Mapping hotspots of potential ecosystem fragility using commonly available spatial data

Alexandre Génin, Steven Lee, Eric Berlow, Steven Ostoja, Sonia Kéfi

To cite this version:

Alexandre Génin, Steven Lee, Eric Berlow, Steven Ostoja, Sonia Kéfi. Mapping hotspots of potential ecosystem fragility using commonly available spatial data. Biological Conservation, 2020, 241, pp.108388. 10.1016/j.biocon.2019.108388 . hal-02996937

HAL Id: hal-02996937
https://hal.umontpellier.fr/hal-02996937

Submitted on 7 Mar 2022

Mapping hotspots of potential ecosystem fragility using commonly available spatial data.

Alexandre Génin (1)*, Steven R. Lee (2), Eric L. Berlow (3), Steven M. Ostoja (4), Sonia Kéfi (1)

1. Institut des Sciences de l'Evolution, Université de Montpellier, CNRS, IRD, EPHE, CC 065, Place Eugène Bataillon, 34095 Montpellier Cedex 05, France

2. US Geological Survey, Western Ecological Research Center- Yosemite Field Station. Oakhurst, CA, USA

3. Vibrant Data Labs, Berkeley, CA, USA

4. USDA California Climate Hub, Agricultural Research Service, University of California Davis, John Muir Institute of the Environment. Davis, CA, USA

* corresponding author: alexandre.genin@umontpellier.fr. Office phone: +33 4 67 14 40 81. Fax: +33 4 67 14 40 61

Abstract

Effective conservation requires prioritizing areas that are vulnerable to large, irreversible changes. Unfortunately, rigorously documenting these changes with experiments and long-term monitoring is not only costly, but may provide evidence that is too late to facilitate proactive decisions.

We use a simple model to illustrate that commonly available short-term spatial, “snapshot”, data from a given ecosystem along an environmental gradient can be used to identify environmental conditions under which different ecosystem states (e.g. different species compositions) co-occur in space. These environmental conditions are those under which future perturbations have the potential for discontinuous large, sometimes irreversible, effects; and can be mapped in space to *predict* potential spatial hotspots of ecosystem fragility.

We apply these insights to ecologically important high-elevation subalpine meadows of the Sierra Nevada (California). Our analysis reveals specific areas within meadows that may be more vulnerable than others because their plant communities have the potential to shift to a different state. These shifts can be mechanistically explained by interactions between the vegetation and the local water regimes and/or the upper soil conditions.

Our study provides a simple workflow using commonly available data to help prioritize conservation areas based on their potential sensitivity to upcoming perturbations. Such an approach could be very valuable to make most efficient use of conservation and management resources in the context of ongoing global changes.

Keywords

global changes; management; subalpine ecosystems; alternative states; perturbations; spatial data

Introduction

While many ecosystems respond gradually to environmental changes, others can exhibit ecological shifts, where disproportional changes in an ecosystem state occur following gradual changes in environmental conditions (Scheffer et al. 2001, Suding et al. 2004). This represents a global conservation concern, as a wide array of terrestrial and marine ecosystems have shown such large degradation events, including coral reefs (Hughes 1994), fish stocks (Sguotti et al. 2019), arid ecosystems (Kéfi et al. 2007), shallow lakes (Scheffer et al. 1993) and kelp forests (Steneck et al. 2013). In general, because these systems have attracted a large research effort, the ecological mechanisms underpinning these shifts have been identified. Yet, this knowledge does not always translate into actionable conservation directives because local factors can alter ecological dynamics in seemingly idiosyncratic ways (Suding and Hobbs 2009). For example, in rangelands, erosion and fire regimes can tip ecosystems between grass and shrub-dominated states (Suding et al. 2004). Yet, because such dynamics are altered by local factors (such as topography or grazing pressure), it remains difficult to pinpoint the ecological conditions at which a state shift will occur for a given piece of land (Bestelmeyer 2006). Such uncertainty can severely hinder effective conservation efforts, because the potential of some areas to undergo large changes and become degraded can be left overlooked. Conversely, degraded areas that would respond favorably to restoration may be left unattended. For informed land management in the current context of global changes, it is important to identify, *a priori*, how different zones in a given area of interest (e.g. a conservation area) may respond to perturbations. Rigorously and precisely documenting large changes after they happen is too late. Approximately forecasting its potential to occur can help prioritize resources for proactive decisions.

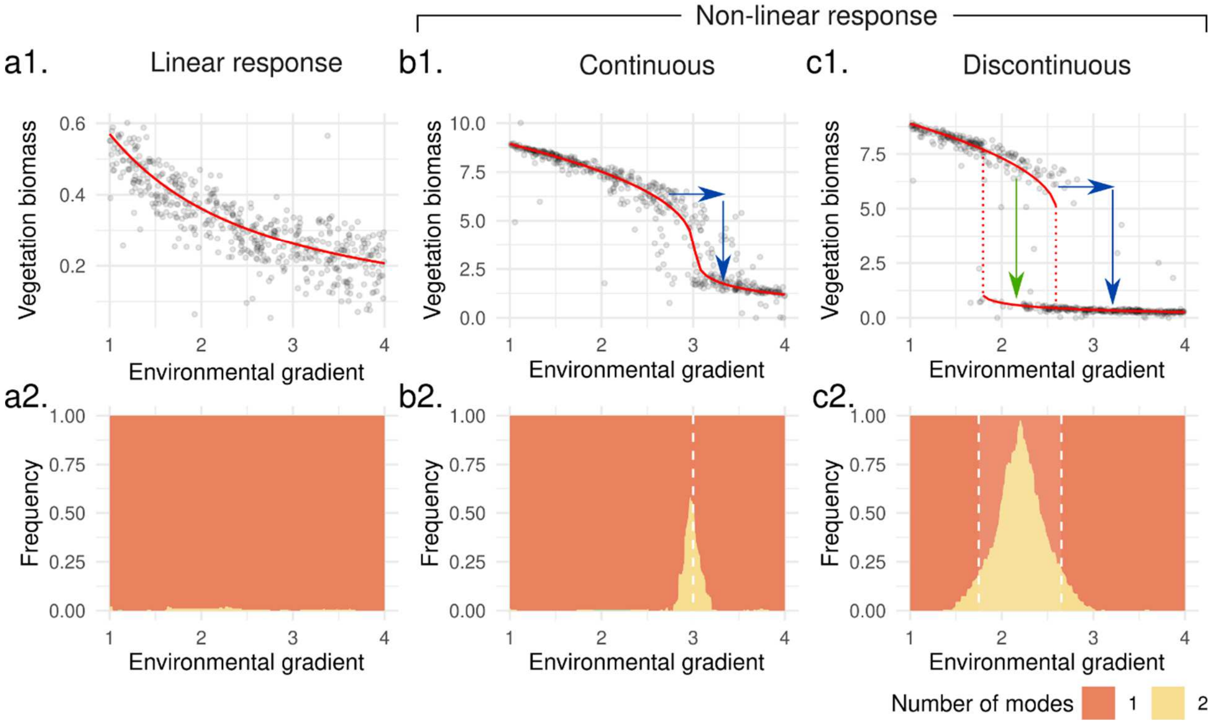
To improve the practical forecast of large changes in ecosystems, ecologists have sought to build more operational frameworks and tools to identify from empirical data if a given ecosystem can exhibit shifts, and if so, for which environmental conditions (Scheffer and Carpenter 2003, Andersen et al. 2009, Suding and Hobbs 2009, Ratajczak et al. 2018). However, not all methods can always be applied to a focal ecosystem. Ideally, efforts to identify upcoming shifts would involve documenting temporal responses to changes in environmental conditions, or perturbations (Bestelmeyer et al. 2011, Petraitis 2013), and then applying predictions based on those observations to new locations where environmental changes have yet to occur. Such an approach relies on time-series from either observations or experiments (Scheffer and Carpenter 2003). However, for many locations, obtaining such datasets, or setting up experiments at the appropriate scales can be costly and require perturbation experiments that are often incompatible with management policies. Conservation areas more commonly have inventory data, where descriptive attributes (e.g. abundance of species, soil characteristics, etc.) are surveyed along known environmental gradients (effectively replacing temporal sampling by spatial sampling -- a ‘space-for-time’ approach). These surveys are often not repeated, resulting in data collected from multiple locations along environmental gradients, but at a single or several close points in time (what we name hereafter ‘snapshot’ data). As an example, such data could correspond to plant surveys along altitudinal or aridity gradients during a single year. We show here how such data can help land managers prioritize their conservation efforts by predicting where perturbations *may* cause potentially large changes in an ecosystem.

Our work focuses on ecologically important subalpine meadows in the Sierra Nevada (California, USA). Many of these meadows are located in two major national parks, Yosemite and Sequoia, and exemplify conservation areas where, due to their remoteness, it is often impractical to regularly monitor for temporal changes, and the use of snapshot data to gain insight on the potential effect of perturbations can be particularly useful. The small-scale species composition of meadow plant communities is known to be tightly linked to water availability (Benedict 1983, Allen-Diaz 1991, Lowry et al. 2011). However, this general response can be altered, for example because of local changes in soil organic matter content (Ankenbauer and Loheide 2017), or because of a locally-increased importance of erosion processes (Klikoff 1965). Due to such local effects, some meadow areas may be more fragile than others to similar perturbations. We show in this study how a novel approach to the analysis of snapshot datasets may help better prioritize conservation and research efforts to protect these habitats.

Theoretically, three broad types of changes in a variable of interest (e.g. productivity) can be observed along an environmental gradient: linear, continuous non-linear or discontinuous non-linear with discrete alternative states (Box 1 panels a1, b1, c1) (Ratajczak et al. 2018). Such variable can be a population-level (e.g. abundance of a focal species), community-level (e.g. species composition) or a more ecosystemic attribute (e.g. productivity), depending on what is most relevant for management. When a linear response is observed, the observed state changes proportionally to changes in environmental conditions along the gradient (panel a1 in Box 1). In contrast, with a continuous non-linear response, small changes in conditions result in an apparent large shift as an ecological threshold is crossed (panel b1 in Box 1). A specific case arises when the response exhibits alternative states (panel c1 in Box 1), a case in which several discrete states can be observed for a range of environmental conditions along the gradient. These states can arise from differences in an environmental variable not captured by the main gradient (e.g. different local soil types), or because ecological processes make the ecosystem persist in one of several contrasted states despite similar environmental conditions. In this latter case, the discrete states are effectively alternative stable states, which can arise due to a variety of ecological mechanisms, such as priority effects, or reinforcing feedbacks (Beisner et al. 2003, Petraitis 2013). Regardless of the ecological origin of non-linear responses along a gradient, for practical conservation purposes, the areas where such responses are observed can be considered more fragile, either because they are susceptible to larger shifts than elsewhere, or because the community types in those areas strongly depend on a specific, narrower set of environmental conditions not reflected in the main gradient. Finding ways to identify where they occur is thus essential for informed management.

Previous studies have suggested approaches to identify the ranges of external conditions where discrete states are observed, based on detecting distinct ecosystem states in time series (Livina et al. 2010, Vasilakopoulos and Marshall 2015) or ecosystem data along spatial gradients (Hirota et al. 2011, Scheffer et al. 2012). These approaches have mostly been used for cases where a single, well-known index (e.g., tree cover), characterizes the observed ecosystem state. When the latter cannot be adequately characterized by a single index, it is more challenging to visualize discontinuous changes in data (Magurran et al. 2010), and fewer statistical tools are available to identify discrete states. Yet many cases of interest to both conservation and ecological theory require identifying discontinuous changes in complex multi-species communities. For example, when focusing on changes in species compositions, the observed state is accurately characterized by a set of species abundances, which cannot be reduced into a single index without losing information (Barros et al. 2016). We show here how non-linear responses can be detected along environmental gradients by applying such previous approaches to inventory data. We apply our work to detect areas in subalpine meadows that could be more fragile to upcoming perturbations, and discuss the implications of such fragility for the local conservation of habitats and species.

Box1: Identifying non-linear responses along a gradient from the modality of observations



Proof of concept model: to illustrate the link between multimodality in observations and the type of response of an ecosystem along a gradient, we can consider a simple model of ecosystem dynamics (Noy-Meir 1975, May 1977), which can produce different types of deterministic ecosystem responses (red lines) along an environmental gradient. When a linear response is observed (a1), the average ecosystem state (which could here represent the vegetation biomass; red line, y-axis) changes proportionally to the changes in environmental conditions (x-axis). This is not the case for non-linear responses (b1, c1), where large changes in the observed state can arise following small changes in the environmental conditions captured by the gradient (blue arrows), or between alternative states after a perturbation (green arrow).

We modified the model to have similar dynamics, but subjected to typical sources of ecological noise. We retained the final state of a set of independent simulations with varying parameters to simulate a sampling scheme that would lead to snapshot data (i.e. sampling independent instances of the same system in different environmental conditions, but at a single point in time). This dataset is represented by the black points in panels a1, b1, c1, each point corresponding to one observation. Using a moving-window approach, we can investigate how the distribution of these observations changes along the grazing gradient, and thereby characterize how their ‘modality’ relates to the type of ecosystem response (panels a2-c2). Wherever the observations tend to exhibit two modes along the gradient, their distribution is better described by a density with two modes, and thus the frequency of counts for two modes increases (see *Methods* for details about this approach). As this arises in the vicinity of non-linear responses (panels b2 and c2), modality can be used as a criterion to identify ranges of environmental conditions where apparent non-linear responses occur.

It is important to note that a non-linear response along a gradient does not mean that intrinsic ecological processes drive the response: external variables not captured by the main gradient could also explain such a pattern (e.g. different soil types producing alternative states). More information about the model is provided in Supplementary Material S2.

2 Methods

2.1 Data source

We used a dataset resulting from a study documenting meadow plant communities in 47 meadows located in Yosemite and Sequoia National Parks in 2011 and 2012 (Lee et al. 2017). Meadows were situated between 2517m and 3355m a.s.l. and had an area between 0.42ha and 22.1ha. Data were collected for each meadow from 2x2m plots, regularly spaced (20 m intervals) along belt transects that were oriented perpendicularly to the main meadow drainage, spaced 40 m apart and ran from meadow edge to meadow edge. Volumetric water content of the top 12 cm of soil was obtained for all plots using a handheld TDR soil moisture probe. In every third plot, ocular cover estimates of all plant species were recorded in 8 25x25cm sub-plots, and averaged to estimate the relative abundance of each species in the plot. In addition, ocular estimates of the cover of 28 variables describing the local water regime, upper soil characteristics and disturbance by herbivores were measured (e.g. cover of bare ground, rock bed, etc., see Supplementary Material S1). Overall, sampling resulted in 2860 plots with soil moisture and local environmental conditions; of these, 1287 also had species composition recorded.

We focused on the variations of a community-level attribute, the species composition of vegetation communities, along the natural moisture gradient in the meadows. Decades of work show that moisture is a key factor in determining meadow species composition, as within a meadow, plant species sort along the gradient (Allen-Diaz 1991, Lowry et al. 2011). We checked that this aspect of meadows was also captured by our dataset (Supplementary Material S2, section 2), and used water content as the main gradient. We attenuated the effect of year-to-year variations in precipitation on the measurements of the moisture gradient by standardizing the measured water content within year. As a result, in what follows, water content values are given in standard deviations (s.d.) to the annual mean.

2.2 Analyzing changes in modality along a gradient

When a set of observations is unimodal, observations tend to cluster around a single mean. For example, a unimodal set of observations of a given species' abundance can be effectively characterized by a mean abundance, from which most observations would deviate only slightly (Figure 1a). When a set of observations is multi-modal, observations cannot be well-summarized by a single mean and instead, several, different clusters are present in the data. This could for instance be the case when a species is either locally very rare or very abundant (these abundances being *modes*, Figure 1b), but rarely observed with intermediate abundances.

Characterizing the multimodality for a set of observations can be done using the density of observations. The density is a function that captures the relative frequency of different observed values: those that are more frequent in a given set of observations will have higher density values. Conversely, values that are scarcely observed correspond to lower density values. This results in local peaks ('modes') in the density around values that are frequently observed in a given set of observations. A unimodal set of observations will tend to have a single peak in its density, i.e. most observations cluster around a single, average value (Figure 1 a, c). A multi-modal set of observations will have a density with multiple well-defined peaks (Figure 1 b, d). Computing the density of a given set of observations, and estimating the number of well-separated peaks in it is thus a natural way to characterize the number of modes in a set of observations (its *modality*, Silverman (1981), Müller and Sawitzki (1991)).

For each subset of observations, we computed its density, and considered n_{cut} thresholds regularly spaced between zero and the maximum value of the density. For each threshold, we considered a "horizontal slice" of the density and counted the number of observed disjunct parts of the density above the threshold (λ-clusters *sensu* Müller and Sawitzki (1991), see Figure 1). This number of clusters provides an estimate of the number of well-separated modes in a given density above a certain threshold (Müller and Sawitzki 1991, Fisher 2001, Scott 2015). For a given density, i.e. a set of observations, this process yields a set of n_{cut} integer values, each corresponding to an estimated number of modes. How these counts vary along an environmental gradient informs about changes in the modality of the distribution of states. A typical example could be that of a density switching from unimodal to bimodal around a certain value of the environmental gradient, which would be reflected in a higher number of counts for two well-separated modes.

This non-parametric method can be applied to univariate densities (Figure 1, top; Box 1), which corresponds to the case where the state of the system is defined by a single value. It can also be applied to multivariate densities, obtained through multi-dimensional kernel density estimation (Figure 1b), which makes it applicable to cases where each observation is multi-dimensional (e.g. when each observation consists of abundances of individual species). It only depends on a single parameter (the density bandwidth) and it is more robust to outliers than previous approaches for the detection of multimodality (Fisher 2001). n_{cut} only affects the precision of computations, and was set to a large value to identify differences in height of the density (we used 512 here, which was enough to detect changes in modality along the gradient).

Given a fixed set of observations, this procedure allows estimating the number of modes in the density, but it needs to be adapted to investigate change in this number along a gradient. To do so, we adopted a moving-window approach. We chose a window size *w*, here corresponding to a range of water content values, and estimated the number of modes for each subset of observations falling within this range, each time ’moving’ the window by a small amount along the gradient. This allows determining how the number of modes (i.e. the *modality*) of observations change along an environmental gradient.

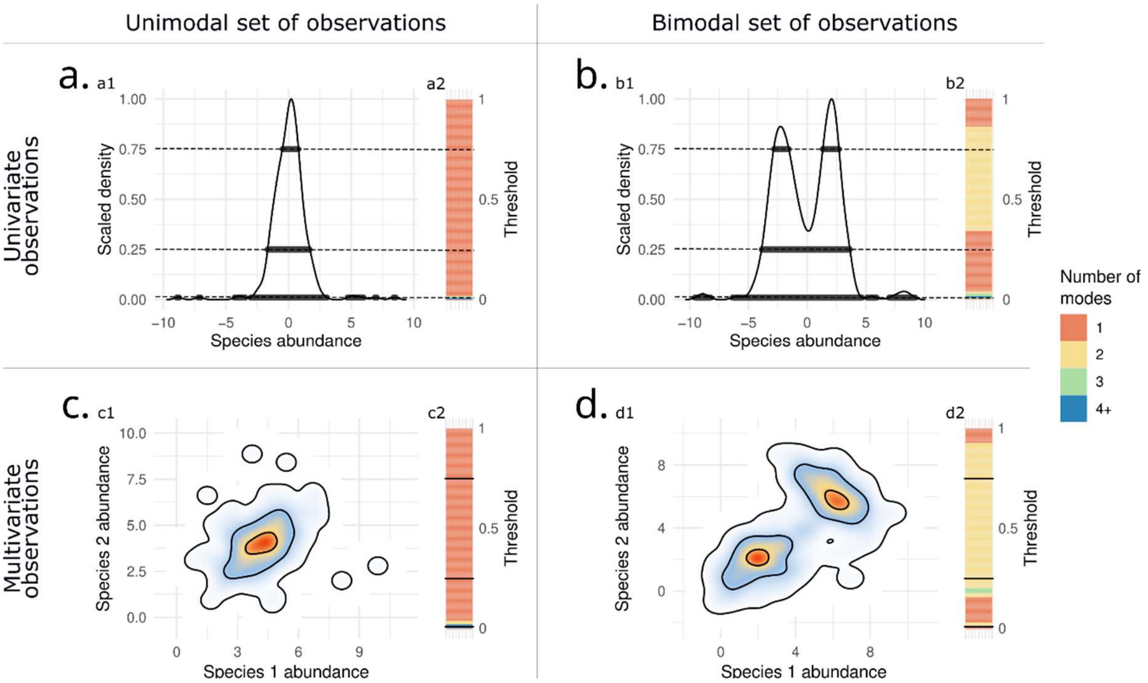


Figure 1: Principle of the measure of modality based on synthetic data. With univariate observations (i.e. that are described by a single variable such as the abundance of a single species), the density can be represented as a one-dimensional function of that variable (a1, b1). This function may have one peak, and thus be unimodal (a1) or two and be bimodal (b1). When an observation is multivariate (c1, d1), i.e. when two or more variables are used for each observation, such as abundances of several species, the density is represented as a two-dimensional plot (c1, d1) along these variables, where red values represent high density values. Again, multivariate densities can be unimodal (c1) or multi-modal (d1). To estimate the number of modes, we can consider the number of disjunct parts of a given density with values above a certain threshold (dashed lines, a1, b1 and continuous lines in c1 and d1). Doing so for a range of thresholds between zero and the maximum value of the density yields a majority of counts for one mode for unimodal densities (red area in panels a2, c2). For bimodal densities, the same analysis yields an increased number of counts for two modes (yellow areas, b2, d2). Note that very low thresholds identify spurious clusters because of outliers in the density, but this effect only concerns a reduced number of threshold values, so the influence of outliers on this estimate of modality is minimal. Data was generated

by drawing samples out of a univariate or multivariate Gaussian distribution (a,c), or a mixture of two Gaussian distributions (b,d; see Supplementary Material S2 section 5 for the distribution parameters).

2.3 Empirical data analyses

Classical univariate synthetic indices (*e.g.* total species richness or total cover) fail to capture the response of meadow vegetation communities to environmental parameters (Lee et al. 2017). We thus defined the state of meadow vegetation communities as the set of relative abundances of all species in a given plot (Barros et al. 2016). Each single observation had thus a dimension equal to the total number of species in the dataset (257). Because carrying kernel density estimation cannot be done on data with this many dimensions (typically, computational costs become too high after 5-10 dimensions (Blonder et al. 2014), we summarized the variations of vegetation species compositions to a few axes using an ordination. In principle, using Principal Coordinate Analysis (also named metric Multi-Dimensional Scaling) is the method of choice as it preserves dissimilarities between sites in the resulting axes. However, empirically it often produces strong arch artefacts on vegetation data, in particular when species turnover is high along a gradient (Legendre and Legendre 2012). Non-metric multidimensional scaling could be considered as an alternative but, because it preserves only the ranks of the original dissimilarities between sites, the resulting axes summarizing species turnover cannot be used quantitatively. We therefore chose a Correspondence-analysis-based method – Detrended-Correspondence Analysis – which (i) produces axes that can be used as quantitative variables and (ii) is less subject to arch-like artifacts. We carried out sensitivity analyses to confirm that our results were robust to the choice of dimensionality reduction method (Supplementary Material S2, section 4).

We used the two first axes of the DCA and measured changes in the modality of observations along the moisture gradient using the previously-described approach. We set the window size to 1 and used the bivariate diagonal density estimation kernel [0.70, 0; 0, 0.62], as obtained from automatic bandwidth selection (function *bw.nrd* in R). Because the sample size affects the shape of the density (and thus the counts of modes), we always computed densities using a fixed number of observations (150). Because there were often more samples falling within a window of water content, we repeated the analysis 100 times for each window, using a random subsample of 150 observations.

Based on results from the previous analysis, patterns identified two ranges of water content with increased support for two well-separated modes, dry areas and wet areas. We extracted these subsets from the dataset and investigated soil attributes that could underpin their more bimodal distributions. We refer to dry areas for sites where water content was below -1 s.d. and wet areas for those where water content was above 1 s.d. Because the multimodality analysis suggested a dominance of two discrete plant community types, we clustered each subset into two groups based on species composition, using hierarchical clustering with a Ward distance on a Bray-Curtis distance matrix. We contrasted the environmental covariates of each group to discuss possible underlying mechanisms for this apparent non-linear response along the water-content gradient. We used two-sample Mann-Whitney tests (*wilcox.test* in R) to contrast the distribution of each plot's characteristics across groups, adjusting p-values for multiple tests using the Benjamini-Hochberg correction.

Because the sampling was regularly spaced in meadows, we computed for each meadow an estimate of the proportion of area with increased multimodality as the proportion of plots laid in that meadow with a measured standardized water content above 1 or below -1. To map in space where this increase of modality occurred, we interpolated the standardized soil water content within meadows over a regular grid of points using ordinary kriging (using a spherical variogram for each meadow separately). We then modelled the empirical relationship between water content and the proportion of counts for more than one cluster as a one-dimensional generalized additive model. We used this empirical relationship to infer from the interpolated water content the proportion of counts for more than one cluster, and used it as an index of meadow areas with apparent non-linear responses (Supplementary Material S2, section 6). Analyses were conducted in R (version 3.6.1, R Core Team 2019), along with the R package 'vegan' v2.5 (Oksanen et al. 2018), 'gstat' v2.0.3 (Pebesma 2004) and 'mgcv' v1.8.28 (Wood 2004).

Results

We used a proof-of-concept model (Box 1) to investigate how the modality of observations, i.e. how many peaks in the distribution are identified, changes along gradients for different types of ecosystem responses: linear, non-linear continuous and discontinuous. Simulations show that, in the case of a single ecosystem state along a gradual/linear ecosystem response, the distribution is predominantly characterized by a single mode, while multimodality arises when the state changes non-linearly with environmental conditions (Box 1, panels b2, c2).

The vegetation composition of the sites of our data set was summarized to the two first axis of a Detrended-Correspondence Analysis (Figure 2). The analysis of changes in modality in DCA scores along the moisture gradient highlighted ranges of the moisture gradient with an increased modality in the species composition of the meadows (Figure 2): dry areas and wet areas. In these two ranges of the water gradient, we found an increased count for two well-defined modes (Figure 3b), highlighting an apparent non-linear response of species composition to changes in water content.

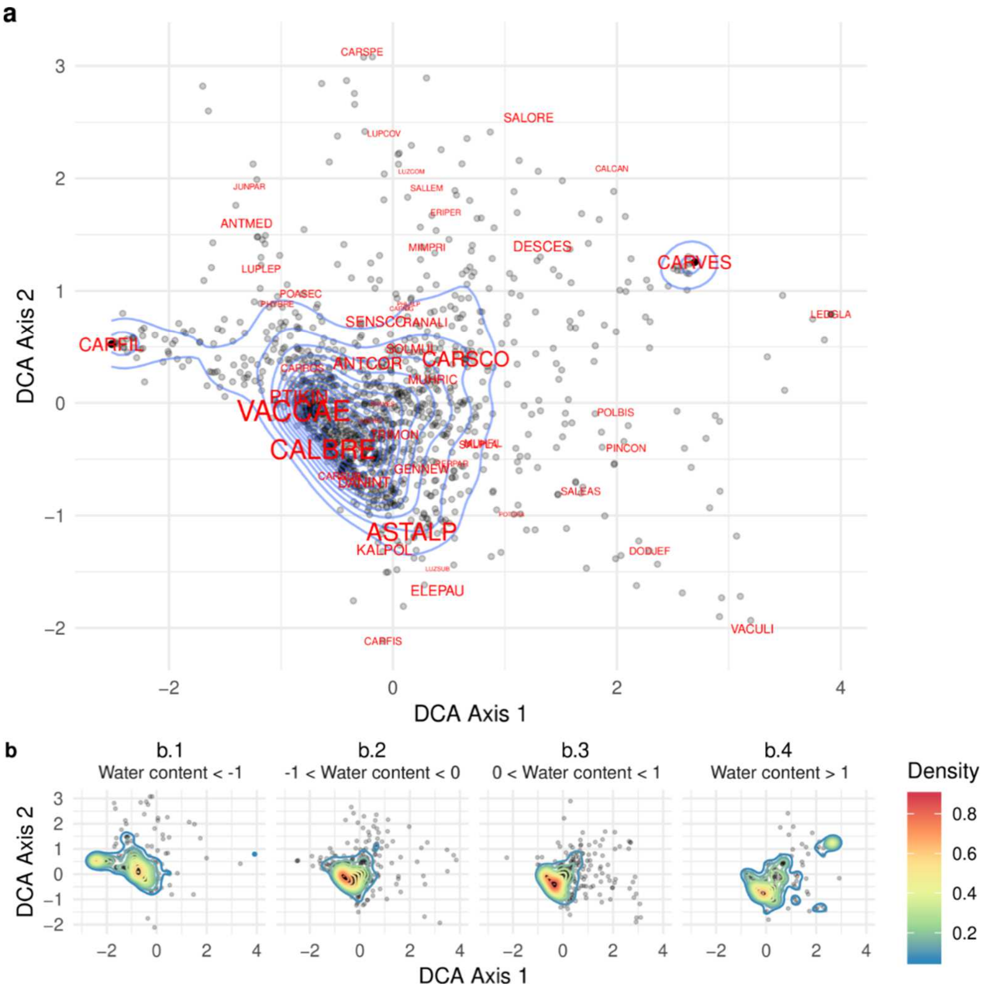


Figure 2: Detrended Correspondence Analysis results on the subalpine dataset (a). The distribution of each point in the two axes is based on similarity of species composition (two close points tend to have a more similar composition). Species labels indicate points (plots) where a given species is the most abundant, with font size being proportional to the overall abundance of the species in the dataset (see supplementary materialS2 for a list of species codes). The bottom panels (b1-4) show the same graph, but only with points comprised within different ranges of water content. Increased modality (i.e. multimodality) can be observed for water contents below -1 s.d. and above +1 s.d.

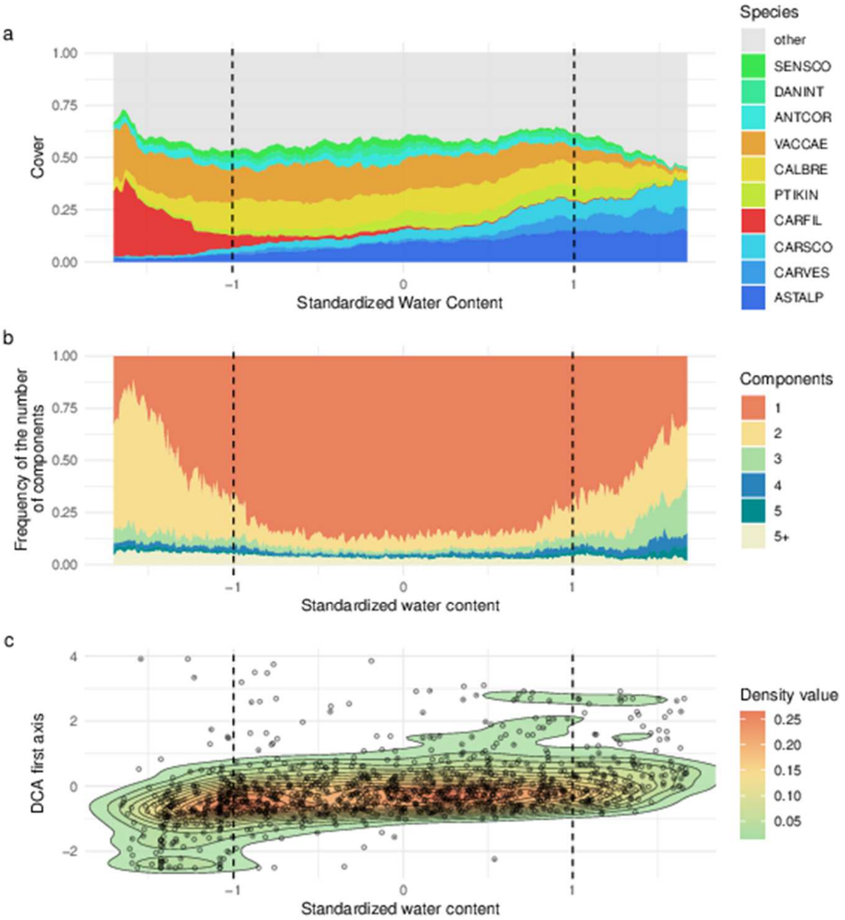


Figure 3: (a) Moving averages (window width: 0.4 s.d.) of the relative covers of the ten most abundant species in the subalpine meadow dataset along the gradient of soil water content. (b) results of modality analysis along the same gradient. (c) Changes in the first axis values of the Detrended Correspondence Analysis (DCA) along the water moisture gradient (each point corresponds to a summarized species composition), and density contour highlighting local modes.

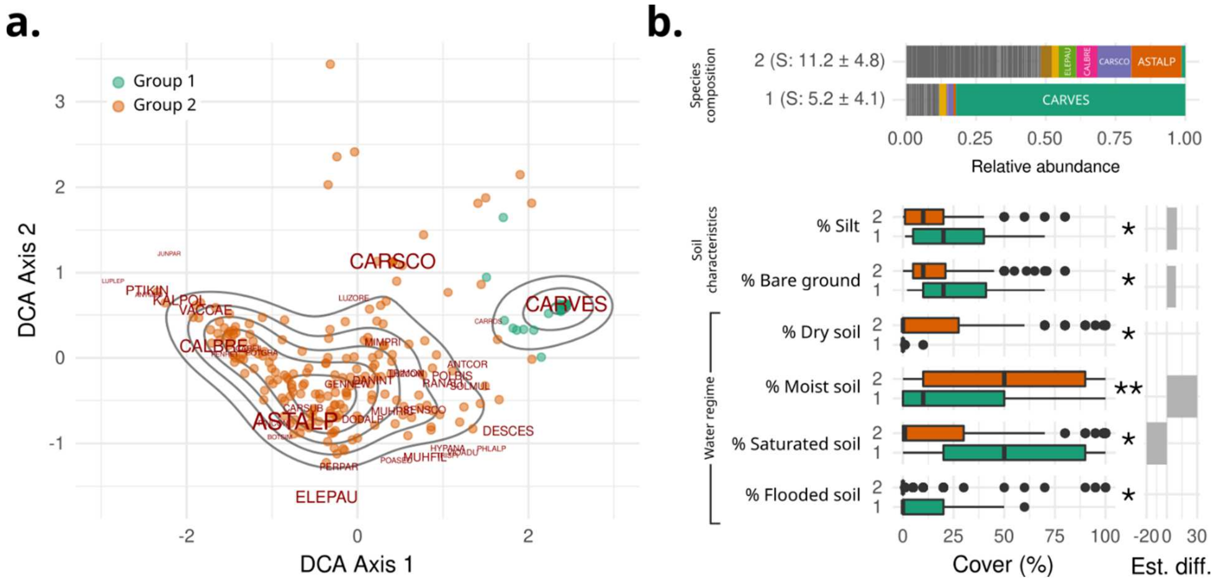
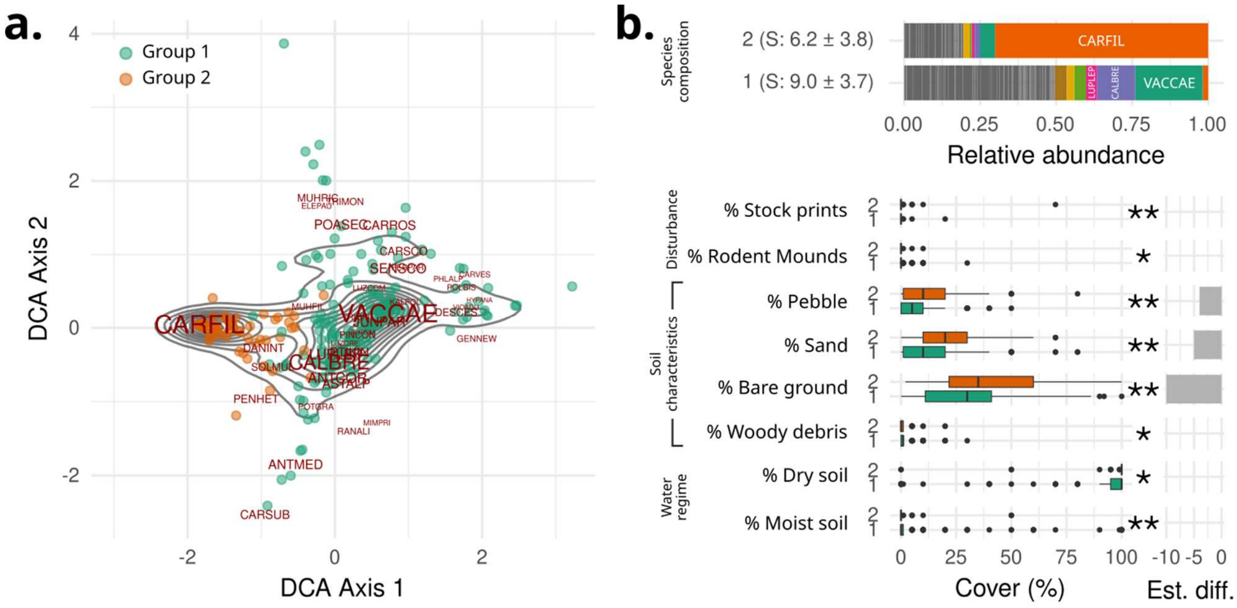


Figure 4: (a) Detrended Correspondence Analysis results on the wet meadow subset. Species labels indicate where individual species are dominant, with label size proportional to the total abundance of the species in the subset (see supplementary material S2). A density estimate is added on top of the points. (b) Differences in species composition, number of species *S*, and in plot attributes between the two groups (only attributes with significant differences are shown). Stars represent the significance of a Mann-Whitney test (one star for P < 0.05 and two for P < 0.01). The grey bars show the estimated difference in medians between the two groups.



.

Figure 5: (a) Detrended Correspondence Analysis results on the dry meadow subset. Species labels indicate where individual species are dominant, with label size proportional to the total abundance of the species in the subset. (b) Differences in species composition, number of species *S*, and in plot attributes between the two groups (only attributes with significant differences are shown), stars represent the significance of a Mann-Whitney test (one star for $P < 0.05$ and two for $P < 0.01$). The grey bars show the estimated difference in medians between the two groups.

In wet areas (plots with water content above 1 s.d.), the two groups defined based on species composition were contrasted. One was dominated by *Carex vesicaria (CARVES)*, while the other was a more mesic community that included *Oreostemma alpigenum* (Figure 4a, b-top). The *C. vesicaria*-dominated group represented 25% of all the plots in wet areas and had different water regime-related attributes: a more saturated and less moist soil, as well as a higher observed cover of silt and bare ground (Figure 4b). The covers of dry and flooded soils were significantly different between the two groups but the estimated effect size was close to zero.

In dry areas (plots with water content below -1 s.d.), one of the two groups was dominated by *Carex filifolia (CARFIL)* (10% of plots in dry areas), while the other was dominated by more mesic plants (e.g. *Vaccinium caespitosum (VACCAE)*; *Calmagrostis brewerii (CALBRE)*; Figure 5a). The *C. filifolia*-dominated group had a higher ocular cover of coarse particles (sand, pebble) and bare ground (Figure 5b). It also showed a significant difference in water-regime attributes (cover of dry and moist soil), and disturbance attributes (stock prints, rodent mounds), although they had a very small estimated cover difference (below 10^{-4}).

Overall, the proportion of area with increased multimodality in meadows represented on average 22% of a given meadow area of (across surveyed meadows, this number reached from 0 to 77%) and could be identified on maps based on the spatial interpolation of soil moisture (Figure 6).

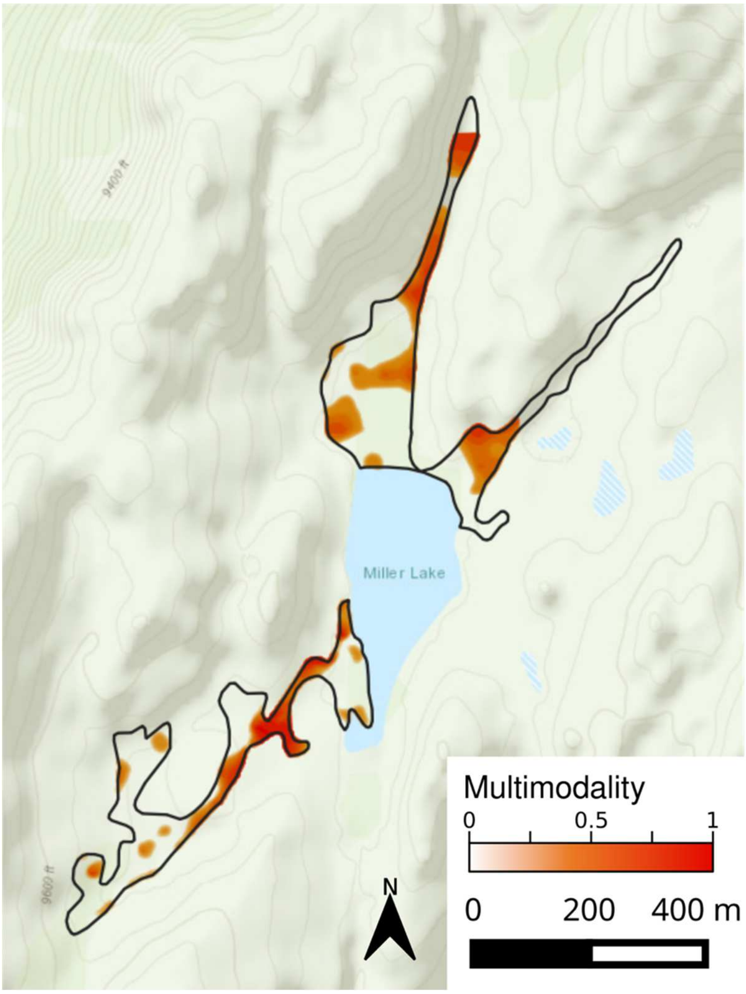


Figure 6: Example of spatial use of a multimodality index (the empirical support, between zero and one, for more than one community type for a given soil moisture level, see *Methods*). Zones in red represent areas where non-linear responses to changes in soil water content are present, and could thus be more fragile to perturbations.

Discussion

Our analysis revealed that in wet and dry areas of meadows, there was no smooth response of vegetation composition to the moisture gradient but instead, discrete vegetation types were present. These types were associated with contrasted local soil characteristics. As a result, current and upcoming perturbations may have an unexpected effect in these areas because general changes in moisture regime may alter the specific environmental factors on which vegetation states depend. Reviewing the factors that co-varied with species composition (Figure 4 and 5) helps identify specific environmental variables and potential ecological processes that are critical to the conservation of these habitats, and provide informed possible scenarios for their response to current perturbations in the Sierra Nevada.

Flagging potentially fragile meadow regions.

In wet areas of the meadows, a *Carex-vesicaria*-dominated community contrasted with a more diverse mesic-wet meadow community. While both communities co-occurred in regions of similar average surface soil moisture (as measured by 12 cm TDR probes); the *C. vesicaria*-dominated community had a larger amount of flooded or saturated soil and more silt, which are typical characteristics of streamside flooded pools where *Carex spp.* dominate (Baldwin and Hickman 2012). The mesic-wet community had a more diverse species composition, and occurred more frequently outside of areas with standing water. Here, the discrete vegetation states seemed to depend on the absence or presence of standing water, despite the fact that average, empirically measured surface soil moisture was similar. Local changes in water regimes could thus have a strong effect on species composition – or vice versa. These local changes can arise following changes in regional climate (see below), but may also be caused by local factors. Grazing for example has been shown to increase run-off through compaction of the soil (Kauffman and Krueger 1984, Ostoja et al. 2014), which could diminish the probability of creation of areas of standing water in meadows. However, given the overall low levels of grazing, and the relatively small importance of stock disturbance in our dataset (Figure 4, 5), their potential impacts are most likely minor within parks compared to regional-scale perturbations (Holmquist et al. 2014, Lee et al. 2017).

In dry meadow areas, our analysis identified a *Carex filifolia*-dominated community contrasting against a more diverse mesic community which both occur in areas of similar empirically measured soil moisture. The *C. filifolia* community was observed to have coarser upper soil particles (sand, pebble), and a higher amount of bare ground. The two vegetation states differed significantly in visible disturbance by herbivores and water regime. This difference is consistent with the fact that dry communities are known to respond to external parameters such as rodent disturbance, erosion processes and soil nutrients (Klikoff 1965). Differences in local soil surface conditions that are not captured by the main gradient could thus explain the presence of different discrete states. However, these local soil conditions are likely tied to the vegetation state as plant-soil feedbacks have been shown to exist in these areas. For example, sods formed by mesic meadow species can improve local microhabitat conditions by trapping smaller soil particles which hold nutrients and water during the dry season (Wood 1975, Ankenbauer and Loheide 2017). These plant-soil feedbacks could be strong enough to impair a rapid regrowth of the vegetation after soil disturbance, so that an acute disturbance of dry areas could have much longer-term effects than expected (e.g. through trampling). It is unknown whether such feedbacks could be strong enough to impair the vegetation regrowth completely following disturbance (effectively underpinning alternative stable states). However, the evidence as a whole highlights the need to adopt a conservation perspective that takes into account the possibility of long-term, poorly-reversible degradation: based on the available evidence, mesic communities in dry areas are likely to be very fragile to upcoming changes in conditions and recovering from acute degradation following an extreme event would likely take decades, if possible at all without restoration efforts (Ratliff and Westfall 1992).

More broadly, in both wet and dry areas, soil water-related parameters stood out as possible drivers of vegetation shifts. With the increase in frequency and intensity of drought events associated with climate change, snow to rain ratios and total spring snowpack are decreasing across the Sierra Nevada (Knowles et al. 2006, Barnett et al. 2008). These long-term changes in climate dynamics have the potential to extensively alter meadow hydrology, to which the wet and dry plant communities appear to be the most sensitive. As these communities represent on average 20% of meadow areas (up to 77%), this could constitute a significant concern for subalpine meadows across the central and southern Sierra Nevada. The Sierra Nevada has recently experienced an intense drought event (Belmecheri et al. 2015), which included the four driest years (2012-2015) of the last 2000 years (Adams et al. 2015). While the repercussions of the drought have clearly manifested across Sierra Nevada forest ecosystems (Potter 2016), the impacts on meadow ecosystems have yet to be thoroughly evaluated. Monitoring changes and setting up experimental research in potentially fragile areas occur could inform us greatly about such events. Our approach allows planning this next step by creating maps identifying these areas.

A workflow to identify and map fragile areas

With this work, we showed that investigating snapshot data for changes in modality of ecosystem states can provide information on the potential fragility of specific meadow areas. Importantly, our results show that we may identify these areas more accurately by taking into account the full state of the community (here, species composition) as we would not have detected changes in modality based on simple indices such as richness or total plant cover (Supplementary Material S2). It is a relatively simple analysis that can be carried out on commonly-available inventory data to investigate the possible dynamics of the ecological system under focus. An example workflow could be the following:

1. Gather a dataset of observations describing how the state of an ecosystem varies along known influential environmental gradients. This can be done over a relatively short period of time, but the sample size of observations must be sufficiently large (at least 10-50 points per level of the environmental gradient). If no prior knowledge is available to identify main environmental gradients, this can be done

based on redundancy analysis or canonical correspondence analysis (Supplementary Material S2, section 2).

2. When the state is described using a high number of variables (e.g. species abundances), reduce the dimensionality of the dataset (e.g. using Detrended Correspondence Analysis) to 1-5 dimensions (variables). It is noteworthy that reducing the dataset to 2 dimensions or less drastically speeds up the numerical computations.

3. Analyze the dataset for changes in modality along the main gradient, to identify areas along the gradient with increased modality, i.e. where apparent non-linear responses are present

4. For these areas, use local environmental data to identify the source of increased multimodality. Such increase may arise as a combination of (1) effects of secondary environmental drivers of plant community composition (e.g. different soil types) and (2) intrinsic ecological processes driving non-linear responses (e.g. reinforcing feedbacks, Wilson and Agnew 1992). For conservation purposes, in both cases, these areas may be considered fragile because the local state (e.g. species composition) depends on a narrower set of environmental conditions and/or the response to perturbations may be non-linear (possibly exhibiting ecological shifts).

5. Use spatial data to map where these areas are present, *i.e.* where the effect of perturbations may be larger. It is important to note that identifying the ranges of environmental conditions where non-linear responses occur only requires a reduced subset of data. For example, in our work, only 45% of the plots had a recorded species composition. However, the measurement of water content was carried out for a much higher number of plots, which allowed for higher resolution in our maps of community fragility.

The measurement of modality is an open area of statistical research (Scott 2015), and other approaches, not based on the use of a density, could be useful to evaluate multimodality in a set of observations (Clarke et al. 2008) and determine the statistical significance of patterns. Computing multivariate densities can be computation intensive, although approximate numerical methods may help (Blonder et al. 2014). Comparing different methods is beyond the scope of this article, but it is important to carry out sensitivity analyses to test for the robustness of the result (e.g. to the bandwidth used, or the method of ordination to reduce dimensionality, Supplementary Information S2, section 4).

Detecting large ecological changes for conservation

The observation of discrete community states does not mean that a given ecosystem necessarily exhibits irreversible shifts between alternative stable states. For example, discrete states can be also due to changes in an environmental parameter not captured by the ones measured. This is probably the case for the different community states observed on the wetter end of the gradient in the data set studied here. Nonetheless, complementing this approach with other sources of evidence can suggest where non-linear and/or poorly-reversible degradation may be possible in a given ecosystem of interest. Such additional knowledge can come from previously-known ecological mechanisms which drive irreversible shifts (*e.g.* known reinforcing feedbacks between vegetation and soil quality like here), historical information about shifts that occurred in the past and appear irreversible, or evidence from independent approaches (*e.g.* experiments, or based on indicators of ecosystem shifts; Dakos et al. 2012, Kéfi et al. 2014, Nijp et al. 2019). While no approach by itself is likely to prove the existence of irreversible shifts, from a conservation perspective it is particularly important to consider this possibility when the alternative community is considered to be degraded (e.g. because it is poorer in species or because it represents a loss of conservation-critical habitat). It is better to imprecisely predict a poorly-reversible shift than to precisely document it after it happens.

As the conditions for the emergence of alternative stable states are quite restrictive, ecosystems probably respond more often to gradual changes in a gradual way than in an abrupt way (e.g. because of spatial heterogeneity that favors gradual responses (van Nes and Scheffer 2005)). Cases of irreversible shifts have however been shown to be possible in a broad range of ecosystems and in a response to a number of drivers (lake eutrophication due to increase nutrient loading, coral reef degradation due to warming and/or water pollution, dryland desertification due to increased grazing and/or droughts). Because abrupt ecosystem responses can happen in about any ecosystems and their occurrence depends on a joint set of conditions which are not straightforward to monitor, it is extremely useful to find approaches that allow rapid predictions of where they are most likely to happen. The approach proposed here is promising because it provides a way to do so for species-rich communities (where most theoretical developments so far focus on species-poor models), and because it can be based on data typically gathered in conservation/protected areas.

References

Adams, K. D., R. M. Negrini, E. R. Cook, and S. Rajagopal. 2015. Annually resolved late Holocene paleohydrology of the southern Sierra Nevada and Tulare Lake, California. Water Resources Research 51:9708–9724.

Allen-Diaz, B. H. 1991. Water Table and Plant Species Relationships in Sierra Nevada Meadows. American Midland Naturalist 126:30.

Andersen, T., J. Carstensen, E. Hernández-García, and C. M. Duarte. 2009. Ecological thresholds and regime shifts: approaches to identification. Trends in Ecology & Evolution 24:49–57.

Ankenbauer, K. J., and S. P. Loheide. 2017. The effects of soil organic matter on soil water retention and plant water use in a meadow of the Sierra Nevada, CA: Soil organic matter affects plant water use. Hydrological Processes 31:891–901.

Baldwin, B. G., and J. C. Hickman, editors. 2012. The Jepson manual: vascular plants of California. 2. ed. University of California Press, Berkeley, Calif.

Barnett, T. P., D. W. Pierce, H. G. Hidalgo, C. Bonfils, B. D. Santer, T. Das, G. Bala, A. W. Wood, T. Nozawa, A. A. Mirin, D. R. Cayan, and M. D. Dettinger. 2008. Human-Induced Changes in the Hydrology of the Western United States. Science 319:1080–1083.

Barros, C., W. Thuiller, D. Georges, I. Boulangeat, and T. Münkemüller. 2016. N-dimensional hypervolumes to study stability of complex ecosystems. Ecology Letters 19:729–742.

Beisner, B., D. Haydon, and K. Cuddington. 2003. Alternative stable states in ecology. Frontiers in Ecology and the Environment 1:376–382.

Belmecheri, S., F. Babst, E. R. Wahl, D. W. Stahle, and V. Trouet. 2015. Multi-century evaluation of Sierra Nevada snowpack. Nature Climate Change 6:2–3.

Benedict, N. B. 1983. Plant Associations of Subalpine Meadows, Sequoia National Park, California. Arctic and Alpine Research 15:383.

Bestelmeyer, B. T. 2006. Threshold concepts and their use in rangeland management and restoration: the good, the bad, and the insidious. Restoration Ecology 14:325–329.

Bestelmeyer, B. T., A. M. Ellison, W. R. Fraser, K. B. Gorman, S. J. Holbrook, C. M. Laney, M. D. Ohman, D. P. C. Peters, F. C. Pillsbury, A. Rassweiler, R. J. Schmitt, and S. Sharma. 2011. Analysis of abrupt transitions in ecological systems. Ecosphere 2:art129.

Blonder, B., C. Lamanna, C. Violle, and B. J. Enquist. 2014. The n-dimensional hypervolume. Global Ecology and Biogeography 23:595–609.

Clarke, K. R., P. J. Somerfield, and R. N. Gorley. 2008. Testing of null hypotheses in exploratory community analyses: similarity profiles and biota-environment linkage. Journal of Experimental Marine Biology and Ecology 366:56–69.

Dakos, V., S. R. Carpenter, W. A. Brock, A. M. Ellison, V. Guttal, A. R. Ives, S. Kéfi, V. Livina, D. A. Seekell, E. H. van Nes, and M. Scheffer. 2012. Methods for Detecting Early Warnings of Critical Transitions in Time Series Illustrated Using Simulated Ecological Data. PLoS ONE 7:e41010.

Fisher, N. I. 2001. Mode testing via the excess mass estimate. Biometrika 88:499–517.

Hirota, M., M. Holmgren, E. H. Van Nes, and M. Scheffer. 2011. Global Resilience of Tropical Forest and Savanna to Critical Transitions. Science 334:232–235.

Holmquist, J. G., J. Schmidt-Gengenbach, and E. A. Ballenger. 2014. Patch-Scale Effects of Equine Disturbance on Arthropod Assemblages and Vegetation Structure in Subalpine Wetlands. Environmental Management 53:1109–1118.

Hughes, T. P. 1994. Catastrophes, Phase Shifts, and Large-Scale Degradation of a Caribbean Coral Reef. Science 265:1547–1551.

Kauffman, J. B., and W. C. Krueger. 1984. Livestock Impacts on Riparian Ecosystems and Streamside Management Implications... A Review. Journal of Range Management 37:430.

Kéfi, S., V. Guttal, W. A. Brock, S. R. Carpenter, A. M. Ellison, V. N. Livina, D. A. Seekell, M. Scheffer, E. H. van Nes, and V. Dakos. 2014. Early Warning Signals of Ecological Transitions: Methods for Spatial Patterns. PLoS ONE 9:e92097.

Kéfi, S., M. Rietkerk, C. L. Alados, Y. Pueyo, V. P. Papanastasis, A. ElAich, and P. C. de Ruiter. 2007. Spatial vegetation patterns and imminent desertification in Mediterranean arid ecosystems. Nature 449:213–217.

Klikoff, L. G. 1965. Microenvironmental Influence on Vegetational Pattern near Timberline in the Central Sierra Nevada. Ecological Monographs 35:187–211.

Knowles, N., M. D. Dettinger, and D. R. Cayan. 2006. Trends in Snowfall versus Rainfall in the Western United States. Journal of Climate 19:4545–4559.

Lee, S. R., E. L. Berlow, S. M. Ostoja, M. L. Brooks, A. Génin, J. R. Matchett, and S. C. Hart. 2017. A multi-scale evaluation of pack stock effects on subalpine meadow plant communities in the Sierra Nevada. PLOS ONE 12:e0178536.

Legendre, P., and L. Legendre. 2012. Numerical ecology. Third English edition. Elsevier, Amsterdam.

Livina, V. N., F. Kwasniok, and T. M. Lenton. 2010. Potential analysis reveals changing number of climate states during the last 60 kyr. Climate of the Past 6:77–82.

Lowry, C. S., S. P. Loheide, C. E. Moore, and J. D. Lundquist. 2011. Groundwater controls on vegetation composition and patterning in mountain meadows. Water Resources Research 47:n/a-n/a.

Magurran, A. E., S. R. Baillie, S. T. Buckland, J. McP. Dick, D. A. Elston, E. M. Scott, R. I. Smith, P. J. Somerfield, and A. D. Watt. 2010. Long-term datasets in biodiversity research and monitoring: assessing change in ecological communities through time. Trends in Ecology & Evolution 25:574–582.

May, R. 1977. Thresholds and breakpoints in ecosystems with a multiplicity of stable states. Nature 269:471–477.

Müller, D. W., and G. Sawitzki. 1991. Excess Mass Estimates and Tests for Multimodality. Journal of the American Statistical Association 86:738–746.

van Nes, E. H., and M. Scheffer. 2005. Implications of spatial heterogeneity for catastrophic regime shifts in ecosystems. Ecology 86:1797–1807.

Nijp, J. J., A. J. A. M. Temme, G. A. K. Voorn, L. Kooistra, G. M. Hengeveld, M. B. Soons, A. J. Teuling, and J. Wallinga. 2019. Spatial early warning signals for impending regime shifts: A practical framework for application in real-world landscapes. Global Change Biology 25:1905–1921.

Noy-Meir, I. 1975. Stability of Grazing Systems: An Application of Predator-Prey Graphs. The Journal of Ecology 63:459.

Oksanen, J., F. G. Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlinn, P. R. Minchin, R. B. O'Hara, G. L. Simpson, P. Solymos, M. H. H. Stevens, E. Szoecs, and H. Wagner. 2018. vegan: Community Ecology Package.

Ostoja, S. M., M. L. Brooks, P. E. Moore, E. L. Berlow, R. Blank, J. Roche, J. Chase, and S. Haultain. 2014. Potential environmental effects of pack stock on meadow ecosystems of the Sierra Nevada, USA. The Rangeland Journal 36:411.

Pebesma, E. J. 2004. Multivariable geostatistics in S: the gstat package. Computers & Geosciences 30:683–691.

Petraitis, P. 2013. Multiple stable states in natural ecosystems. Oxford University Press.

Potter, C. S. 2016. Landsat Image Analysis of Tree Mortality in the Southern Sierra Nevada Region of California during the 2013-2015 Drought. Journal of Earth Science & Climatic Change.

R Core Team. 2019. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Ratajczak, Z., S. R. Carpenter, A. R. Ives, C. J. Kucharik, T. Ramiadantsoa, M. A. Stegner, J. W. Williams, J. Zhang, and M. G. Turner. 2018. Abrupt Change in Ecological Systems: Inference and Diagnosis. Trends in Ecology & Evolution.

Ratliff, R. D., and S. E. Westfall. 1992. Restoring plant cover on high-elevation gravel areas, Sequoia National Park, California. Biological Conservation 60:189–195.

Scheffer, M., S. Carpenter, J. A. Foley, C. Folke, and B. Walker. 2001. Catastrophic shifts in ecosystems. Nature 413:591–596.

Scheffer, M., and S. R. Carpenter. 2003. Catastrophic regime shifts in ecosystems: linking theory to observation. Trends in Ecology & Evolution 18:648–656.

Scheffer, M., M. Hirota, M. Holmgren, E. H. Van Nes, and F. S. Chapin. 2012. Thresholds for boreal biome transitions. Proceedings of the National Academy of Sciences 109:21384–21389.

Scheffer, M., S. H. Hosper, M.-L. Meijer, B. Moss, and E. Jeppesen. 1993. Alternative equilibria in shallow lakes. Trends in Ecology & Evolution 8:275–279.

Scott, D. W. 2015. Multivariate density estimation: theory, practice, and visualization. Second edition. John Wiley & Sons, Inc, Hoboken, New Jersey.

Sguotti, C., S. A. Otto, R. Frelat, T. J. Langbehn, M. P. Ryberg, M. Lindegren, J. M. Durant, N. Chr. Stenseth, and C. Möllmann. 2019. Catastrophic dynamics limit Atlantic cod recovery. Proceedings of the Royal Society B: Biological Sciences 286:20182877.

Silverman, B. W. 1981. Using Kernel Density Estimates to Investigate Multimodality. Journal of the Royal Statistical Society. Series B (Methodological) 43:97–99.

Steneck, R. S., A. Leland, D. C. McNaught, and J. Vavrinec. 2013. Ecosystem Flips, Locks, and Feedbacks: the Lasting Effects of Fisheries on Maine’s Kelp Forest Ecosystem. Bulletin of Marine Science 89:31–55.

Suding, K. N., K. L. Gross, and G. R. Houseman. 2004. Alternative states and positive feedbacks in restoration ecology. Trends in Ecology & Evolution 19:46–53.

Suding, K. N., and R. J. Hobbs. 2009. Threshold models in restoration and conservation: a developing framework. Trends in Ecology & Evolution 24:271–279.

Vasilakopoulos, P., and C. T. Marshall. 2015. Resilience and tipping points of an exploited fish population over six decades. Global Change Biology 21:1834–1847.

Wilson, J. B., and A. D. Q. Agnew. 1992. Positive-feedback Switches in Plant Communities. Pages 263–336 *in* M. B. and A. H. Fitter, editor. Advances in Ecological Research. Academic Press.

Wood, S. H. 1975. Holocene stratigraphy and chronology of Mountain Meadows, Sierra Nevada, California. phd, California Institute of Technology.

Wood, S. N. 2004. Stable and efficient multiple smoothing parameter estimation for generalized additive models. Journal of the American Statistical Association 99:673–686.