ARTICLE IN PRESS

Ecological Indicators xxx (xxxx) xxxx

Contents lists available at ScienceDirect

Ecological Indicators

journal homepage: www.elsevier.com/locate/ecolind

Original Articles

# A general framework for quantifying the effects of land-use history on ecosystem dynamics

Leen Depauw[a,*], Dries Landuyt[a], Michael P. Perring[a,b], Haben Blondeel[a], Sybryn L. Maes[a], Martin Kopecký[c,d], František Máliš[e,f], Margot Vanhellemont[a], Kris Verheyen[a]

[a] Forest & Nature Lab, Campus Gontrode, Ghent University, Geraardsbergsesteenweg 267, BE-9090 Melle-Gontrode, Belgium
[b] Ecosystem Restoration and Intervention Ecology Research Group, School of Biological Sciences, The University of Western Australia, Crawley, WA 6009, Australia
[c] Department of GIS and Remote Sensing, Institute of Botany, The Czech Academy of Sciences, Zámek 1, CZ-252 43, Průhonice, Czech Republic
[d] Department of Forest Ecology, Faculty of Forestry and Wood Sciences, Czech University of Life Sciences Prague, Kamýcká 129, CZ-165 00 Praha 6 – Suchdol, Czech Republic
[e] Faculty of Forestry, Technical University in Zvolen, T. G. Masaryka 24, 960 53 Zvolen, Slovakia
[f] National Forestry Centre, Forest Research Institute Zvolen, T. G. Masaryka 22, 960 92 Zvolen, Slovakia

ARTICLE INFO

*Keywords:*
Land-use legacies
Land-use history
Markov chains
Ecological memory
Understorey composition
Light availability

ABSTRACT

Land-use legacies are important for explaining present-day ecological patterns and processes. However, an overarching approach to quantify land-use history effects on ecosystem properties is lacking, mainly due to the scarcity of high-quality, complete and detailed data on past land use. We propose a general framework for quantifying the effects of land-use history on ecosystem properties, which is applicable (i) to different ecological processes in various ecosystem types and across trophic levels; and (ii) when historical data are incomplete or of variable quality.

The conceptual foundation of our framework is that past land use affects current (and future) ecosystem properties through altering the past values of resources and conditions that are the driving variables of ecosystem responses. We describe and illustrate how Markov chains can be applied to derive past time series of driving variables, and how these time series can be used to improve our understanding of present-day ecosystem properties.

We present our framework in a stepwise manner, elucidating its general nature. We illustrate its application through a case study on the importance of past light levels for the contemporary understorey composition of temperate deciduous forest. We found that the understorey shows legacies of past forest management: high past light availability lead to a low proportion of typical forest species in the understorey. Our framework can be a useful tool for quantifying the effect of past land use on ecological patterns and processes and enhancing our understanding of ecosystem dynamics by including legacy effects which have often been ignored.

## 1. Introduction

*Ecological memory* is defined as '*the capacity of past states or experiences to influence present or future responses of the community*' (Padisák, 1992), and as '*the degree to which an ecological process is shaped by past modifications of a landscape*' (Peterson, 2002). The importance of ecological memory in plant and ecosystem processes has been demonstrated in a recent study by Ogle et al. (2015), who showed that various ecosystem processes, across biological, temporal and/or spatial scales, were better explained when models take into account antecedent conditions on top of contemporary conditions. Similar patterns have been observed in other ecosystems (Barron-Gafford et al., 2014; Cable et al., 2013; Hawkins and Ellis, 2010; Leuning et al., 2005; Oesterheld et al., 2001; Sala et al., 2012). An ecosystem's ecological memory is (among other factors) caused by the past land use of the system, which influences the past conditions of the system (Schaefer, 2009; Sun et al., 2013).

Past land use can affect ecosystems for decades to centuries (Foster et al., 2003; Lunt and Spooner, 2005). The system properties resulting from past land use are called land-use legacies (Foster et al., 2003; Kopecký and Vojta, 2009; Perring et al., 2016). Examples of species and communities affected by past land use include plant community composition in forests (De Frenne et al., 2011; Dupouey et al., 2002; Flinn and Marks, 2007; Peterken and Game, 1984), grasshoppers in

* Corresponding author.
*E-mail address:* Leen.Depauw@UGent.be (L. Depauw).

https://doi.org/10.1016/j.ecolind.2019.05.026
Received 16 August 2018; Received in revised form 30 April 2019; Accepted 12 May 2019
1470-160X/ 

Please cite this article as: Leen Depauw, et al., Ecological Indicators, https://doi.org/10.1016/j.ecolind.2019.05.026

woodlands (Hahn and Orrock, 2015), butterflies in grasslands (Moranz et al., 2012), fish and invertebrates in streams (Harding et al., 1998), and birds in Mediterranean forests and shrublands (De Cáceres et al., 2013). In general, there is increasing evidence that past land use can affect future biodiversity over decades to centuries (Bürgi et al., 2017; Essl et al., 2015).

Given the importance of past land use for explaining current and future ecosystem properties, a standardized method to quantify the effects of past land use is needed. Most existing classification schemes or indices for land use consider only contemporary land-use intensity and are developed for one specific ecosystem type, such as forest, grassland or agricultural land (e.g. Blüthgen et al., 2012; Dietrich et al., 2012; Kahl and Bauhus, 2014; Luyssaert et al., 2011; Schall and Ammer, 2013). They do not capture past land use or historical land-use changes and lack general applicability. More general frameworks for quantifying ecological memory (e.g. Ogle et al., 2015) require a lot of data. Such data, including continuous time series, are often lacking for long-term processes (e.g. time scales of decades or even centuries).

We propose a framework that can help resolve the above-mentioned restrictions, by quantifying the effect of land-use history on ecological processes in different ecosystem types, even when data on past land use is incomplete, uncertain and of low quality or resolution. We do not intend to replace existing methods such as the modelling approach from Ogle et al. (2015); our framework can support and complement existing methods through developing the well-needed and often lacking time series of environmental variables. Our basic postulate is that past land use affects current (and future) ecological properties. This occurs through the past land use altering resources and conditions that are the driving variables of ecosystem and community responses (Perring et al., 2016) (Fig. 1). Testing this postulate would be aided by time series data of the driving variables, but such series are rarely available. Trajectories of past land use, even if uncertain, are more frequently known (e.g. McGrath et al., 2015).

Here, we provide a general framework to derive time series of driving variables from known land-use history. By defining the driving variables case-specifically, the framework can be used for a wide range of ecological processes and properties within different ecosystems. In this paper, we describe how Markov chains can be applied to derive time series of driving variables given the known land-use history. Additionally, we provide an illustration of how past values of driving variables can be used to explain current ecosystem properties. Our framework is based on Markov-chain modelling (box 1), a stochastic modelling approach that is often used to model temporal ecosystem changes, such as successional vegetation change, based on temporal autocorrelation in time series (Balzter, 2000; Golroo et al., 2012; Horn, 1975; Logofet and Lesnaya, 2000; Usher, 1981). Markov chains can deal with different types of data as well as uncertainties or missing data, and can incorporate expert knowledge to describe causal relations in the network when long-term data series are lacking (Golroo et al., 2012) (as also implemented in Bayesian belief network modelling (Aguilera et al., 2011; Pollino et al., 2007)). Hence, Markov chains are highly suitable when land-use history data are incomplete or uncertain, which is often the case.

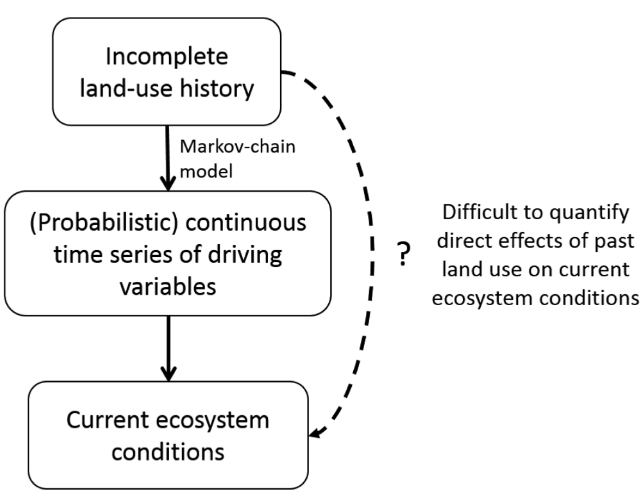



**Fig. 1.** Rationale of the proposed framework: past land use affects current ecosystem conditions through alteration of the resources and conditions that are the driving variables of ecosystem responses. We describe and illustrate how Markov chains can be applied to derive time series of driving variables given the known land-use history, and we provide an illustration of how time series of driving variables can be used to explain current ecosystem conditions.

We describe our framework step-by-step (Section 2, Fig. 3). In each step, we provide a general description of the modelling approach, and illustrate the proposed approach with a specific case study about the effects of past forest management practices on the current understorey composition in temperate forests. We outline some of the main strengths and opportunities of the framework, describe how the model performance could be improved, and discuss the applicability of the framework to assess how past land use influences current ecosystem properties (Section 3).

## 2. Stepwise explanation and illustration of the modelling framework

In our framework, a Markov chain models the dynamics of the driving variables of the studied ecosystem process. A variable representing the land-use history (called *land-use variable*) is added to the chain as an auxiliary variable (cf. box 1, Fig. 2). The final model represents the dynamics of a driving variable, under the assumption that its present state is directly influenced by the current land-use state, and indirectly by past land use, through the past states of the driving variable (Fig. 2).

Below, we describe the modelling approach step-by-step. Each step contains a general explanation and a specific application for a case study. In the case study, we aim to assess the effect of past forest management practices on the current understorey composition, in terms of the proportion of typical forest species (i.e. species found mainly in closed forest, as defined for the lowlands of the Czech Republic, cf. Heinken, unpublished results). We use 29 forest plots from Koda Wood (Czech Republic), Zvolen (central Slovakia) and Slovak Karst (south-eastern Slovakia). For each plot, a description of the management history since 1950 and two vegetation surveys (the first in the 1950s, 60s or 70s, depending on the region, and the second in 2015) are available (see Appendix A). The plots were originally established in mostly oak-dominated forests managed either as coppice, coppice-with-standards or high forests. In each region, we resurveyed plots from all three management categories to cover the historical management variability. Between the surveys, the intensity of forest management generally decreased and shifted from historically dominant coppicing to presently high forest management or no regular management in forest reserves. The change in management resulted in a general decline of plant species richness and a spatial homogenization of the vegetation (Hédl et al., 2010; Kopecký et al., 2013). The species that showed the strongest decline were light-demanding species typical for open oak forests such as *Bupleurum falcatum, Carex montana, Silene nutans, Veronica chamaedrys* agg., *Ajuga genevensis*, *Lotus corniculatus, Campanula persicifolia* and *Tanacetum corymbosum*. In contrast, shade-tolerant, mesic and nutrient-demanding species such as *Aliaria petiolata*, *Asarum europaeum, Hepatica nobilis*, *Mercurialis perennis*, *Galium aparine* and *Neotia nidus-avis* became dominant in the understorey. The annual *Impatiens parviflora* was the only invasive alien species with higher occurrence across the studied plots. The majority of the species in the study plots were perennials (full species list in Appendix G). Tree species regeneration became more abundant, particularly of shade-tolerant tree species such as *Fagus sylvatica* and *Carpinus betulus* (Máliš et al.,

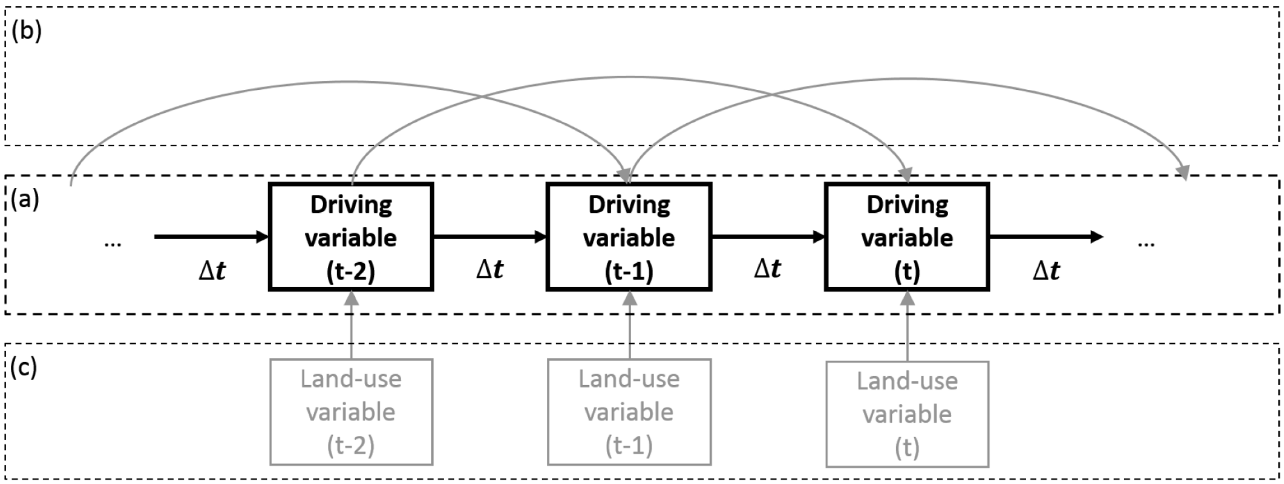



**Fig. 2.** The Markov-chain model used within the framework presented in this paper, consisting of a first-order Markov chain (a) with an additional direct effect (b) between the state variable at t − 2 and the state variable at t (i.e. second-order Markov chain) and an additional auxiliary variable (c) representing the land-use history of the system.

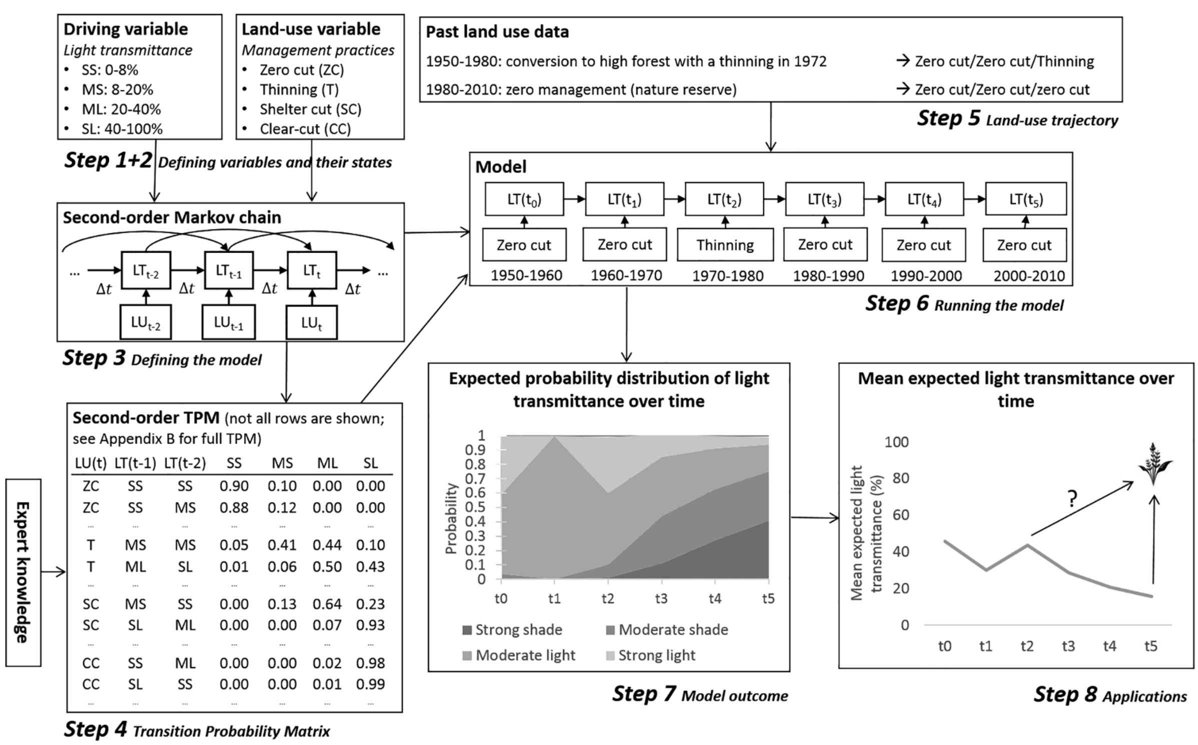



**Fig. 3.** Flowchart illustrating the steps of the framework, applied to our case study. Note that only a few rows of the TPM are shown here as an illustration. The full second-order TPM, with 64 rows, can be found in Appendix B. The data and graphs shown for step 5, 6, 7 and 8 are based on a hypothetical plot with a land-use history as described in Step 5 of the figure. With LT light transmittance, LU land use, SS strong shade, MS moderate shade, ML moderate light, SL strong light, ZC zero cut, T thinning, SC shelter cut and CC clear-cut.

2016).

### 2.1. Step 1: Defining variables

The ecological process of interest is scrutinised to identify its main *driving variables*. For example, soil pH, soil moisture content, nutrient availability, and light availability are important driving variables for plant community composition trajectories (Klanderud et al., 2015), whereas soil temperature and moisture content are among the main driving variables for soil respiration rates (Ogle et al., 2015). Making an informed choice in this first step is vital, as the chosen driving variable (s) should enable the user to evaluate how land use affects the ecological process of interest. We only consider one driving variable in the further description and illustration of the framework, but the entire process can be repeated for the multiple variables that drive the same ecological process.

In our case study, the ecological process of interest is the shaping of the forest understorey community. We selected light transmittance as the driving variable because the understorey composition changes observed in our study regions were strongly related to the light requirement of understorey plants (Hédl et al., 2010; Kopecký et al., 2013) and light availability is one of the main environmental factors controlling the establishment and growth of plant species in forests (Baeten et al., 2009; Thomaes et al., 2013; Tinya and Ódor, 2016). Several studies have observed time lags in vegetation response to understorey light conditions (Dölle and Schmidt, 2009; Thomas et al., 1999), suggesting that past values of light transmittance can be important for current understorey composition. Light transmittance is defined as the ratio of the amount of solar radiation reaching the understorey to the total incident radiation at the top of the canopy (Parker, 2014). It is a common assumption that using light transmittance (%) rather than absolute values of radiation allows for predictions or estimations without knowledge on specific climate and weather conditions (Balandier et al., 2009). Light transmittance depends on forest architecture, and is, as such, mostly uninfluenced by the absolute amount of light at the top of the canopy. Light transmittance depends on canopy closure and hence on the time of the year. In the further description of our case study, we consider the light transmittance in July.

After identifying the process-specific driving variable, a suitable variable representing the land use of the system is defined. The chosen *land-use variable* can be related to one or more of the various aspects comprising land use, such as land cover (e.g. grassland, arable land, forest, heathland), fertilizer type and fertilization intensity, soil manipulation (e.g. ploughing, tilling), harvesting (e.g. crop type in arable fields, different management regimes for timber production in forests, litter raking in forests), and should have a potential effect on the driving variable. For example, past fertilization type and intensity can be suitable land-use variables when soil pH is chosen as the driving variable (Koerner et al., 1997).

As the land-use variable in our case study, we selected forest management, given its possible impact on the canopy composition and structure and hence on light transmittance (Thomaes et al., 2014) and the forest understorey (e.g. Kopecký et al., 2013; Perring et al., 2018; Ujházy et al., 2017; Van Calster et al., 2008). We did not consider other factors affecting light transmittance, such as tree species and phenology, but kept in mind that these could influence the interpretation of the results.

### 2.2. Step 2: Discretization of variables

First, to be able to use a driving variable in our Markov chain, the variable needs to be discretized (cf. box 1) by defining a finite set of ecologically relevant, representative states (Carpinone et al., 2015; Shamshad et al., 2005). In our case study, we defined sensible discrete states for light transmittance, looking at the relationship between light transmittance and understorey community composition in temperate deciduous forests in Europe. We used three threshold values between four light transmittance states: strong shade (0–8%), moderate shade (8–20%), moderate light (20–40%) and strong light ($>$ 40%). Many understorey species of temperate deciduous forest benefit from light levels below 8%, when the survival of certain competitors is limited (De Keersmaeker et al., 2004). For some forest understorey species, the survival is higher under moderate levels of shade (8–20%) than under strong shade ($\leq$8%) (Thomaes, 2014). Understorey cover reaches an asymptotic maximum at around 40% light transmittance (Balandier et al., 2009).

Second, similar to the driving variable, also the land-use variable needs to be discretized. In our case study, we defined four states of forest management (further on referred to as *land-use states*) that cover a gradient in management intensity, and encompass the typical forest management actions in our study regions:

- Zero cut: no tree fellings or removals, forest under a zero management system or forest in a period in between two interventions of a rotation system;
- Thinning: the removal of a proportion of trees to allow more growing space for the final crop trees (den Ouden et al., 2010) or management actions with similar effects on the canopy structure, such as selection felling of single trees;
- Shelter cut: a method of securing natural tree regeneration under the sparse shelter of old trees that are removed by successive cuttings to admit a gradually increasing amount of light to the seedlings (den Ouden et al., 2010) or the cutting phase in a coppice-with-standards system resulting in a similar forest structure;
- Clear-cut: most or all trees in an area are cut, e.g. the harvesting phase of coppice systems or high forest systems with a clear-felling management.

Third, the magnitude of the time step ($\Delta$t) in the chain should be clearly defined. The time step can vary from less than seconds to more than years, depending on the chosen driving and land-use variables, the ecological process considered, and the availability of land-use history data (Carpinone et al., 2015). In our case study, the time step ($\Delta$t) is mainly constrained by the temporal resolution of the available land-use history data (Section 2.5) and set at 10 years. The 10-year time step corresponds well to the typical management cycles in temperate forests (den Ouden et al., 2010; Kerr and Haufe, 2011), but might be too long to detect short-term temporal dynamics in understorey composition. Smaller time steps would have been better to predict light dynamics that drive understorey composition. However, due to the absence of high-resolution land-use history data, high-resolution predictions of light dynamics would be highly uncertain and therefore contain no additional information compared to the light availability data derived from the model with $\Delta t = 10$ years.

### 2.3. Step 3: Defining the model

One can adjust the proposed Markov-chain model to the system and the driving variable of interest by defining the appropriate order of the Markov chain. The order of a Markov chain is the number of time steps in the past that can directly influence the current state (Shamshad et al., 2005). In a simple first-order Markov chain, the present state of the modelled variable only depends on the previous state of that variable. However, for some ecological processes, it might be necessary to include higher-order terms to the chain, to account for the possible ecological memory in the dynamics of the driving variables controlling the processes. For example, adding a second-order arrow to the chain, implies that the state of the driving variable at time t can depend both on the previous state ($t-1$) and the state before that ($t-2$) (box 1) (Usher, 1979). The order that should be used when applying the framework will be case-specific, and depends on the expected ecological memory of the driving variable that is modelled, as well as on the level of complexity that can be dealt with in the Transition Probability Matrix (TPM; see Section 2.4). When validation data are available, results from chains with different orders can be compared to assess how long influences of the past remain important for contemporary states. In addition, mathematical methods are available to identify whether second-order relations are sufficiently important to include when compared to the first-order relations in the model (BayesFusion, 2017). We show later (see Section 2.4) that in our particular case study, a first order model was sufficient to model the light dynamics over time.

### 2.4. Step 4: Transition probability matrix

The Transition Probability Matrix (TPM; box 1) quantifies the causal relations between the different variables in the Markov chain (Logofet and Lesnaya, 2000; Shamshad et al., 2005). In the context of this study, expert-based approaches are best suited to derive the TPM. Experts are asked to complete a TPM according to their knowledge and expectations, and to report their confidence in each estimate (Kuhnert et al., 2010; Pollino et al., 2007). These confidence levels are then used to weight the estimates of all experts in a final TPM (Pollino et al., 2007). It is important to clearly define the investigated process and boundary conditions to ensure that different expert estimates are based on the same assumptions and thus comparable.

In our case study, the second-order TPM describes the probability for light transmittance (LT) at time t being in one of the four defined states, given the light transmittance state of the system at time $t-1$ (i.e. ten years ago) and $t-2$ (i.e. twenty years ago), and the land-use state (i.e. forest management) at time t ($LU_t$). Since both variables (light transmittance and forest management) have four possible states, the second-order TPM contains 64 scenarios = 4 ($LT_{t-2}$) × 4 ($LT_{t-1}$) × 4 ($LU_t$). A team of six experts (all author of this paper) provided a probability distribution and a confidence level for this probability distribution for each of these 64 scenarios, resulting in one second-order TPM (see Appendix B). Clear guidelines, definitions, boundary conditions and assumptions were provided to all experts (Appendix C). Based on the second-order TPM, we calculated the *strength of influence* between nodes (see box 1) in the Markov chain. We found a strength of influence of 0.03 for the second-order relation (influence of $LT_{t-2}$ on $LT_t$) and 0.35 for the first-order relation (influence of $LT_{t-1}$ on $LT_t$). Light transmittance at *t* thus mainly depended on light transmittance at $t-1$, and less on light transmittance at $t-2$. The strength of influence of $LU_t$ on $LT_t$ was 0.49. We concluded that a first-order Markov chain is sufficient to model the light dynamics over time given the land-use trajectory. All further results and figures are from the first-order Markov chain. We derived a first-order TPM by marginalization (i.e. grouping scenarios with the same light transmittance state at $t-1$ (thus: only differing in the light transmittance state at $t-2$) and calculating the average probability distribution for each group of scenarios) (Table 1, Appendix B). The first-order TPM describes the probability for light transmittance (LT) at time *t* being in one of the four defined states, given the light transmittance state of the system at time $t-1$ (i.e. ten years ago) and the land-use state at time *t* ($LU_t$), and thus contains 16 scenarios = 4 ($LT_{t-1}$) × 4 ($LU_t$).

### 2.5. Step 5: Land-use trajectory

Knowledge on past land use can be gathered from natural archives, such as tree-ring series or soil properties, and cultural archives, such as old aerial pictures, historical maps, old management plans, and face-to-face interviews with locals, land owners or managers. The land-use trajectory comprises the translation of what is known about the past land use of the system into a sequence of the possible land-use states defined in Section 2.2 (step 2). Thus, for each time step in the chain, the land-use state that best describes the situation at that time needs to be determined, and will be entered in the Markov chain as evidence. This can, depending on the certainty of the land-use trajectory, either be done as *hard evidence*, assigning a 100% probability to the assumed land-use state at each time step, or as *soft evidence*, providing probabilities for the different states of the land-use variable that sum up to 100% (box 1).

For our case study, two authors of this paper, each with detailed knowledge of the case study regions, investigated the management history of the 29 plots and completed a standardized land-use history questionnaire (Appendix D). The historical information was used to assign a land-use state to each 10-year time step for each plot, starting in 1950 (Appendix E). Some assumptions were necessary, due to variations in the level of detail of the available historical data (Appendix E). To illustrate the possibility of including an uncertain land-use trajectory in the model, we defined three alternative trajectories for one of the Czech plots (Plot KO775; Table 2). The historical information for this plot mentioned sanitary thinnings of standards in the period 1900–2010. We assumed that every 30 years one of these thinnings affected the plot and used a different timing of this thinning frequency in the three alternative land-use trajectories. Presuming that each alternative is equally likely, each time step between 1950 and 2010 has a 66.6% probability of 'zero cut' and 33.3% probability of 'thinning', which can be included in the model as soft evidence.

### 2.6. Step 6: Running the model

Numerous software packages can be used to implement and run Markov-chain models. Aside from software packages that are often used for Markov-chain modelling (e.g. R (Spedicato, 2017), MARCA (Stewart, 1996), PRISM (Kwiatkowska et al., 2011)), also software packages primarily designed for Bayesian belief network modelling can be highly suitable (e.g. Netica (Norsys, 1998), Hugin (Hugin, 2008) and GeNie (Druzdzel 1999; http://www.bayesfusion.com)) (Landuyt et al., 2013). In our case study, models were implemented and run using the free software package GeNie. We built the model structure (a first-order Markov chain with one auxiliary variable), and entered the weighted-average TPM of the six experts (cf. Appendix B). Then, we entered the assumed land-use state for each considered time step, first as hard evidence (i.e. assigning a 100% probability to the assumed land-use state) for all 29 plots, and then as soft evidence (i.e. providing probabilities for the different states of the land-use variable that sum up to 100%) for one of the plots, to illustrate how using hard vs. soft evidence

**Table 1**

The first-order Transition Probability Matrix (TPM) derived from the second-order TPM by marginalization. The pie charts represent the average expected probability distribution of light transmittance at t for the 16 different scenarios (i.e. 16 combinations of the land-use state at t and the light transmittance state at (t − 1). The full first- and second-order TPMs can be found in Appendix B.

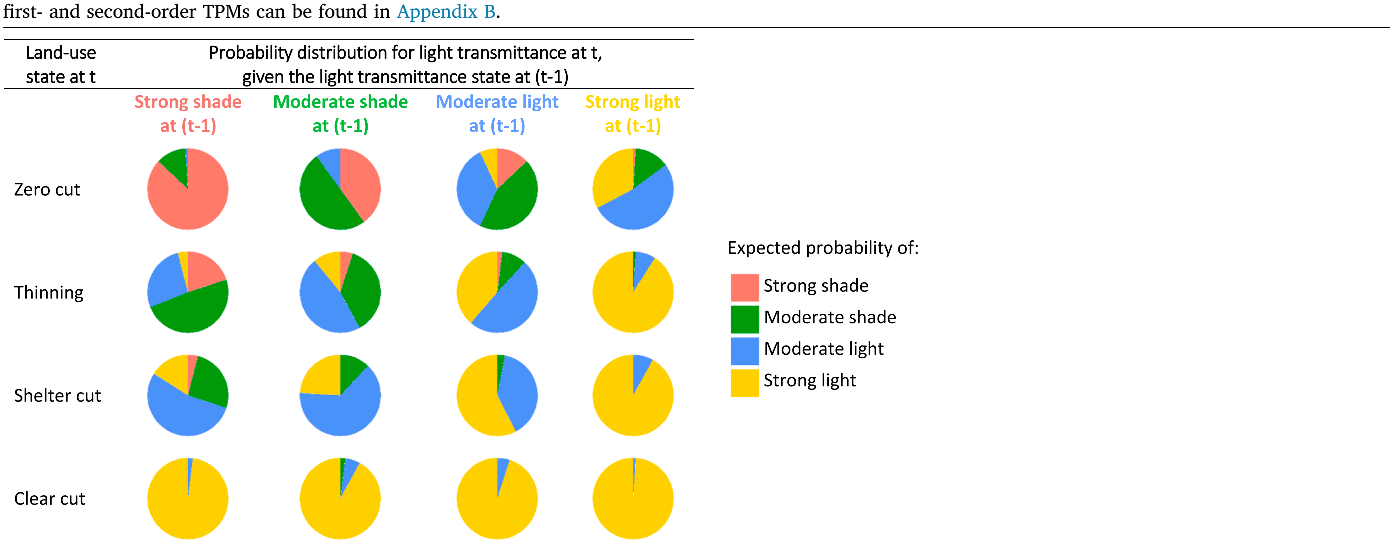
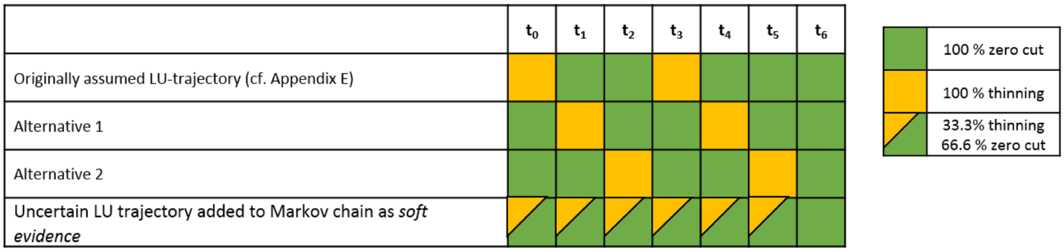


**Table 2**
Three alternative land-use (LU) trajectories for one of the Czech plots (KO775), with shifted timings for the sanitary thinnings that took place between 1900 and 2010. The last row shows how alternative trajectories can be combined into one uncertain land-use trajectory, which can be entered in the model as soft evidence.




influences the results. For each of the 29 study plots, the model then calculated the probability of each light transmittance state to occur at each time step (for seven time steps of 10 years; from 1950 to 2020), given the specific land-use trajectory of the plot.

Note that the model can be updated with evidence on the state of the driving variable at certain time steps (in case these data are available). In our case study, we have light transmittance data for time step $t_6$ (2010–2020). We first used these data to evaluate the model outcomes (Section 2.7) and then updated the model using the light transmittance data as evidence to generate model outcomes for further analysis (see Section 2.8 for details).

### 2.7. Step 7: Evaluation of model outcomes

The final model output is a probability distribution of the different states of the driving variable at each time step. In other words, the probability for each possible state of the driving variable at each time step is predicted based on the land-use history data and the TPM (Fig. 3). From the probability distribution output, a user can derive several variables to use in further analyses. Time series of, for instance, the mean expected value, the most probable state to occur or the probability for a certain state to occur (e.g. Dlamini, 2010; Smith et al., 2007) can be used to further investigate and analyse ecological process dynamics. In our case study, we calculated the mean expected value of light transmittance at each time step based on the probability distribution at each time step and the mean value of each light transmittance (LT) state:

$$\begin{aligned}\text{mean expected } LT &= P_{SS}.\ \bar{SS} + P_{MS}.\ \bar{MS} + P_{ML}.\ \bar{ML} + P_{SL}.\ \bar{SL} \\ &= P_{SS}.\ 4\% + P_{MS}.\ 14\% + P_{ML}.\ 30\% + P_{SL}.\ 70\% \end{aligned} \tag{1}$$

with $\bar{SS}$, $\bar{MS}$, $\bar{ML}$ and $\bar{SL}$ the class means of respectively strong shade, moderate shade, moderate light and strong light; and with P the probability for a light transmittance state to occur.

Metrics to evaluate the performance of models that produce a probabilistic output include confusion tables, *k*-fold cross-validation, receiver operating characteristic curves, and several performance indices such as spherical pay-off, Schwarz' Bayesian information criterion, and true skill statistic (Marcot, 2012). Another commonly used approach is based on comparing the model performance to the expected percentage of correct classifications if the prediction was made in a random manner (i.e. by a model called *random classifier* or *baseline classifier*) (e.g. Genc and Dag 2016). In our case study, we used light transmittance data obtained from the 2015 survey that took place in each of the 29 plots (Appendix A) to evaluate the model performance. We measured light transmittance with a spherical densiometer (Forestry Suppliers, 2008; Lemmon, 1957). For the time step $t_6$ (2010–2020) for which observed light transmittance data are available, we compared model predictions against predictions of an indifferent baseline classifier (uniform distribution). For each plot, the model performance was expressed as the predicted probability of the *observed* light transmittance state at the survey time, minus the baseline probability of that state. Since the defined light transmittance classes were unbalanced, baseline probabilities, derived from a uniform distribution, were set to 8%, 12%, 20% and 60%, for the states 'strong shade', 'moderate shade', 'moderate light' and 'strong light', respectively. Positive model performance values, where predicted probability values are higher than their baseline, indicate that model predictions are informative.

In our case study, the model performance differed between plots (Fig. 4), and for the majority of the plots, the informed model was performing better than the random (baseline) model (more positive than negative values in Fig. 4). Many of the plots for which the model performed badly were thinned within the 20 years prior to the survey. Thinning events close to the survey hence seemed to decrease the model's performance. Two possible explanations for this observation are: (i) the documented thinnings might not have taken place in or close to the plot, and (ii) the experts who completed the TPM might have wrong expectations about the effect of thinnings on light levels. The experts generally assumed thinnings to increase light levels, but a recent study showed that light levels at the forest floor can be similar in forests with a dense vs. a more open canopy, due to a higher shrub density in the more open forests (Sercu et al., 2017).

Including uncertainty in the timing of thinning events in our model resulted in a more gradual change in predicted average light transmittance over time compared to the cyclic behaviour of light transmittance for thinning events with a certain timing (Fig. 5). Yet, the general trend, i.e. an overall decrease in light transmittance over time, was similar for certain and uncertain land-use trajectories.

### 2.8. Step 8: Application of model outcomes

For the 29 plots of our case study, we have vegetation data from two surveys (the first survey in the 1950s, 60s or 70s, depending on the region, and the second in 2015; see Appendix A). The survey data

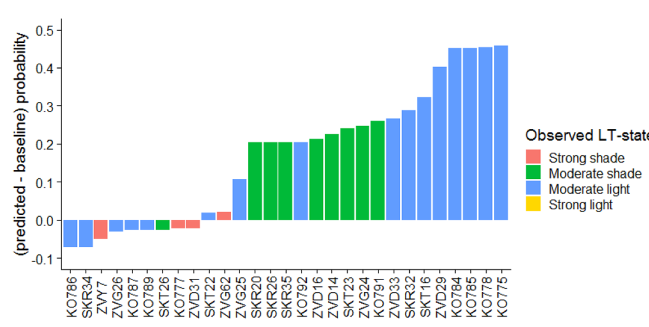



**Fig. 4.** Measure of model performance for the 29 plots of our case study, calculated as the predicted probability of the observed light transmittance state (at the 2015 survey) minus the baseline probability of that state (based on a uniform distribution). The more positive the value, the better the model predictions. The colours of the bars indicate the observed light transmittance state during the 2015 survey.

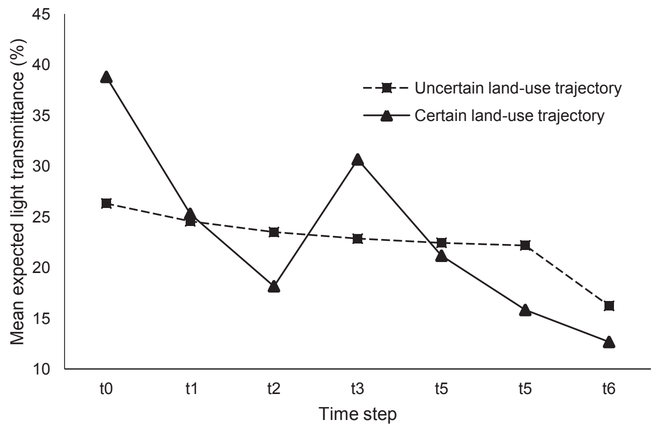



**Fig. 5.** Comparison between the results of a first-order Markov chain, with and without accounting for uncertainty in the land-use (LU) trajectory (see Table 2), for one plot from our case study (KO 775) and seven 10-year time steps during 1950–2020.

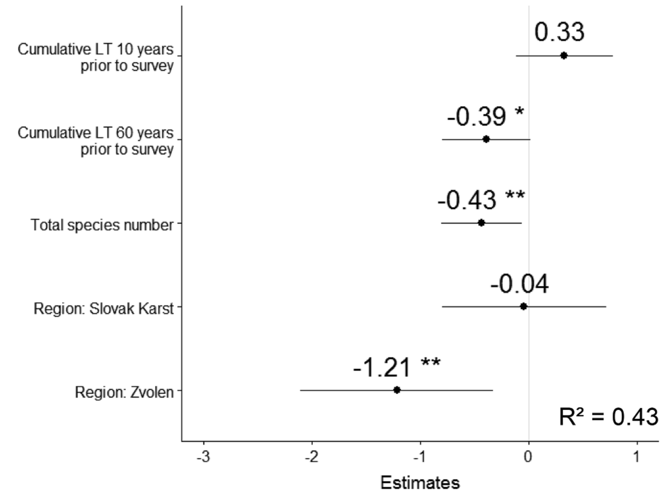



**Fig. 7.** Effect sizes of cumulative light transmittance (LT) over the past 10 and 60 years prior to the survey for the proportion of typical forest species in the total species pool. Significant effects are indicated with '*' ($p < 0.10$). The effect sizes of the covariates 'total species number' and 'region' are also shown.

comprise an estimated cover (in %) for each species in three separate layers, i.e. tree layer (all trees taller than half of the height of the canopy trees), shrub layer (all woody plants taller than 1.3 m not included in the tree layer) and understorey (all plants smaller than 1.3 m). We have data on light transmittance for the 2015 survey, measured with a spherical densiometer, and derived estimates of light transmittance for the first survey through the relationship between the light transmittance and tree and shrub cover data of the second survey (see Appendix F). We included the light transmittance data of both time steps (the two survey times) as evidence in our model to calculate a time series of mean expected light transmittance for each plot. We expect that including evidence will make the model results more informative, but we cannot quantify this effect, as there is no validation data available. We did not include uncertainty in the land-use trajectory to obtain the estimated light transmittance over time. We used the obtained time series, combined with the vegetation data from the 2015 survey, to assess the importance of past light levels on the current understorey community composition.

The data from the two surveys provide light transmittance values at two time points, as well as an estimation for light transmittance values in between both surveys, given we assume linear dynamics (Fig. 6a). Our framework, however, allows uncovering the light transmittance in between surveys, demonstrating that two plots with very similar light levels during both surveys may have experienced completely different light regimes in between surveys (Fig. 6b).

We used a simple linear model to explore the importance of past light levels for understorey community composition. The response variable was the proportion of typical forest species (i.e. species found mainly in closed forest, as defined for the lowlands of the Czech Republic, cf. Heinken, unpublished results) in the understorey community (all plants smaller than 1.3 m height, including tree species) in the 2015 survey. The explanatory variables were the cumulative light transmittance, i.e. the area under the curve of estimated light transmittance over time (Fig. 6c), for 10 and 60 years prior to the 2015 survey. As covariates, we included the total number of species present in 2015 and the region (i.e. Koda Wood, Zvolen, or Slovak Karst – see Section 2) of a plot.

We found that the cumulative light transmittance over a period of 60 years prior to the survey was a better predictor of the proportion of typical forest species in a plot's understorey community ($p = 0.07$), compared to the cumulative light transmittance of the recent past (i.e. 10 years prior to the survey) ($p = 0.16$) (Fig. 7). This suggests that the current understorey composition is better explained by cumulative light levels over the past 60 years than by the more recently prevailing light levels. Study plots with a higher number of species in the understorey had a lower proportion of typical forest species, and the plots in Zvolen had a lower proportion of typical forest species than in the other two regions. The model explains 43% of the variation in the proportion of typical forest species ($R^2 = 0.43$); an acceptable $R^2$-value for ecological processes. Our findings suggest that management legacies are present in forest understoreys and are in accordance with Thomas et al. (1999)

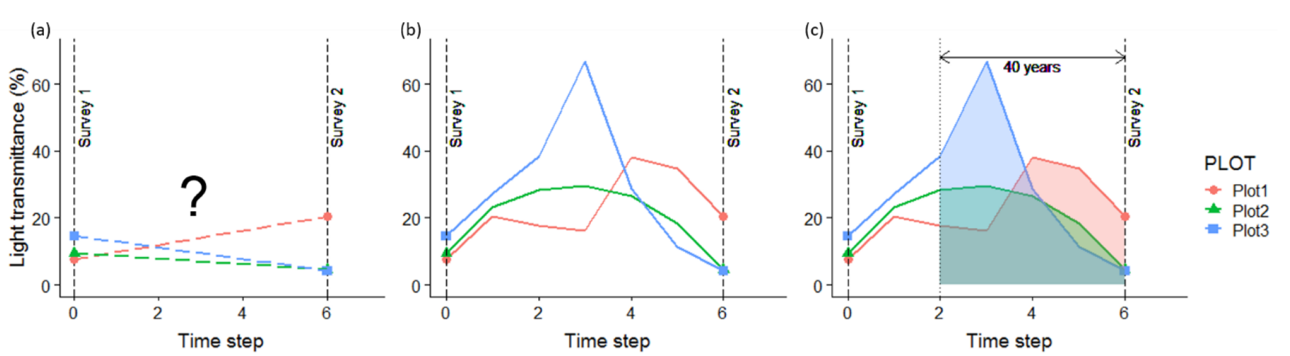



**Fig. 6.** Graphical illustration of the added value of our framework for a resurvey study, using 3 of our 29 study plots. Light transmittance values are only available at the two survey times. (a) Light transmittance between both surveys can be estimated through linear interpolation. (b) Using our framework, light transmittance in between surveys can be estimated, demonstrating that two plots with similar light levels for both surveys may have experienced completely different light regimes between the surveys. (c) The projected time series of light transmittance can be used to calculate, for example, the cumulative light transmittance over the 40 years before a survey.

and Dölle and Schmidt (2009), who found that the light-vegetation relationship might be better explained by past light regimes than by current light conditions because of the slowness of plant community changes. Note that our findings are limited by (i) the small sample size and (ii) possible correlation structures among plots in each region that are not accounted for in our simple analysis. All analyses were performed in R 3.3.2 (R Core Team, 2017).

## 3. Discussion

We proposed a framework based on the hypothesis that past land use affects current ecosystem properties through its impact on past values of driving variables (Fig. 1). We used our framework to model the temporal dynamics of one such driving variable (i.e. light transmittance) based on land-use history data, to look for effects of past land use on current understorey composition in temperate forests. To more thoroughly estimate the past resources and conditions of an ecosystem, the modelling could be repeated for other driving variables relevant for the particular study system.

### *3.1. Strengths of the framework*

The strength of the framework is its applicability to different types of ecological processes and ecosystems, while previously developed indices or classification schemes for quantifying land-use legacies were only applicable to specific ecosystems, such as forests (e.g. Schall and Ammer 2013; Kahl and Bauhus 2014), grasslands (e.g. Blüthgen et al., 2012), or agricultural fields (e.g. Dietrich et al., 2012). The modelling framework of Ogle et al. (2015) for quantifying ecological memory is also applicable in different ecosystems, but has the disadvantage of requiring long continuous time series. When such long-term data are unavailable or incomplete, which is often the case, our framework offers the opportunity to derive time series of biologically meaningful driving variables from uncertain or incomplete land-use data.

Markov chains offer the advantage that they can handle low-quality land-use data with high uncertainties since both hard evidence (100% certainty about the land use at a certain time point, e.g. based on photographs) and soft evidence (probabilistic information about the land use at a certain time point, e.g. based on expert information) can be inserted (Jensen and Nielsen, 2007). The general applicability of the proposed framework is further improved by allowing the user to adjust the order of the Markov chain, depending on the expected extent of influences of the past. For our case study, where we model light transmittance over time for a given land-use trajectory, we found very small influences of the second-order term of the Markov chain (based on the Transition Probability Matrix (TPM)), suggesting that light transmittance at the forest floor mainly depended on more recent management events.

### *3.2. Opportunities for improving model performance*

The poor model performance that we observed for some of the plots in our case study can have several reasons. We believe the most important reason is the high uncertainty of the data on past land use. As the exact timing of management interventions was often unknown, especially at the plot level, we can't expect to be able to accurately predict light transmittance values at a specific point in time. In addition, the resolution of the Markov chain in the application (i.e. time intervals of 10 years) might be too low to capture small fluctuations in light availability that might have had an impact on the understorey. However, when the aim of the model is to derive general trends in the dynamics of a driving variable, such as cumulative light availability, this bias can be considered less problematic. We illustrated this with one of the plots from our case study (Fig. 5), where similar general trends were predicted with and without accounting for uncertainty in the land-use trajectory.

Another potential weakness of the framework is the strong dependence of the model output on the quality of the Transition Probability Matrix (TPM), which depends on the knowledge of the consulted experts. However, the TPM might be improved by including literature data and data-learning techniques to estimate the conditional probabilities. The latter, however, requires extensive long-term data, which are often not available. Providing experts with clear guidelines and background information on the investigated process and boundary conditions is key for obtaining high-quality TPMs. In addition, when multiple experts have provided a TPM, running the model with each separate TPM instead of the (weighted) average TPM can provide information on the dependency of the model results on the TPM, and can reveal how some TPMs better fit the data (assuming qualitative validation data is available) than others and should therefore be given more weight in the final TPM.

Finally, information loss through strong simplifications due to the discrete nature of Markov chains can decrease model performance. There is a trade-off between accuracy and complexity, as an increase in the number of states will also increase the number of rows of the TPM. By using ecologically relevant thresholds, information loss through discretization can be minimized.

To deal with the abovementioned issues, a lot can be learned from recent advances in the field of Bayesian belief network modelling, a modelling technique that also works with discrete variables and an identical probabilistic knowledge base that is often derived from a combination of literature data, field data and expert knowledge (see, for example, Murphy (2002)). Within this field, expert knowledge elicitation techniques (e.g. Kuhnert et al., 2010; Pollino et al., 2007), and data assimilation techniques (e.g. Chen and Pollino, 2012; Marcot et al., 2006) to combine different data sources have been developed and optimized.

Marcot (2012) suggests that Bayesian belief networks may best be developed stepwise, starting from a less ambitious model based on expert knowledge, testing and calibrating the model, updating the structure of the model and retesting it until a satisfying performance is reached. In this paper, we used Markov chains, which are related to Bayesian belief networks and also offer the flexibility to update the model with auxiliary variables, such as the land-use variable in Fig. 2. They can easily be extended even further, depending on the complexity of the ecological processes that are studied. For example, if next to land use, other variables influence the state of the driving variable, these can be added to the chain as well, and model performance can be tested again. Of course, this will only work if we have temporal data on this additional auxiliary variable and if the relation between this variable and the driving variable can be quantified through experts or data. Besides, the improvement of model performance can only be tested when qualitative validation data is available.

### *3.3. Applicability of the framework*

With our framework, we are able to predict time trends of driving variables of ecological processes and properties, for a given land-use history. We believe this is a key step leading to further investigation of how past land use affects current ecosystems. Long time series of measured past resources and conditions are often not available. With the time trends we model, we can reveal some of the likely past behaviour of these resources and conditions (cf. Fig. 6), allowing us to detect why systems with seemingly similar contemporary resources and conditions can display different properties. In our case study, we derived past light dynamics to assess how current herb layer communities are (partially) shaped by past light availability, and revealed why forest plots with similar current light conditions have different herb layer communities. Several other drivers, such as soil pH, nutrient availability and soil moisture content also affect herb layer communities (Klanderud et al., 2015). It would therefore be interesting to apply the proposed framework on the other important driving variables, which might be influenced by other land-use variables. It may not always be feasible to determine all driving variables of an ecological process, but gaining insight into the dynamics of a subset of the driving variables will already improve our understanding of the process and its dependence on past land use.

We hope our framework will provide an opportunity for further studies on how past ecosystem properties (i.e. past levels of resources and conditions), controlled by past land use, are affecting contemporary ecological properties and patterns. The modelling approach can easily be translated to different driving variables and different land-use variables and can be extended or adapted depending on the complexity of the study system. We therefore believe the proposed approach is widely applicable in studies where researchers have (some) data on past land use and want to take those into account to achieve a better understanding and better predictions of the contemporary or future ecological state.

**Box 1**

Theoretical background of Markov chains.

**Markov chains** are graphical, multivariate, statistical models, representing dynamic systems wherein variables can go from one state to another over time, with a transition probability that depends on preceding conditions (see Fig. 1 box 1). A Markov chain consists of **nodes**, representing the system's variables, and **arrows**, representing the causal relations among these variables. Each variable is discrete and characterized by a set of **states** it can manifest (numerical values, discrete classes or qualitative levels) and a **probability distribution** that quantifies the probability of being in one of the states. If such a probability distribution depends on the state of another variable, it is referred to as a **conditional probability**, which quantifies the causal relation represented by an arrow. Through **probabilistic inference**, a Markov chain can infer the probability distribution for a given variable conditional on the state of the other variables in the model (Jensen and Nielsen, 2007).

**The order of a Markov chain** (Fig. 1 box 1) is the number of time steps in the past that influence the probability of the current state (Shamshad et al., 2005).

**Auxiliary variables** can be added to Markov chains to model more complex processes with multiple variables. For example, in Fig. 2 box 1, the state of the variable X at each time step depends on the state of X at the previous time step (first-order Markov chain), and on the state of the auxiliary variable Y at the current time step.

**The Transition Probability Matrix (TPM)** is the core of a Markov chain, in which each element represents the probability that a variable is in a certain state, at a certain time step, given the state of the previous time step(s) (Golroo et al., 2012; Logofet and Lesnaya, 2000; Shamshad et al., 2005).

Let $X$ be a variable, possessing discrete states $S$ ($S=\{1,2,\ldots,m\}$). In general, for a given sequence of time points $t_1 < t_2 < \cdots < t_{n-1} < t_n$, the conditional probability for $X$ to be in a certain state at time $t_n$ is (Balzter, 2000; Logofet and Lesnaya, 2000; Shamshad et al., 2005):

$$P(X_s(t_n)|X(t_1), X(t_2), \cdots, X(t_{n-1})) \tag{1}$$

In formula (1), $X(t_n)$ depends on the state of X at all previous time steps $t_1, \cdots, t_{n-1}$, representing a Markov chain of order $n-1$. Formulae (2) and (3) show the conditional probabilities for a first- and second-order Markov chain:

$$P(X(t_n)|X(t_{n-1})) \tag{2}$$

$$P(X(t_n)|X(t_{n-2}), X(t_{n-1})) \tag{3}$$

These conditional probabilities make up the TPM. For m states, the first-order TPM takes the form (Shamshad et al., 2005):

$$TPM = \begin{bmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,m} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,m} \\ \vdots & \vdots & p_{i,j} & \vdots \\ p_{m,1} & p_{m,2} & \cdots & p_{m,m} \end{bmatrix} \tag{4}$$

with $p_{i,j}$ the probability of state $i$, if the previous state was $j$.

Similarly, the second-order TPM takes the form (Shamshad et al., 2005):

$$TPM = \begin{bmatrix} p_{1.1,1} & p_{1.1,2} & \cdots & p_{1.1,m} \\ p_{1.2,1} & p_{1.2,2} & \cdots & p_{1.2,m} \\ \vdots & \vdots & p_{i,j,k} & \vdots \\ p_{1.m,1} & p_{1.m,2} & \cdots & p_{1.m,m} \\ p_{2.1,1} & p_{2.1,2} & \cdots & p_{2.1,m} \\ p_{2.2,1} & p_{2.2,2} & \cdots & p_{2.2,m} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m.m,1} & p_{m.m,2} & \cdots & p_{m.m,m} \end{bmatrix} \tag{5}$$

with $p_{i,j,k}$ the probability of state $i$, if the states at the two previous time steps were (in chronological order) $k$ and $j$.

In Markov chain studies, a TPM is often derived from empirical evidence or machine learning (Balzter, 2000; Logofet and Lesnaya, 2000; Usher, 1981). However, transition probabilities can also be derived from expert knowledge (Aguilera et al., 2011), a particularly suitable approach when long-term data series are lacking (Golroo et al., 2012; Pollino et al., 2007).

**The strength of influence** can be calculated for each arrow in a Markov chain based on the Transition Probability Matrix (TPM), and represents a measure for the extra information that is obtained by knowing the value of the parent (i.e. the node where the arrow starts from) (Theijssen et al., 2013). In other words, it quantifies how much the value of the parent node affects the value of the child node (i.e. the node where the arrow arrives).

**Belief updating** is the process of inserting new information (evidence) on the status of one of the variables in a Markov chain. This will change the probability distribution of other variables in the network, and lower the uncertainty in the model output (Jensen and Nielsen, 2007). The process of inserting **hard evidence** into the network is called **instantiation**, and comprises assigning a 100% probability to one of the states of a variable. **Soft evidence** provides probabilistic information on the status of a variable (Jensen and Nielsen, 2007).

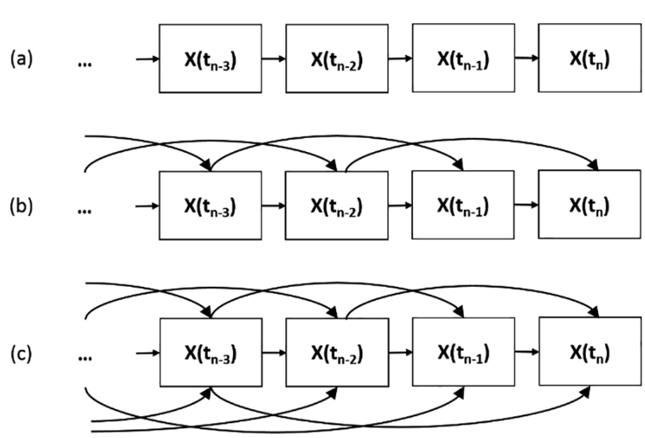



**Fig. 1 box 1.** In the first-order Markov chain (a), the state of X only depends on the state of X at the previous time step. In the second- and third-order Markov chains (b, c), the state of X depends on the state of X at the two and three previous time steps, respectively.

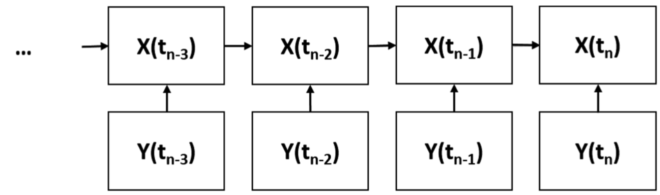



**Fig. 2 box 1.** First-order Markov chain with one auxiliary variable. The state of the variable X at each time step depends on the state of X at the previous time step, and on the state of the auxiliary variable Y at the current time step.

## 4. Author contribution statement

LD, DL, MPP and KV conceived and designed the study with significant contributions from SLM and HB; KV and MPP provided insight and ideas for the statistical analyses; SLM, HB, MPP, MV, MK and FM assisted with the data collection; LD led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

## 5. Data accessibility

We intend to archive all data used in this paper on our public website: www.pastforward.ugent.be.

## Acknowledgements

We thank the European Research Council (ERC Consolidator grant no. 614839: PASTFORWARD) for funding the scientific research and fieldwork. DL was supported by a postdoctoral fellowship of the Research Foundation-Flanders (FWO). MK was supported by institutional support RVO 67985939. FM was supported by APVV-15-0270 and APVV-15-0176. We thank Martin Macek and Karol Ujházy for assisting us with the relocation of the plots and the determination of plant species. We also thank Robbe De Beelde, Sanne Govaert and Bram Bauwens for their support with the fieldwork, and Emiel De Lombaerde and Pieter Vangansbeke for their input on the data analysis. We thank three anonymous reviewers and the Editor, Christine Fürst, for their constructive comments that improved and clarified the manuscript.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ecolind.2019.05.026.

## References

Aguilera, P.A., Fernández, A., Fernández, R., Rumí, R., Salmerón, A., 2011. Bayesian networks in environmental modelling. Environ. Model. Softw. 26, 1376–1388. https://doi.org/10.1016/j.envsoft.2011.06.004.

Baeten, L., Bauwens, B., De Schrijver, A., De Keersmaeker, L., Van Calster, H., Vandekerkhove, K., Roelandt, B., Beeckman, H., Verheyen, K., 2009. Herb layer changes (1954–2000) related to the conversion of coppice-with-standards forest and soil acidification. Appl. Veg. Sci. 12, 187–197.

Balandier, P., Marquier, A., Dumas, Y., Gaudio, N., Philippe, G., Da Silva, D., Adam, A., Ginisty, C., Sinoquet, H., 2009. Light sharing among different forest strata for sustainable management of vegetation and regeneration. For. Achiev. Millenn. Goals 81–86.

Balzter, H., 2000. Markov chain models for vegetation dynamics. Ecol. Modell. 126, 139–154.

Barron-Gafford, G.A., Cable, J.M., Bentley, L.P., Scott, R.L., Huxman, T.E., Jenerette, G.D., Ogle, K., 2014. Quantifying the timescales over which exogenous and endogenous conditions affect soil respiration. New Phytol. 202, 442–454. https://doi.org/10.1111/nph.12675.

BayesFusion, L., 2017. GeNIe Modeler: User manual.

Blüthgen, N., Dormann, C.F., Prati, D., Klaus, V.H., Kleinebecker, T., Hölzel, N., Alt, F., Boch, S., Gockel, S., Hemp, A., Müller, J., Nieschulze, J., Renner, S.C., Schöning, I., Schumacher, U., Socher, S.A., Wells, K., Birkhofer, K., Buscot, F., Oelmann, Y., Rothenwöhrer, C., Scherber, C., Tscharntke, T., Weiner, C.N., Fischer, M., Kalko, E.K.V., Linsenmair, K.E., Schulze, E.-D., Weisser, W.W., 2012. A quantitative index of land-use intensity in grasslands: Integrating mowing, grazing and fertilization. Basic Appl. Ecol. 13, 207–220. https://doi.org/10.1016/j.baae.2012.04.001.

Bürgi, M., Östlund, L., Mladenoff, D.J., 2017. Legacy effects of human land use: ecosystems as time-lagged systems. Ecosystems 20, 94–103. https://doi.org/10.1007/s10021-016-0051-6.

Cable, J.M., Ogle, K., Barron-Gafford, G.A., Bentley, L.P., Cable, W.L., Scott, R.L., Williams, D.G., Huxman, T.E., 2013. Antecedent conditions influence soil respiration differences in shrub and grass patches. Ecosystems 16, 1230–1247. https://doi.org/10.1007/s10021-013-9679-7.

Carpinone, A., Giorgio, M., Langella, R., Testa, A., 2015. Markov chain modeling for very-short-term wind power forecasting. Electr. Power Syst. Res. 122, 152–158.

Chen, S.H., Pollino, C.A., 2012. Good practice in Bayesian network modelling. Environ. Model. Softw. 37, 134–145. https://doi.org/10.1016/j.envsoft.2012.03.012.

De Cáceres, M., Brotons, L., Aquilué, N., Fortin, M.J., 2013. The combined effects of land-use legacies and novel fire regimes on bird distributions in the Mediterranean. J. Biogeogr. 40, 1535–1547. https://doi.org/10.1111/jbi.12111.

De Frenne, P., Baeten, L., Graae, B.J., Brunet, J., Wulf, M., Orczewska, A., Kolb, A., Jansen, I., Jamoneau, A., Jacquemyn, H., Hermy, M., Diekmann, M., De Schrijver, A., De Sanctis, M., Decocq, G., Cousins, S.A.O., Verheyen, K., 2011. Interregional variation in the floristic recovery of post-agricultural forests. J. Ecol. 99, 600–609. https://doi.org/10.1111/j.1365-2745.2010.01768.x.

De Keersmaeker, L., Martens, L., Verheyen, K., Hermy, M., De Schrijver, A., Lust, N., 2004. Impact of soil fertility and isolation on diversity of herbaceous woodland species colonizing afforestations in Muizen Forest. Belgium. For. Ecol. Manage. 188, 291–304.

den Ouden, J., Muys, B., Mohren, F., Verheyen, K., 2010. Bosecologie en bosbeheer. Acco.

Dietrich, J.P., Schmitz, C., Müller, C., Fader, M., Lotze-Campen, H., Popp, A., 2012. Measuring agricultural land-use intensity – A global analysis using a model-assisted approach. Ecol. Modell. 232, 109–118. https://doi.org/10.1016/j.ecolmodel.2012.03.002.

Dlamini, W.M., 2010. A Bayesian belief network analysis of factors influencing wildfire occurrence in Swaziland. Environ. Model. Softw. 25, 199–208. https://doi.org/10.1016/j.envsoft.2009.08.002.

Dölle, M., Schmidt, W., 2009. Impact of tree species on nutrient and light availability: evidence from a permanent plot study of old-field succession. Plant Ecol. 203, 273–287. https://doi.org/10.1007/s11258-008-9547-2.

Druzdzel, M.J., 1999. SMILE: Structural Modeling, Inference, and Learning Engine and GeNIe: A Development Environment for Graphical Decision-Theoretic Models. In: Proceedings of the Sixteenth National Conference on Artificial Intelligence (AAAI-99). pp. 342–343.

Dupouey, J.L., Dambrine, E., Laffite, J.D., Moares, C., 2002. Irreversible impact of past land use on forest soils and biodiversity. Ecology 83, 2978–2984. https://doi.org/10.2307/3071833.

Essl, F., Dullinger, S., Rabitsch, W., Hulme, P.E., Pyšek, P., Wilson, J.R.U., Richardson, D.M., 2015. Historical legacies accumulate to shape future biodiversity in an era of rapid global change. Divers. Distrib. 21, 534–547. https://doi.org/10.1111/ddi.12312.

Flinn, K.M., Marks, P.L., 2007. Agricultural legacies in forest environments: tree communities, soil properties, and light availability. Ecol. Appl. 17, 452–463. https://doi.org/10.1890/05-1963.

Forestry Suppliers, 2008. Using Forest Densiometers.

Foster, D.R., Swanson, F., Aber, J., Burke, I., Brokaw, N., Tilman, D., Knapp, A., 2003. The importance of land-use legacies to ecology and conservation. Bioscience 53, 77–88. https://doi.org/10.1641/0006-3568(2003) 053[0077:TIOLUL]2.0.CO;2.

Genc, O., Dag, A., 2016. A Bayesian network-based data analytical approach to predict velocity distribution in small streams. J. Hydroinformatics 18, 466–480. https://doi.org/10.2166/hydro.2015.110.

Golroo, A., Ph, D., Eng, P., Tighe, S.L., 2012. Development of pervious concrete pavement performance models using expert opinionsJ. Transp. Eng. 138, 634–648. https://doi.org/10.1061/(ASCE)TE.1943-5436.0000356.

Hahn, P.G., Orrock, J.L., 2015. Land-use history alters contemporary insect herbivore community composition and decouples plant-herbivore relationships. J. Anim. Ecol. 84, 745–754. https://doi.org/10.1111/1365-2656.12311.

Harding, J.S., Benfield, E.F., Bolstad, P.V., Helfman, G.S., Jones, E.B.D., 1998. Stream biodiversity: the ghost of land use past. Proc. Natl. Acad. Sci. U.S.A. 95, 14843–14847. https://doi.org/10.1073/pnas.95.25.14843.

Hawkins, T.W., Ellis, A.W., 2010. The dependence of streamflow on antecedent subsurface moisture in an arid climate. J. Arid Environ. 74, 75–86. https://doi.org/10.1016/j.jaridenv.2009.07.003.

Hédl, R., Kopecký, M., Komárek, J., 2010. Half a century of succession in a temperate oakwood: from species-rich community to mesic forest. Divers. Distrib. 16, 267–276. https://doi.org/10.1111/j.1472-4642.2010.00637.x.

Horn, H.S., 1975. Markovian properties of forest successions. In: Cody, M.L., Diamond, J.M. (Eds.), Ecology and Evolution of. Communities. Belknap press, pp. 196–211.

Hugin, 2008. Hugin Researcher, Version 6.9: User's Guide. Hugin Expert, Aalborg, Denmark.

Jensen, F.V., Nielsen, T.D., 2007. Bayesian Networks and Decision Graphs, 2nd ed. Springer Science + Business Media LLC.

Kahl, T., Bauhus, J., 2014. An index of forest management intensity based on assessment of harvested tree volume, tree species composition and dead wood origin. Nat. Conserv. 7, 15–27. https://doi.org/10.3897/natureconservation.7.7281.

Kerr, G., Haufe, J., 2011. Thinning Practice – A Silvicultural Guide.

Klanderud, K., Vandvik, V., Goldberg, D., 2015. The importance of biotic vs. abiotic drivers of local plant community composition along regional bioclimatic gradients. PLoS One 10, 1–14. https://doi.org/10.1371/journal.pone.0130205.

Koerner, W., Dupouey, J.L., Dambrine, E., Benoit, M., 1997. Influence of past land use on the vegetation and soils of present day forest in the Vosges mountains. France. J. Ecol. 85, 351–358.

Kopecký, M., Hédl, R., Szabó, P., 2013. Non-random extinctions dominate plant community changes in abandoned coppices. J. Appl. Ecol. 50, 79–87. https://doi.org/10.1111/1365-2664.12010.

Kopecký, M., Vojta, J., 2009. Land use legacies in post-agricultural forests in the Doupovské Mountains, Czech Republic. Appl. Veg. Sci. 12, 251–260.

Kuhnert, P.M., Martin, T.G., Griffiths, S.P., 2010. A guide to eliciting and using expert knowledge in Bayesian ecological models. Ecol. Lett. 13, 900–914. https://doi.org/10.1111/j.1461-0248.2010.01477.x.

Kwiatkowska, M., Normann, G., Parker, D., 2011. PRISM 4.0: Verification of Embedded Real-time Systems. In: 23rd International Conference on Computer Aided Verification (CAV'11). Springer, pp. 585–591. https://doi.org/10.1007/978-3-658-09994-7_1.

Landuyt, D., Broekx, S., D'hondt, R., Engelen, G., Aertsens, J., Goethals, P.L.M., 2013. A

review of Bayesian belief networks in ecosystem service modelling. Environ. Model. Softw. 46, 1–11. https://doi.org/10.1016/j.envsoft.2013.03.011.
Lemmon, P.E., 1957. A new instrument for measuring forest overstory density. J. For. 55, 667–668.
Leuning, R., Cleugh, H.A., Zegelin, S.J., Hughes, D., 2005. Carbon and water fluxes over a temperate Eucalyptus forest and a tropical wet/dry savanna in Australia: measurements and comparison with MODIS remote sensing estimates. Agric. For. Meteorol. 129, 151–173. https://doi.org/10.1016/j.agrformet.2004.12.004.
Logofet, D.O., Lesnaya, E.V., 2000. The mathematics of Markov models: what Markov chains can really predict in forest successions. Ecol. Modell. 126, 285–298.
Lunt, I.D., Spooner, P.G., 2005. Using historical ecology to understand patterns of biodiversity in fragmented agricultural landscapes. J. Biogeogr. 32, 1859–1873. https://doi.org/10.1111/j.1365-2699.2005.01296.x.
Luyssaert, S., Hessenmöller, D., Von Lüpke, N., Kaiser, S., Schulze, E.D., 2011. Quantifying land use and disturbance intensity in forestry, based on the self-thinning relationship. Ecol. Appl. 21, 3272–3284. https://doi.org/10.1890/10-2395.1.
Máliš, F., Kopecký, M., Petřík, P., Vladovič, J., Merganič, J., Vida, T., 2016. Life-stage, not climate change, explains observed tree range shifts. Glob. Chang. Biol. 22, 1904–1914. https://doi.org/10.1111/gcb.13210.Life-stage.
Marcot, B.G., 2012. Metrics for evaluating performance and uncertainty of Bayesian network models. Ecol. Modell. 230, 50–62. https://doi.org/10.1016/j.ecolmodel.2012.01.013.
Marcot, B.G., Steventon, J.D., Sutherland, G.D., Mccann, R.K., 2006. Guidelines for developing and updating Bayesian belief networks applied to ecological modeling and conservation. Can. J. For. Res. 36, 3063–3074. https://doi.org/10.1139/X06-135.
McGrath, M.J., Luyssaert, S., Meyfroidt, P., Kaplan, J.O., Bürgi, M., Chen, Y., Erb, K., Gimmi, U., McInerney, D., Naudts, K., Otto, J., Pasztor, F., Ryder, J., Schelhaas, M.J., Valade, A., 2015. Reconstructing European forest management from 1600 to 2010. Biogeosciences 12, 4291–4316. https://doi.org/10.5194/bg-12-4291-2015.
Moranz, R.A., Debinski, D.M., McGranahan, D.A., Engle, D.M., Miller, J.R., 2012. Untangling the effects of fire, grazing, and land-use legacies on grassland butterfly communities. Biodivers. Conserv. 21, 2719–2746. https://doi.org/10.1007/s10531-012-0330-2.
Murphy, K.P., 2002. Dynamic Bayesian Networks: Representation, Inference and Learning.
Norsys, 1998. Netica Application User's Guide. Norsys Software Corporation, Vancouver, BC.
Oesterheld, M., Loreti, J., Semmartin, M., Sala, O.E., Oesterheld, M., Semmartin, M., 2001. Inter-annual variation in primary production of a semi-arid grassland related to previous-year production. J. Veg. Sci. 12, 137–142. https://doi.org/10.2307/3236681.
Ogle, K., Barber, J.J., Barron-Gafford, G.A., Bentley, L.P., Young, J.M., Huxman, T.E., Loik, M.E., Tissue, D.T., 2015. Quantifying ecological memory in plant and ecosystem processes. Ecol. Lett. 18, 221–235. https://doi.org/10.1111/ele.12399.
Padisák, J., 1992. Seasonal succession of phytoplankton in a large shallow lake (Balaton, Hungary) – A dynamic approach to ecological memory, its possible role and mechanisms. J. Ecol. 80, 217–230.
Parker, W.C., 2014. The relationship of stand structure with canopy transmittance: Simple models and practical methods for managing understory light conditions in eastern white pine (Pinus strobus L.) -dominated forests. For. Chron. 90, 489–497.
Perring, M.P., Bernhardt-Römermann, M., Baeten, L., Midolo, G., Blondeel, H., Depauw, L., Landuyt, D., Maes, S.L., De Lombaerde, E., Carón, M.M., Vellend, M., Brunet, J., Chudomelová, M., Decocq, G., Diekmann, M., Dirnböck, T., Dörfler, I., Durak, T., De Frenne, P., Gilliam, F.S., Hédl, R., Heinken, T., Hommel, P., Jaroszewicz, B., Kirby, K.J., Kopecký, M., Lenoir, J., Li, D., Máliš, F., Mitchell, F.J.G., Naaf, T., Newman, M., Petřík, P., Reczyńska, K., Schmidt, W., Standovár, T., Świerkosz, K., Van Calster, H., Vild, O., Wagner, E.R., Wulf, M., Verheyen, K., 2018. Global environmental change effects on plant community composition trajectories depend upon management legacies. Glob. Chang. Biol. 24, 1722–1740. https://doi.org/10.1111/gcb.14030.
Perring, M.P., De Frenne, P., Baeten, L., Maes, S.L., Depauw, L., Blondeel, H., Carón, M.M., Verheyen, K., 2016. Global environmental change effects on ecosystems: The importance of land-use legacies. Glob. Chang. Biol. 22, 1361–1371. https://doi.org/10.1111/gcb.13146.
Peterken, G.F., Game, M., 1984. Historical factors affecting the number and distribution of vascular plant species in the woodlands of central Lincolnshire. J. Ecol. 72, 155–182. https://doi.org/10.1111/j.1365-2664.2005.01116.x.
Peterson, G.D., 2002. Contagious disturbance, ecological memory, and the emergence of landscape pattern. Ecosystems 5, 329–338.
Pollino, C.A., Woodberry, O., Nicholson, A., Korb, K., Hart, B.T., 2007. Parameterisation and evaluation of a Bayesian network for use in an ecological risk assessment. Environ. Model. Softw. 22, 1140–1152. https://doi.org/10.1016/j.envsoft.2006.03.006.
R Core Team, 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.
Sala, O.E., Gherardi, L.A., Reichmann, L., Jobbágy, E., Peters, D., 2012. Legacies of precipitation fluctuations on primary production: theory and data synthesis. Philos. Trans. R. Soc. B 367, 3135–3144. https://doi.org/10.1098/rstb.2011.0347.
Schaefer, V., 2009. Alien invasions, ecological restoration in cities and the loss of ecological memory. Restoration Ecol. 17, 171–176. https://doi.org/10.1111/j.1526-100X.2008.00513.x.
Schall, P., Ammer, C., 2013. How to quantify forest management intensity in Central European forests. Eur. J. For. Res. 132, 379–396. https://doi.org/10.1007/s10342-013-0681-6.
Sercu, B.K., Baeten, L., van Coillie, F., Martel, A., Lens, L., Verheyen, K., Bonte, D., 2017. How tree species identity and diversity affect light transmittance to the understory in mature temperate forests. Ecol. Evol. 7, 10861–10870. https://doi.org/10.1002/ece3.3528.
Shamshad, A., Bawadi, M.A., Wan Hussin, W.M.A., Majid, T.A., Sanusi, S.A.M., 2005. First and second order Markov chain models for synthetic generation of wind speed time series. Energy 30, 693–708. https://doi.org/10.1016/j.energy.2004.05.026.
Smith, C.S., Howes, A.L., Price, B., McAlpine, C.A., 2007. Using a Bayesian belief network to predict suitable habitat of an endangered mammal – The Julia Creek dunnart (Sminthopsis douglasi). Biol. Conserv. 139, 333–347. https://doi.org/10.1016/j.biocon.2007.06.025.
Spedicato, G.A., 2017. Discrete time Markov chains with R. R J. 9 (2), 84–104. https://doi.org/10.1002/9781118477793.ch5.
Stewart, W.J., 1996. Marca: Markov Chain Analyzer - A Software Package for Markov Modelling Version 3.0.
Sun, Z., Ren, H., Schaefer, V., Lu, H., Wang, J., Li, L., Liu, N., 2013. Quantifying ecological memory during forest succession: a case study from lower subtropical forest ecosystems in South China. Ecol. Indic. 34, 192–203. https://doi.org/10.1016/j.ecolind.2013.05.010.
Theijssen, D., ten Bosch, L., Boves, L., Cranen, B., van Halteren, H., 2013. Choosing alternatives: Using Bayesian Networks and memory-based learning to study the dative alternation. Corpus Linguist. Linguist. Theory 9, 227–262. https://doi.org/10.1515/cllt-2013-0007.
Thomaes, A., 2014. Tree species effects on herb layer development in post-agricultural forests. PhD thesis.
Thomaes, A., De Keersmaeker, L., De Schrijver, A., Baeten, L., Vandekerkhove, K., Verstraeten, G., Verheyen, K., 2013. Can soil acidity and light help to explain tree species effects on forest herb layer performance in post-agricultural forests? Plant Soil 373, 183–199. https://doi.org/10.1007/s11104-013-1786-x.
Thomaes, A., De Keersmaeker, L., Verschelde, P., Vandekerkhove, K., Verheyen, K., 2014. Tree species determine the colonisation success of forest herbs in post-agricultural forests: results from a 9 yr introduction experiment. Biol. Conserv. 169, 238–247. https://doi.org/10.1016/j.biocon.2013.10.024.
Thomas, S.C., Halpern, C.B., Falk, D.A., Liguori, D.A., Austin, K.A., 1999. Plant diversity in managed forests: understory responses to thinning and fertilization. Ecol. Appl. 9, 864–879. https://doi.org/10.7326/0003-4819-136-9-200205070-00007.
Tinya, F., Ódor, P., 2016. Congruence of the spatial pattern of light and understory vegetation in an old-growth, temperate mixed forest. For. Ecol. Manage. 381, 84–92. https://doi.org/10.1016/j.foreco.2016.09.027.
Ujházy, K., Hederová, L., Máliš, F., Ujházyová, M., Bosela, M., Čiliak, M., 2017. Overstorey dynamics controls plant diversity in age-class temperate forests. For. Ecol. Manage. 391, 96–105. https://doi.org/10.1016/j.foreco.2017.02.010.
Usher, M., 1981. Modelling ecological succession, with particular reference to Markovian models. Plant Ecol. 46, 11–18. https://doi.org/10.1007/BF00118380.
Usher, M.B., 1979. Markovian approaches to ecological succession. J. 48, 413–426.
Van Calster, H., Baeten, L., Verheyen, K., De Keersmaeker, L., Dekeyser, S., Rogister, J.E., Hermy, M., 2008. Diverging effects of overstorey conversion scenarios on the understorey vegetation in a former coppice-with-standards forest. For. Ecol. Manage. 256, 519–528. https://doi.org/10.1016/j.foreco.2008.04.042.