequential states of variables, the
247 Markovian properties are often used. In a Markovian process, the next state de-
248 pends only on the current state and not on the sequence of events that preceded
249 it. This allows to efficiently learn the probability of any particular sequence of
250 states by computing only the probability of transition between individual states.
251 As a matter of fact, most of the approaches similar to those presented in section
252 2.1.2 use these approaches.

253 One of the most frequently used Markovian models are Bayesian Networks
254 (BN) (Friedman and Koller, 2003; Heckerman et al., 1995) which are today one
255 of the most promising approaches to Data Mining and Knowledge Discovery in
256 databases. A BN is a graph (structure of the network) where each node is a
257 random variable (for instance the crop grown on a particular field on a given
258 year) and each edge represents the degree of dependence between the random
259 variables (the probability of transition between states). Figure 3 illustrates some

260 examples of BN.

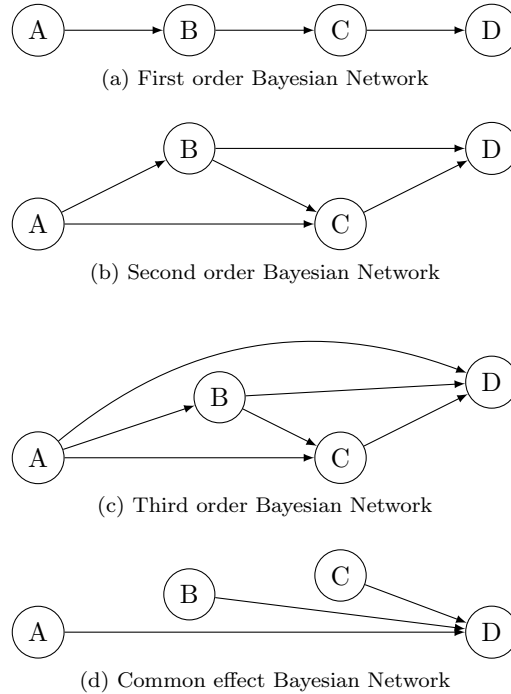


Figure 3: Examples of Bayesian Networks.

261 A Markov Random Field, MRF, (or Markovian Network, MN) is similar
262 to a BN in its representation of dependencies ([Kindermann et al., 1980](#)); the
263 differences being that BN are directed and acyclic, whereas MN are undirected
264 and may be cyclic. Thus, a MN can represent certain dependencies that a BN
265 cannot (such as cyclic dependencies); on the other hand, it cannot represent
266 certain dependencies that a BN can (such as induced dependencies).

267 BN and MRF need probability estimates which can be learnt from data.
268 However, they cannot easily incorporate other types of knowledge as for in-
269 stance logic rules. For instance, in the case of crop rotations, a new regulation
270 about nitrates can change the patterns of the sequences. Changes in prices or a

271 reorientation towards bio-fuel production can lead to yet bigger changes. These
272 expected changes can be expressed with rules, but no data is available for learn-
273 ing until several agricultural seasons have passed. Furthermore, in some cases,
274 the knowledge is easier to express in terms of a set of sentences or formulas in
275 first-order logic (*if-then* rules), rather than in terms of transition probabilities
276 between states. Therefore, an alternative or an extension to BN and MRF is
277 needed.

278 3.2. Markov Logic

279 To combine knowledge from databases and knowledge from experts, inference
280 approaches which are able to combine probabilistic learning and rule-based logic
281 reasoning are needed. Combining probability and first-order logic in a single rep-
282 resentation has long been a goal of Artificial Intelligence. Probabilistic graphical
283 models like BN make it possible to efficiently handle uncertainty. First-order
284 logic allows to compactly represent a wide variety of knowledge. The combi-
285 nation of probabilistic and propositional models has been one research area of
286 important activity since the mid 1990's (Cussens, 2001; Puech and Muggleton,
287 2003).

288 Recently, Markov Logic (ML) (Richardson and Domingos, 2006) was in-
289 troduced as a simple approach to combining first-order logic and probabilistic
290 graphical models in a single representation. A Markov Logic Network (MLN)
291 is a first-order knowledge base (KB) with a weight attached to each formula¹.
292 Together with a set of constants representing objects in the domain, it specifies

¹Logic formulas are also called rules or clauses.

293 a ground MN² containing one feature for each possible grounding of a first-order
294 formula in the KB, with the corresponding weight. Inference in MLNs is per-
295 formed by Monte Carlo Markov Chains (MCMC) over the minimal subset of the
296 ground network required for answering the query. Weights are efficiently learned
297 from relational databases by iteratively optimizing a pseudo-likelihood measure.
298 Optionally, additional clauses are learned using inductive logic programming
299 techniques. Also, clauses can be added if some prior or expert knowledge is
300 available.

301 A first-order logic KB can be seen as a set of hard constraints on the set of
302 possible worlds: if a world does not respect one single formula, it has zero prob-
303 ability. In MLN, these constraints are softened: if a world does not verify one
304 formula in the KB it has a lower probability, but not zero. The more formulas a
305 world respects, the more probable it is. Each formula has an associated weight
306 that reflects how strong a constraint is: the higher the weight, the greater the
307 difference in probability between a world that satisfies the formula and one that
308 does not. The weights are not limited in range as probability values are.

309 Models like MRF and BN can still be represented compactly by MLNs, by
310 defining formulas for the corresponding factors.

311 Efficient algorithms for learning the structure of the networks and the weights
312 associated to the rules exist ([Singla and Domingos, 2005](#)) and they are made
313 available by the authors as a free and open source software implementation ([Kok](#)

²A ground MN is a MN without free variables in the logic formulas. It is also usually referred as a *possible world*.

314 [et al., 2006](#)) which makes possible the assessment of the approach for our needs.

315 *3.3. The proposed approach*

316 We propose to model each rotation of interest as one rule and use a MLN
317 for the inference. Therefore, the rules do not need to be learned, but only their
318 weights. Using data for a set of years, the weights of each rule are learned. The
319 approach is validated by applying the inference.

320 The crops of interest for our experiments are wheat, barley, corn, rapeseed
321 and sunflower, which represent 78% of the surface in the study area. The rules
322 are expressed as follows in the case of a 4 year rotation cycle:

$$\{C_{n-3}^a, C_{n-2}^b, C_{n-1}^c\} \rightarrow C_n^d, \omega,$$

which means that the rule which says that a sequence of crop a , followed by
crop b , followed by crop c leads to crop d the following year has a weight ω . The
notation can be simplified as

$$\{a, b, c, d, \omega\}.$$

323 The weights ω have to be learned for each possible sequence of crops that has
324 to be modeled. This type of rules corresponds to the same kind of dependency
325 which can be modeled by a common effect BN (figure 3d).

326 **4. Experiments and results**

327 *4.1. Description of the available data and the area of study*

328 *4.1.1. The French RPG LPIS*

329 The information about the crop rotation used for the assessment of the model
330 was obtained from the *Registre Parcellaire Graphique*, RPG, a topographical

331 Land Parcel Information System (LPIS) containing the agricultural parcels and
332 the corresponding crops grown.

333 At the national French level, it contains about 7 million parcels. The system
334 was implemented in 2002 in application of EU directives. It is annually updated
335 by farmers themselves. The information of interest associated to each parcel is:

- 336 • the geographical outline of the parcel and an identifier;
- 337 • the district where the parcel is located;
- 338 • the type of the crop grown a particular year using a 28 class nomenclature;
- 339 • the administrative type of the exploitation;
- 340 • the age class of the owner for individual owners.

341 One particularity of the RPG is that the parcels may correspond either to
342 individual fields or to groups of small fields. These groups may be composed by
343 fields where different crops are present. In these cases, the spatial distribution
344 is not given and only the proportion of each crop surface is known.

345 For the experiments presented here, only individual fields where a single
346 crop is grown were used. This made the analysis easier and the amount of
347 data remained sufficient for the statistical approach to be robust. However, a
348 statistical bias might appear because of the use of a subset of the fields. To
349 solve this issue, techniques have been proposed for the estimation of the spatial
350 distribution of the crops within a group of fields ([Inglada et al., 2012](#)) and they
351 could be used in the future.

352 It is also worth noting that the RPG was used here to have access to a
353 very large geographical area during several years and assess the properties of
354 the proposed model, but other sources of data, as for instance land-cover maps
355 from previous years as the one illustrated in figure 1, could be used without loss
356 of generality.

357 *4.1.2. Study area and time frame*

358 For our study, we used 7 years of data (2006-2012) over a large region in the
359 South of France (figure 4). This amount of data allowed us to assess the model
360 in terms of temporal stability, temporal depth of the rotations as well as spatial
361 homogeneity of the areas.

362 We used 3 areas of study which are depicted on figure 4:

- 363 1. a small area of 20 km \times 20 km (red rectangle) which has rather homoge-
364 neous pedo-climatic conditions with about 1700 parcels studied;
- 365 2. a medium sized region (dark gray area including the small area) with
366 about 15500 parcels studied and where soils have different types and a
367 sensible North-South climatic gradient is present;
- 368 3. a large sized area (light gray area plus the 2 previous ones) with about
369 72000 parcels studied and presenting a wide variety of soils, landscapes
370 and climatic conditions.

371 *4.2. Experimental setup*

372 *4.2.1. Assessment*

373 To assess the capabilities of MLN to give useful information for forecasting
374 the grown crops at the field level, we used the data base presented in section

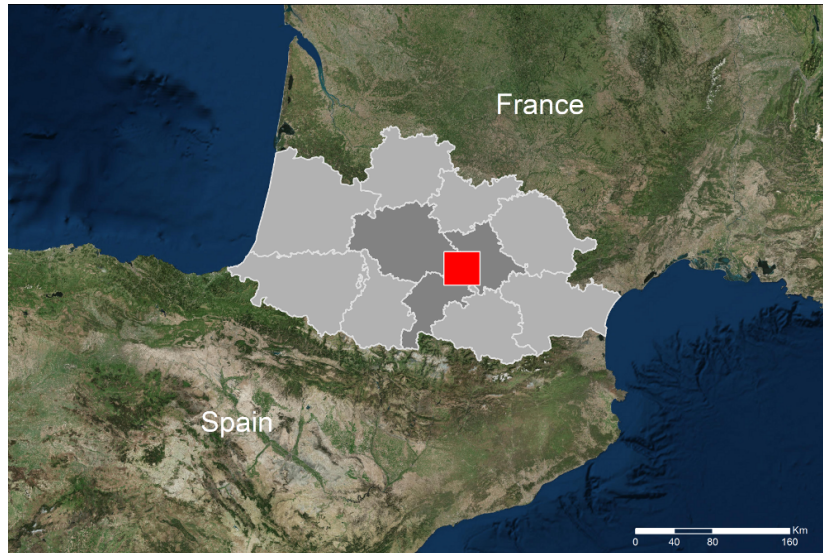


Figure 4: The 3 study regions: in red a $20\text{ km} \times 20\text{ km}$ area (small), in dark gray the medium area and in light gray the large area.

375 4.1. We studied the influence of the length of the considered rotations as well
 376 as the extent of the area over which the modeling was performed.

377 To assess the influence of the rotation length, we analyzed 3 different cases:
 378 4 year rotations (that is knowledge of the previous 3 years to forecast the forth
 379 one), 5 year rotations and 6 year rotations.

380 Finally, to assess the impact of the extent of the area (eco-climatic conditions,
 381 pedology, etc.), we used the 3 regions presented in figure 4.

382 4.2.2. Evaluation

383 To evaluate the quality of the crop prediction, classical tools from the ma-
 384 chine learning field were used: the confusion matrix and the Kappa coefficient.

385 The confusion matrix (also known as contingency table) is a double entry
 386 table where row entries are the actual classes (crop in the reference data) and

387 column entries are the predicted classes. Each cell of the table contains the
388 number of elements of the row class predicted by the classifier as belonging to
389 the column class.

390 The diagonal elements in the matrix represent the number of correctly pre-
391 dicted individuals of each class, i.e. the number of ground truth (reference)
392 individuals with a certain class label that actually obtained the same class label
393 during prediction.

394 The off-diagonal elements represent misclassified individuals or the classi-
395 fication errors, i.e. the number of ground truth individuals that ended up in
396 another class during classification.

397 Part of the agreement between the classifier's output and the reference data
398 can be due to chance. The Kappa coefficient (κ) expresses a relative difference
399 between the observed agreement P_o and the random agreement which can be
400 expected if the classifier was random, P_e .

$$\kappa = \frac{P_o - P_e}{1 - P_e}$$

where

$$P_o = \frac{1}{n} \sum_{i=1}^r n_{ii}$$

is the agreement and

$$P_e = \frac{1}{n^2} \sum_{i=1}^r n_{i.} n_{.i}$$

401 κ is a real number between -1 and 1 and can be interpreted as follows:

	Agreement	κ
	Excellent	> 0.81
	Good	$0.80 - 0.61$
402	Moderate	$0.60 - 0.41$
	Weak	$0.40 - 0.21$
	Bad	$0.20 - 0.0$
	Very bad	< 0

403 4.3. Assessment of the proposed approach

404 4.3.1. Examples of obtained rotations

405 To give the reader a sense of the difference between crop rotation frequency
406 and the knowledge modeled by the MLN, the 20 most frequent rotations in the
407 small study area for a 4 year cycle are presented in table 1, and the 20 rules
408 with the highest weights for the same area and the same period are presented
409 in table 2.

410 In terms of frequency of the rotations, the first thing we note is that the first
411 and the second rotations are the same with a shift of one year. It is interesting
412 to note that these 2 rotations have very high weights in table 2 and these weight
413 are not very different if we take into account that there is a 1.6 ratio in terms
414 of frequency. We can also see that the corn mono-culture is very frequent and
415 the corresponding rule has also a very high weight.

416 Looking at the first 3 rows of both tables, one may deduce that rule weights
417 yield similar information to frequency of occurrence of rotations. However, this

	2009	2010	2011	2012	number
1	sunflower	wheat	sunflower	wheat	405
2	wheat	sunflower	wheat	sunflower	253
3	corn	corn	corn	corn	113
4	sunflower	wheat	sunflower	barley	46
5	wheat	rapeseed	wheat	rapeseed	46
6	wheat	rapeseed	wheat	sunflower	46
7	wheat	sunflower	wheat	rapeseed	38
8	rapeseed	barley	wheat	rapeseed	34
9	rapeseed	wheat	sunflower	wheat	34
10	sunflower	wheat	rapeseed	wheat	26
11	rapeseed	wheat	rapeseed	wheat	26
12	barley	wheat	sunflower	wheat	26
13	barley	wheat	rapeseed	barley	24
14	sunflower	barley	wheat	sunflower	24
15	wheat	sunflower	barley	sunflower	22
16	sunflower	wheat	wheat	sunflower	21
17	wheat	sunflower	barley	wheat	21
18	barley	wheat	sunflower	barley	19
19	wheat	rapeseed	barley	wheat	18
20	barley	sunflower	wheat	sunflower	16

Table 1: Most frequent rotations in the small area with their corresponding number of occurrences.

418 is not the case, since the rules represent a conditional probability³ of the last
419 crop of the sequence with respect to the sequence of the 3 crops which precede
420 it. For instance, rules where corn is present appear in the table (limited to the
421 20 rules with the highest weights) even if corn is only present in one of the most
422 frequent sequences.

423 4.3.2. Overview of the behavior

424 With the data set used, there were 27 different combinations in terms of
425 area, rotation length and particular sets of years. Tables 3, 4 and 5 give an

³Although weights are not restricted to the $[0 - 1]$ intervals as probabilities are. In the same way, the sum of all weights does not have to be 1 as with probabilities. This latter property allows introducing new knowledge not represented in the data when available.

	2009	2010	2011	2012	weight
1	sunflower	wheat	sunflower	wheat	0.752
2	corn	corn	corn	corn	0.699
3	wheat	sunflower	wheat	sunflower	0.601
4	wheat	barley	wheat	barley	0.355
5	corn	corn	rapeseed	barley	0.333
6	corn	corn	rapeseed	rapeseed	0.331
7	corn	rapeseed	corn	rapeseed	0.322
8	corn	corn	rapeseed	wheat	0.319
9	corn	rapeseed	corn	barley	0.317
10	corn	corn	rapeseed	sunflower	0.312
11	sunflower	wheat	barley	sunflower	0.309
12	rapeseed	corn	corn	rapeseed	0.305
13	corn	corn	barley	barley	0.305
14	wheat	barley	wheat	corn	0.304
15	rapeseed	corn	corn	barley	0.302
16	barley	wheat	sunflower	rapeseed	0.302
17	corn	rapeseed	corn	sunflower	0.3
18	corn	barley	corn	barley	0.3
19	wheat	barley	wheat	rapeseed	0.298
20	barley	wheat	sunflower	corn	0.297

Table 2: Higher weight rules in the small area with their corresponding weights ($\{a, b, c, d, \omega\}$).

Small region				
-	2009	2010	2011	2012
4 years	0.51	0.58	0.54	0.60
5 years	-	0.57	0.53	0.61
6 years	-	-	0.54	0.55

Table 3: κ coefficient values for the small region

426 overview of the results, in terms of κ coefficients, for the small, the medium and
 427 the large regions respectively.

428 The first observation we can make is that most of the κ values were in the
 429 high fifties, which is a moderate to good prediction of the crops. It is not
 430 surprising to note that the predictions for the small area were the best and
 431 those for the large area were the worse, since the eco-pedo-climatic conditions
 432 which govern agricultural practices are more homogeneous in the small area.
 433 However, the results of the medium area were very close to those of the small
 434 area.

435 In terms of rotation length, we can observe that 4 and 5 years were equivalent
 436 for the small and medium regions and that 6 years was worse than 5 which could
 437 be explained by the high number of rotations to model in the longer case (4096
 438 combinations with respect to 1024).

439 Finally, we can observe that the predictions for the year 2011 were the ones
 440 with the lower quality independently of the area and of the length of the rota-
 441 tions. This may be explained by the fact that 2009 suffered from an anomalous
 442 weather which forced many farmers in the South of France to change the planned
 443 winter wheat for a Summer crop like sorghum or sunflower. This modification
 444 of practices impacted the statistical representativity of the data.

Medium region				
-	2009	2010	2011	2012
4 years	0.53	0.57	0.51	0.58
5 years	-	0.57	0.52	0.59
6 years	-	-	0.51	0.54

Table 4: κ coefficient values for the medium region

Large region				
-	2009	2010	2011	2012
4 years	0.50	0.56	0.52	0.58
5 years	-	0.50	0.46	0.53
6 years	-	-	0.43	0.43

Table 5: κ coefficient values for the large region

445 In the following paragraphs, the details of the confusion matrices are ana-
 446 lyzed to gain some insight on the behavior of the model.

447 *4.3.3. Area*

448 We focused our interest on the differences of prediction quality between the
 449 different regions of different size. In order not to multiply the combinations, we
 450 used the results for the length of 5 years and analyzed the confusion matrices
 451 which resulted from the averaging the results of the predictions for 3 years (2010
 452 to 2012).

453 The confusion matrices for the small, the medium and the large areas are
 454 presented on tables 6, 7 and 8 respectively.

455 The first thing we can highlight is that there were no major differences
 456 between the small and the medium regions as it was already noted in the overall
 457 κ coefficient tables above. The confusion matrices allowed us to check that this
 458 stability was reproduced even at the level of the individual crops and their
 459 specific confusions.

-	Wheat	Corn	Barley	Rapeseed	Sunflower
Wheat	73	5	8	5	9
Corn	5	80	5	4	6
Barley	24	6	32	8	30
Rapeseed	7	5	12	29	46
Sunflower	15	17	28	21	20

Table 6: Confusion matrix for the small region

460 In terms of confusions, we can see that sunflower was the most difficult crop
461 to predict and more so when the area was very large. In this latter case, the
462 prediction accuracy was lower than random (which would be of 20%). During
463 the past decade, sunflower yields have been steadily decreasing in this region
464 and it is increasingly becoming an opportunity crop to use when the planned
465 winter crop could not be sowed.

466 At the opposite, wheat and corn were very well predicted and this was mostly
467 because they are the principal crops grown in the area. Rapeseed was much
468 confused with sunflower, since they are usually chosen for economic reasons
469 rather than for agronomic ones. We also see that barley was often predicted
470 as wheat, which is easy to explain because these 2 crops are both straw cereals
471 (and therefore interchangeable from the agronomic point of view) and as stated
472 before, wheat is the most prominent one of those 2. The confusion was stable
473 between areas, but barley was less well predicted when the area was larger
474 mainly because of increasing confusions with rapeseed. The good prediction of
475 corn remained stable independently of the size of the area.

-	Wheat	Corn	Barley	Rapeseed	Sunflower
Wheat	74	6	7	5	9
Corn	5	80	4	5	6
Barley	23	9	27	11	30
Rapeseed	7	7	11	26	49
Sunflower	18	18	24	22	18

Table 7: Confusion matrix for the medium region

-	Wheat	Corn	Barley	Rapeseed	Sunflower
Wheat	65	8	8	8	10
Corn	6	79	4	5	6
Barley	23	13	22	13	29
Rapeseed	11	10	13	21	45
Sunflower	19	20	24	22	16

Table 8: Confusion matrix for the large region

476 *4.3.4. Length*

477 We limited the study to the medium area and we analyzed the influence of
478 the length of the sequences used for the model (column 2012 of table 4). The
479 results are presented in tables 9, 10 and 11 for the rotations using 4, 5 and 6
480 years respectively.

481 The trends that we observe are the following:

- 482 • the longest sequences were the most difficult to predict, which is not sur-
483 prising, since the number of possible combinations is higher and therefore
484 the probability of each one is lower;
- 485 • the prediction of corn was good and stable for the different rotation
486 lengths, since most of the corn in the area is grown as mono-culture;
- 487 • the prediction of wheat was good but decreased with the length of the
488 sequence;

-	Wheat	Corn	Barley	Rapeseed	Sunflower
Wheat	83	4	3	3	7
Corn	5	75	5	4	11
Barley	27	5	12	13	43
Rapeseed	6	7	8	15	64
Sunflower	17	16	17	23	27

Table 9: Confusion matrix for a 4 year sequence.

-	Wheat	Corn	Barley	Rapeseed	Sunflower
Wheat	80	4	4	4	7
Corn	5	76	4	7	8
Barley	26	7	28	8	31
Rapeseed	6	8	13	24	49
Sunflower	16	15	25	23	21

Table 10: Confusion matrix for a 5 year sequence.

- 489 • rapeseed and sunflower were often confused and their respective prediction
490 accuracies had inverse trends: rapeseed benefited from longer sequences,
491 while sunflower was best predicted with shorter sequences;
- 492 • in the previous paragraphs, we observed an important amount of barley
493 being predicted as wheat, and we saw that this confusion diminished when
494 the areas were larger; here we see that this confusion was stable with
495 respect to the length of the sequence, however the prediction of barley
496 benefited from medium length sequences, mainly because the reduction of
497 the confusion with rapeseed.

498 4.3.5. *Simulating drastic changes*

499 In the previous experiments we showed the ability of MLN to predict the
500 crops knowing the past history of the fields. However, from the application
501 point of view, this kind of use is similar to the use of BN, the main advantage

-	Wheat	Corn	Barley	Rapeseed	Sunflower
Wheat	68	8	7	7	10
Corn	7	75	4	6	8
Barley	24	7	25	10	34
Rapeseed	7	7	12	33	41
Sunflower	19	19	15	27	20

Table 11: Confusion matrix for a 6 year sequence.

502 of MLN being the possibility to have straightforward access to human readable
503 rules instead of having a graphical model which is difficult to interpret when
504 there are many nodes.

505 However, the use of MLN was proposed because they are able to combine
506 statistical learning with first-order logic rules. This particular property of MLN
507 is interesting to introduce knowledge for which no historical data is available. In
508 the case of early crop mapping, this situation may happen due to new regulations
509 or economic reasons, like seed prices.

510 Unfortunately, this kind of behavior was not present in our data set, and
511 therefore, we chose to simulate it. The following experiment was carried out.
512 We assumed that for an arbitrary reason, one type of rotation which had been
513 frequent in the past became nearly non existent from a given point in time. We
514 introduced this expected behavior by strongly modifying the weight of the rule
515 related to this particular rotation. We then analyzed how the probability of the
516 crops to be predicted spread among the possible types of crops.

517 Of course, this kind of event is extreme and not likely to occur as such, but
518 it allowed illustrating the flexibility of the proposed approach.

519 For this experiment, we used the MLN obtained by performing the training

-	Original	Modified
Corn	0.60	0.0014
Wheat	0.11	0.28
Sunflower	0.11	0.28
Rapeseed	0.088	0.22
Barley	0.089	0.23

Table 12: Predicted probabilities for each crop for the rotation $\{C_{n-3}^{corn}, C_{n-2}^{corn}, C_{n-1}^{corn}\} \rightarrow C_n^d$ with the original weight and the modified one.

520 on the medium sized region and using the years from 2008 to 2011 (used to
521 predict the crops in 2012).

522 We chose the sequence $\{corn, corn, corn, corn\}$ whose weight was 0.699 and
523 modified it to have a weight of $-\infty$. It is interesting to note that only this
524 rule was modified. We then analyzed the predicted probability by the MLN for
525 different rotations in the case where we kept the original weight for the rule or
526 we used the modified weight.

527 Table 12 shows the predicted probability for class d on year n for the rules
528 $\{C_{n-3}^{corn}, C_{n-2}^{corn}, C_{n-1}^{corn}\} \rightarrow C_n^d$ for the original (learned from data) weight and
529 the modified one. As one can see, the original setting predicted *corn* with a
530 probability of 0.6, the other classes having a very low probability. In the case
531 where $\{corn, corn, corn, corn\}$ was nearly non existent, *corn* was predicted with
532 a probability which was practically zero, while the other classes were predicted
533 with similar probability, but those which previously had higher probabilities
534 (wheat and sunflower) still had higher chances than rapeseed and barley.

535 It is worth noting that no re-learning from the data had to be done, so this
536 kind of changes can be introduced in the model at no cost.

537 It was also necessary to check that the modification of a particular rule did
538 not have effect on other rules. To verify the correct behavior of the model, we
539 applied the same kind of analysis to other rules. In the case of one of the most
540 frequent rotations of the study area $\{sunflower, wheat, sunflower, wheat\}$,
541 which is described by the rules $\{C_{n-3}^{sunflower}, C_{n-2}^{wheat}, C_{n-1}^{sunflower}\} \rightarrow C_n^d$, there
542 was no modification of the probabilities after changing the weight of the rule
543 $\{corn, corn, corn, corn\}$.

544 The same behavior occurred for the set of rules $\{C_{n-3}^{wheat}, C_{n-2}^{barley}, C_{n-1}^{wheat}\} \rightarrow C_n^d$.
545 Finally, a family of rules containing 2 consecutive years of corn was not modified
546 either.

547 In the case of a BN, this modification would have required to modify the
548 training data and learn the transition probabilities again, since it is impossible
549 to modify the probability of a particular sequence of events without modifying
550 all the rest.

551 The point here is not that the probabilities of the other crops did not change.
552 In a realistic setting, the relative proportion of other crops may evolve due to
553 economic or agronomic reasons. If knowledge about these evolutions is available
554 (for instance, a Summer crop will be replaced by another Summer crop), it can
555 be easily introduced in the model. The main advantage of MLN with respect to
556 other statistical models like BN is that the changes are limited to the particular
557 set of rules directly related to the events and these changes are not propagated
558 to unrelated rules in the model.

559 **5. Conclusions**

560 In this paper we presented a model which allows predicting the crop grown
561 on a field when the crops grown the previous 3 to 5 years are known. This kind
562 of prediction is useful for the production of crop maps at the field level at the
563 beginning of the agricultural season.

564 Our model applies machine learning techniques using a Land Parcel Infor-
565 mation System, or any other kind of land cover maps from previous years, to
566 model crop rotation patterns. With respect to other models existing in the liter-
567 ature, our approach allows combining automatic learning from data with expert
568 knowledge and make predictions at the field level. We have demonstrated with
569 an illustrative example that this property allows introducing constraints that
570 cannot appear in historical data, like for instance new regulations which may
571 change agricultural practices.

572 We assessed the behavior of the model in terms of scale (area covered) and
573 crop rotation length. We concluded that, in terms of statistical accuracy, the
574 results are good and can be used as a first guess for early crop mapping. The
575 obtained results showed that the proposed approach is able to predict the crop
576 type of each field, before the beginning of the crop season, with an accuracy
577 which can go up to 60%, which is better than the results obtained with current
578 approaches based on remote sensing imagery.

579 One application of this model would be to use it to complement other tech-
580 niques for crop mapping as for instance remote sensing image classification.
581 Remote sensing image time series can achieve good results if enough images are

582 available, usually towards the end of the season. The prediction of the most
583 probable crop could allow achieving good results earlier in the season.

584 The results presented here open perspectives in terms of exploitation of the
585 approach, as for instance including other information as digital elevation models,
586 climatic data or soil type maps.

587 **Acknowledgment**

588 The first author acknowledges the funding by CNES, the French Space
589 Agency, and Région Midi-Pyrénées through a 3 year PhD grant.

590 Aurbacher, J. and S. Dabbert

591 2011. Generating crop sequences in land-use models using maximum entropy
592 and Markov chains. *Agricultural Systems*, 104(6):470–479.

593 Bachinger, J. and P. Zander

594 2007. ROTOR, a tool for generating and evaluating crop rotations for organic
595 farming systems. *European Journal of Agronomy*, 26(2):130–143.

596 Boryan, C., Z. Yang, R. Mueller, and M. Craig

597 2011. Monitoring us agriculture: the us department of agriculture, national
598 agricultural statistics service, cropland data layer program. *Geocarto Inter-*
599 *national*, 26(5):341–358.

600 Castellazzi, M., G. Wood, P. J. Burgess, J. Morris, K. Conrad, and J. Perry

601 2008. A systematic representation of crop rotations. *Agricultural Systems*,
602 97(1):26–33.

- 603 Cussens, J.
604 2001. Integrating probabilistic and logical reasoning. In *Foundations of*
605 *Bayesianism*, Pp. 241–260. Springer.
- 606 Detlefsen, N. K. and A. L. Jensen
607 2007. Modelling optimal crop sequences using network flows. *Agricultural*
608 *Systems*, 94(2):566–572.
- 609 Dogliotti, S., W. Rossing, and M. Van Ittersum
610 2003. ROTAT, a tool for systematically generating crop rotations. *European*
611 *Journal of Agronomy*, 19(2):239–250.
- 612 Flatman, G. T. and A. A. Yfantis
613 1984. Geostatistical strategy for soil sampling: The survey and the census.
614 *Environmental monitoring and assessment*, 4(4):335–349.
- 615 Friedman, N. and D. Koller
616 2003. Being Bayesian About Network Structure. A Bayesian Approach to
617 Structure Discovery in Bayesian Networks. *Machine Learning*, 50(1-2):95–
618 125.
- 619 Gebbers, R. and V. I. Adamchuk
620 2010. Precision agriculture and food security. *Science*, 327(5967):828–831.
- 621 Heckerman, D., D. Geiger, and D. Chickering
622 1995. Learning Bayesian networks: The combination of knowledge and sta-
623 tistical data. *Machine Learning*, 20(3):197–243.

- 624 Inglada, J., J. Dejoux, O. Hagolle, and G. Dedieu
625 2012. Multi-temporal remote sensing image segmentation of croplands con-
626 strained by a topographical database. In *Geoscience and Remote Sensing*
627 *Symposium (IGARSS), 2012 IEEE International*, Pp. 6781–6784. IEEE.
- 628 Inglada, J. and S. Garrigues
629 2010. Land-cover maps from partially cloudy multi-temporal image series: op-
630 timal temporal sampling and cloud removal. In *IEEE Intenational Geoscience*
631 *and Remote Sensing Symposium*, Honolulu, Hawaii, USA.
- 632 Kindermann, R., J. L. Snell, et al.
633 1980. *Markov random fields and their applications*, volume 1. American
634 Mathematical Society Providence, RI.
- 635 Klöcking, B., B. Ströbl, S. Knoblauch, U. Maier, B. Pfützner, and A. Gericke
636 2003. Development and allocation of land-use scenarios in agriculture for
637 hydrological impact studies. *Physics and Chemistry of the Earth*, 28(Recent
638 Development in River Basin Research and Management):1311 – 1321.
- 639 Kok, S., M. Sumner, M. Richardson, P. Singla, H. Poon, and P. Domingos
640 2006. The Alchemy system for statistical relational AI (Technical Report).
641 Department of Computer Science and Engineering, University of Washington,
642 Seattle, WA.
- 643 Le Ber, F., M. Benoît, C. Schott, J.-F. Mari, and C. Mignolet
644 2006. Studying crop sequences with CarrotAge, a HMM-based data mining
645 software. *Ecological Modelling*, 191(1):170–185.

- 646 Leteinturier, B., J. Herman, F. d. Longueville, L. Quintin, and R. Oger
647 2006. Adaptation of a crop sequence indicator based on a land parcel man-
648 agement system. *Agriculture, ecosystems & environment*, 112(4):324–334.
- 649 Lorenz, M., C. Fuerst, and E. Thiel
650 2013. A methodological approach for deriving regional crop rotations as basis
651 for the assessment of the impact of agricultural strategies using soil erosion
652 as example. *Journal of environmental management*, 127:S37–S47.
- 653 Osman, J., J. Inglada, J. Dejoux, O. Hagolle, and G. Dedieu
654 2012. Fusion of multi-temporal high resolution optical image series and crop
655 rotation information for land-cover map production. In *Geoscience and Re-
656 mote Sensing Symposium (IGARSS), 2012 IEEE International*, Pp. 6785–
657 6788. IEEE.
- 658 Petitjean, F., J. Inglada, and P. Gancarski
659 2012. Satellite image time series analysis under time warping. *Geoscience
660 and Remote Sensing, IEEE Transactions on*, 50(8):3081–3095.
- 661 Petitjean, F., J. Inglada, and P. Gancarski
662 2014. Assessing the quality of temporal high-resolution classifications with
663 low-resolution satellite image time series. *International Journal of Remote
664 Sensing*, 35(7):2693–2712.
- 665 Puech, A. and S. Muggleton
666 2003. A comparison of stochastic logic programs and Bayesian logic programs.

- 667 In *Proceedings of the IJCAI-2003 Workshop on Learning Statistical Models*
668 *from Relational Data*, Pp. 121–129.
- 669 Resop, J. P., D. H. Fleisher, Q. Wang, D. J. Timlin, and V. R. Reddy
670 2012. Combining explanatory crop models with geospatial data for regional
671 analyses of crop yield using field-scale modeling units. *Computers and Elec-*
672 *tronics in Agriculture*, 89:51–61.
- 673 Richardson, M. and P. Domingos
674 2006. Markov logic networks. *Machine learning*, 62(1-2):107–136.
- 675 Salmon-Monviola, J., P. Durand, F. Ferchaud, F. Oehler, and L. Sorel
676 2012. Modelling spatial dynamics of cropping systems to assess agricultural
677 practices at the catchment scale. *Computers and Electronics in Agriculture*,
678 81:1–13.
- 679 Schönhart, M., E. Schmid, and U. A. Schneider
680 2011. CropRota—A crop rotation model to support integrated land use as-
681 sessments. *European Journal of Agronomy*, 34(4):263–277.
- 682 Singla, P. and P. Domingos
683 2005. Discriminative training of Markov logic networks. In *AAAI*, volume 5,
684 Pp. 868–873.
- 685 Supit, I., C. Van Diepen, A. De Wit, J. Wolf, P. Kabat, B. Baruth, and F. Lud-
686 wig
687 2012. Assessing climate change effects on European crop yields using the crop

688 growth monitoring system and a weather generator. *Agricultural and Forest*
689 *Meteorology*, 164:96–111.

690 Taillandier, P., O. Therond, et al.
691 2011. Use of the Belief Theory to formalize Agent Decision Making Processes:
692 Application to cropping Plan Decision Making. In *European Simulation and*
693 *Modelling Conference*, Pp. 138–142.

694 Tilman, D.
695 1999. Global environmental impacts of agricultural expansion: the need for
696 sustainable and efficient practices. *Proceedings of the National Academy of*
697 *Sciences*, 96(11):5995–6000.

698 Wechsung, F., V. Krysanova, M. Flechsig, and S. Schaphoff
699 2000. May land use change reduce the water deficiency problem caused by re-
700 duced brown coal mining in the state of brandenburg? (english). In *Landscape*
701 *and urban planning*, volume 51, Pp. 177 – 189.

702 Xiao, Y., C. Mignolet, J.-F. Mari, and M. Benoît
703 2014. Modeling the spatial distribution of crop sequences at a large regional
704 scale using land-cover survey data: A case from France. *Computers and*
705 *Electronics in Agriculture*, 102:51–63.

706 You, L., S. Wood, and U. Wood-Sichra
707 2006. Generating global crop distribution maps: from census to grid. In
708 *Selected paper at IAEA 2006 Conference at Brisbane, Australia*, volume 202,
709 Pp. 1–16.