Trajectory analysis in community ecology

Miquel De Cáceres [ID],[1,2,10] Lluís Coll,[1,2,3] Pierre Legendre,[4] Robert B. Allen,[5] Susan K. Wiser,[6] Marie-Josée Fortin,[7] Richard Condit,[8] and Stephen Hubbell[8,9]

[1]*Forest Sciences Center of Catalonia (CTFC), Carretera de Sant Llorenç, km.2 Solsona, Catalonia E-25280 Spain*
[2]*Center for Ecological Research and Forestry Applications (CREAF), Cerdanyola del Vallès, Catalonia E-08193 Spain*
[3]*Department of Agriculture and Forest Engineering (EAGROF), University of Lleida, Lleida E-25198 Spain*
[4]*Département de Sciences Biologiques, Université de Montréal, CP 6128, Succursale Centre-ville, Montreal, Quebec H3C3J7 Canada*
[5]*Independent Researcher, 8 Roblyn Place, Lincoln 7608 New Zealand*
[6]*Manaaki Whenua, Landcare Research, P.O. Box 40, Lincoln 7640 New Zealand*
[7]*Department of Ecology and Evolutionary Biology, University of Toronto, 25 Willcocks Street, Toronto, Ontario M5S3B2 Canada*
[8]*Smithsonian Tropical Research Institute, Box 0843-03092, Balboa, Ancon, Panama*
[9]*Ecology and Evolutionary Biology, University of California, Los Angeles, California 90095, USA*

Abstract. Ecologists have long been interested in how communities change over time. Addressing questions about community dynamics requires ways of representing and comparing the variety of dynamics observed across space. Until now, most analytical frameworks have been based on the comparison of synchronous observations across sites and between repeated surveys. An alternative perspective considers community dynamics as trajectories in a chosen space of community resemblance and utilizes trajectories as objects to be analyzed and compared using their geometry. While methods that take this second perspective exist, for example to test for particular trajectory shapes, there is a need for formal analytical frameworks that fully develop the potential of this approach. By adapting concepts and procedures used for the analysis of spatial trajectories, we present a framework for describing and comparing community trajectories. A key element of our contribution is the means to assess the geometric resemblance between trajectories, which allows users to describe, quantify, and analyze variation in community dynamics. We illustrate the behavior of our framework using simulated data and two spatiotemporal community data sets differing in the community properties of interest (species composition vs. size distribution of individuals). We conclude by evaluating the advantages and limitations of our community trajectory analysis framework, highlighting its broad domain of application and anticipating potential extensions.

Key words: beta diversity; community dynamics; dissimilarity; distance matrices; forest dynamics; stand size structure; trajectory data mining.

Introduction

Ecologists have long been interested in how community dynamics operate in different kinds of ecosystems. The speed and direction of community changes, as well as the biotic and abiotic factors and processes that drive them, are central to research in community ecology (Pickett et al. 1987, McEwan et al. 2011, Vellend 2016). For example, an overarching question in ecology is how do communities respond to disturbances, depending on disturbance intensity, historical legacies, and environmental factors (Pickett and White 1985, Allison 2004, Reyer et al. 2015). A related question revolves around the existence of a single stable state under given environmental conditions, as opposed to the existence of multiple equilibria driven by priority effects (Chase 2003, Smith 2012). More generally, community ecologists are often interested in discerning the relative contribution of drift, selection, speciation, and dispersal processes to observed community dynamics and the resulting diversity patterns (Vellend 2016). Addressing such questions and goals, or others, requires both long-term experimental/observational data sets, and insightful ways of representing, comparing, and analyzing the dynamics of ecological communities.

Several statistical frameworks have been proposed and used to test hypotheses of community dynamics: (1) questions about temporal divergence/convergence of a set of communities can be addressed by testing for between-survey differences in multivariate dispersion (Anderson 2006) among other approaches (Fukami et al. 2005, Walters and Coen 2006); (2) temporal shifts in average composition in a set of monitored communities can be tested using general procedures such as distance-based or permutational MANOVA (Legendre and

Anderson 1999, Anderson 2001, 2017), or specific tests based on null models (Schaefer et al. 2005); (3) space–time interactions and complex spatiotemporal structures can be tested in different ways (Angeler et al. 2009, Legendre et al. 2010, Legendre and Gauthier 2014); (4) multivariate autoregressive models and joint dynamic species distribution models derived from latent variable modeling can be used to explicitly model species interactions and community dynamics (Hampton et al. 2013, Thorson et al. 2016, Ovaskainen et al. 2017).

Complementing statistical frameworks for hypothesis testing, the dynamics of ecological communities have traditionally been *represented* using scatter plots (e.g., ordination diagrams), where the plotted points correspond to the community states observed at sites at different sampling times, and temporal *community trajectories* are indicated using lines, arrows, or numbers (Austin 1977, Fukami et al. 2005, Trexler et al. 2005, Magalhães et al. 2007, Matthews et al. 2013). In accordance with their visualization, the formal (i.e., quantitative) study of community dynamics can be based on the *geometric analysis* of trajectories. For example, the site-to-site geometric comparison of trajectories can be used to group sites that exhibit the same dynamic patterns and to assess the amount of variation in community dynamics among a set of sites. Although they did not present it this way, the "second-stage community analysis" of Clarke et al. (2006) is an example of this geometry-based perspective. Clarke et al. (2006) compared pairs of trajectories by calculating the Mantel correlation between the dissimilarity matrices representing them. The pairwise comparison of all trajectories provides a new space that summarizes their similarity and can be used to explore the factors underlying differences in community dynamics (Clarke et al. 2006). Other approaches can also be included in this geometric perspective because they analyze dissimilarity values between states of a community trajectory (in other words, they analyze the shape of the trajectory) with the aim to distinguish patterns of community dynamics. Examples include "time-lag analysis," a method to detect patterns of temporal variation by comparing community dissimilarity and the time difference between surveys (Collins et al. 2000, Thibault et al. 2004, Takeuchi et al. 2010), as well as methods that analyze the time series of dissimilarity values with respect to a baseline condition (Philippi et al. 1998, Bagchi et al. 2017). Although these approaches are useful, to our knowledge, the full potential of geometrically based methods remains to be explored, especially by expanding the ways community trajectories are described and compared. Such advances would enable the development of comprehensive frameworks to describe, summarize, and interpret the spatial variation of community dynamics.

The analysis of community trajectories has an analogue in the movement of objects through space. In physics and geography, a *spatial trajectory* is a set of positional information for a moving object, ordered in time (e.g., trajectories of animals, humans, or hurricanes). Ecologists have been traditionally interested in the movement of animals and plant propagules in space (Turchin 1998); and a wealth of statistical methods and null models are available for this purpose (Hooten et al. 2017). Moreover, during the last decade, engineers have also further developed procedures for the analysis of movement in geographic space (Lee et al. 2007, Zheng 2015, Besse et al. 2016), and the discipline that deals with large volumes of spatial trajectory data has been called "trajectory data mining" (Buchin et al. 2013, Zheng 2015). Adapting tools designed for the study of movement to the study of community dynamics can offer novel ways of visualizing, analyzing, comparing, and ultimately understanding them. Indeed, movement metrics developed in animal ecology have been already used to categorize community dynamics into fundamental trajectories related to ecological resilience (Bagchi et al. 2017). This transfer of methodological approaches, however, needs to be done with caution, because of the difference between the geographic space in which spatial trajectories occur and the multidimensional, often non-Euclidean, spaces typical of community ecology.

With the aim of expanding the existing statistical toolbox for spatiotemporal community analysis, here, we present a *community trajectory analysis* (CTA) framework based on the geometric analysis and comparison of community trajectories. Our CTA framework includes (1) definition of a community trajectory in a multivariate space as formal representation of the dynamics in a community, (2) geometric properties of a community trajectory, (3) projection of a community state onto a community trajectory, (4) convergence/divergence between a pair of community trajectories, (5) geometric resemblance between a pair of community trajectories, and (6) spatial variation in community dynamics.

We study the behavior of our CTA framework using simulated community dynamics, which allows us to show how the results of the analysis may be influenced by sampling decisions. We then illustrate CTA in practice using two example applications on real spatiotemporal community data sets differing in what component of dynamics is of interest: compositional dynamics in a species-rich tropical forest and a disturbance-related structural dynamics in a species-poor forest. Finally, we discuss the advantages and limitations of CTA in relation to previous work, its domain of application, and ways in which it can be improved or extended.

COMMUNITY TRAJECTORY ANALYSIS

Definition of a community trajectory

Let o_1, o_2, …, o_n be an ordered set of n community observations ($n > 1$) and t_1, t_2, …, t_n be a corresponding set of ordered times (i.e., $t_1 < t_2 < \cdots < t_n$). For all i in {1, 2, …, n}, let x_i contain the coordinates, or *community state*, corresponding to o_i in a chosen multidimensional space Ω

(e.g., of species composition). We define *community trajectory* as the sequence $\mathbf{T} = \{(x_1, t_1), \ldots, (x_n, t_n)\}$, where n is the *size* of the trajectory. Alternatively, the geometry of $\mathbf{T}$ can also be formalized using a set of $n - 1$ *directed segments* $\{s_1, \ldots, s_{n-1}\}$, where $s_i = \{x_i, x_{i+1}\}$ is a segment with endpoints (community states) x_i and x_{i+1} (Fig. 1). A trajectory will often correspond to repeated observations of the same sampling unit (e.g., surveys of a permanent plot), but not necessarily. For example, a trajectory may also represent a forest chronosequence, where $o_1, o_2, \ldots, o_n$ are observations made at different sites and $t_1, t_2, \ldots, t_n$ correspond to forest ages (Foster and Tilman 2000). A trajectory may also represent an average dynamic pathway (e.g., each x_i corresponding to the average composition of a set of observations).

On the multivariate space Ω.—Unlike spatial trajectories that represent the movement of objects in two-dimensional or three-dimensional Euclidean space, community trajectories represent dynamics in spaces that usually have many dimensions and may not be Euclidean. In the past, the geometric study of community trajectories has been sometimes conducted on Ω being defined by two or three axes of an ordination (Matthews et al. 2013), but this has the drawback of discarding the information of trajectories that is contained in additional dimensions.
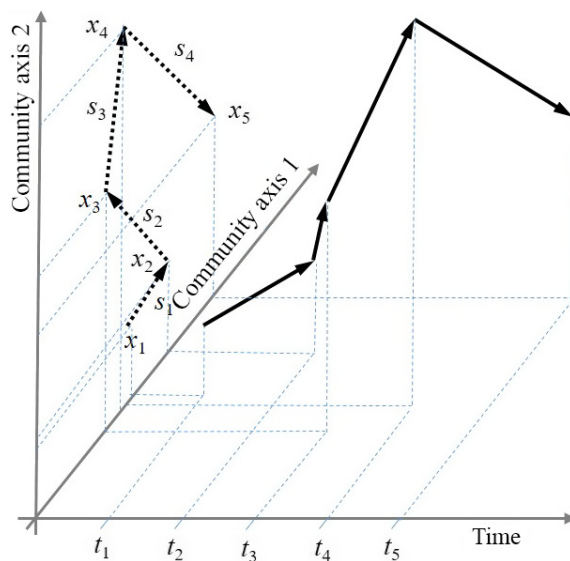

Fig. 1. Example of a community trajectory $\mathbf{T}$ issued from making five community observations ($o_1, \ldots, o_5$) at five ordered points in time ($t_1, \ldots, t_5$). Community observations are represented using a corresponding set of community states ($x_1, \ldots, x_5$) in a multidimensional space (here two principal axes are shown only). The trajectory can also be formalized in terms of four directed segments ($s_1, \ldots, s_4$) in the same space. The x_i are positions (coordinates) and the s_i are vectors that have lengths (i.e., displacement), so that the sum of segment lengths is the total length of the trajectory pathway. The same trajectory can be represented in a three-dimensional plot including the time axis (continuous arrows).

In our framework, we assume that Ω is defined by the resemblance between pairs of community observations, measured using a dissimilarity coefficient d. All analyses that follow are based on the dissimilarity values contained in a distance matrix $\mathbf{\Delta} = [d]$. This approach is not limited in the number of dimensions taken into account.

The study of community dynamics can focus on temporal changes in a wide range of properties (species composition, number and size of individuals, trait distribution, etc.), and the choice of d (hence of Ω) will depend on the property of interest. For example, here we use the percentage difference (Bray-Curtis) coefficient (Bray and Curtis 1957) to define the space Ω in which we study compositional dynamics. Some operations of our framework require d to be a *metric* (i.e., a distance satisfying the triangle inequality). Not all the dissimilarity coefficients used in ecology are metric, but transformations exist that help achieve this property (Gower and Legendre 1986, Legendre and Legendre 2012). Commonly used transformations like taking the square root, however, may strongly modify the angles between consecutive segments, and the overall trajectory shape. In Principal Coordinates Analysis (PCoA; Gower 1966), the usual practice of corrections for negative eigenvalues (Lingoes 1971, Cailliez 1983) can also result in strong shape distortions. In these cases, metric multidimensional scaling (mMDS; Borg and Groenen 2005) will be a better alternative. Note, however, that any Euclidean representation of semi-metric dissimilarities will introduce larger distortions than are needed to ensure metricity only. Moreover, ordination methods modify dissimilarities in different ways depending on the subset of trajectories included in matrix $\mathbf{\Delta}$. While ordination is the correct tool to display trajectories, for the calculation of CTA statistics, we advocate local transformation of semi-metric dissimilarities in every triplet when the triangle inequality is required. This creates the inconsistency that the same distance value may be transformed differently depending on the triangle that is evaluated, but in our opinion has the advantage of introducing a much smaller distortion of overall trajectory shape compared to other approaches.

Geometric properties of a trajectory

These are basic properties that describe the shape of a given trajectory, and their values may be compared between trajectories if they have been defined on comparable Ω spaces (i.e., using the same dissimilarity coefficient and similar spatial and temporal resolution of sampling; see *Simulation Study*).

Length and speed.—Comparing the lengths of community segments and the speeds of trajectories may be interesting, for instance, in studies focusing on the resistance and resilience of communities (Halpern 1989, Allison 2004, Rydgren et al. 2004), or to distinguish between gradual and abrupt changes (Matthews et al. 2013). The

length of a segment s_i is given by the distance between its two endpoints in space Ω, that is $L(s_i) = d(x_i, x_{i+1})$. The *total path length* of a trajectory $L(\mathbf{T})$ is simply the sum of the lengths of all segments that compose it:

$$L(\mathbf{T}) = \sum_{i=1}^{n-1} L(s_i) = \sum_{i=1}^{n-1} d(x_i, x_{i+1}). \qquad (1)$$

Knowing the lengths of segments $s_1, \ldots, s_{n-1}$ and time points $t_1, \ldots, t_n$, one can calculate the *speed* of community change along segments, $S(s_i)$, or the average speed along the trajectory, $S(\mathbf{T})$:

$$S(s_i) = \frac{L(s_i)}{t_{i+1} - t_i} \qquad (2a)$$

$$S(\mathbf{T}) = \frac{L(\mathbf{T})}{t_n - t_1}. \qquad (2b)$$

Both lengths and speeds along trajectories have units that depend on the choice of d. If the Euclidean distance is calculated on individual counts, then lengths can be interpreted as changes in the number of individuals, but other distances may not have interpretable units.

Direction.—Another important geometric feature of a trajectory $\mathbf{T}$ concerns its changes in direction. Unlike in the analysis of movement in space, where angles are often measured on the *X–Y* plane (Turchin 1998), community trajectories are defined in spaces of more than two dimensions and this complicates the study of direction. Let $\{x_i, x_j, x_k\}$ be a triplet of community states of $\mathbf{T}$ that are ordered in time, that is where $t_i < t_j < t_k$. If the set of distances $\{d(x_i, x_j), d(x_i, x_k), d(x_j, x_k)\}$ fulfills the triangle inequality, then angles can be measured on the (Euclidean) plane that contains the three states. We define the angle $0 \leq \theta(x_i, x_j, x_k) \leq 180°$ as the change in direction of vector $\overline{x_j x_k}$ with respect to vector $\overline{x_i x_j}$ in this plane. Values of $0°$ indicate that the three states are completely aligned, and there is no change in direction; whereas values of $180°$ indicate that the orientation of the two vectors is the same but they are opposite in sense (i.e., origins and endpoints are at opposite ends of the vectors). Unlike in spatial movement, it is not useful to distinguish between positive and negative angles (e.g., between clockwise and counterclockwise turns) here, because each triplet forms its own plane. Of particular interest are the angles between consecutive segments s_i and s_{i+1}, that is $\theta(s_i, s_{i+1}) = \theta(x_i, x_{i+1}, x_{i+2})$, because knowing the $n - 1$ sequence of $\theta(s_i, s_{i+1})$ values of a trajectory allows abrupt changes in trajectory direction to be detected (Hughes 1990, Matthews et al. 2013). For example, a large angle (i.e., a sudden change in direction) will often occur between the segment including a disturbance event and the preceding one, due to elevated mortality rates and the recruitment of new individuals/species.

In addition to inspecting the angle between consecutive segments, it is interesting to evaluate the *overall*

directionality of the trajectory (i.e., the degree to which the community is consistently following a particular direction in Ω), which may be useful, for example, to distinguish between communities subjected to stabilizing (non-directional) selection from those influenced by directional or disruptive selection (Matthews et al. [2013]; selection sensu Vellend [2016]). We advocate directionality statistics bounded between 0 and 1, where the maximum value corresponds to a straight trajectory. For two-dimensional spatial trajectories and angles measured from the same reference direction (e.g., north), the *sample mean vector length* of circular statistics (Pewsey et al. 2013) is a good candidate to measure overall directionality because it is bounded between 0 and 1 and its highest value is attained when all trajectory segments are in the same direction. Unfortunately, angles are defined differently in the case of community trajectories, so a different statistic must be found. In a straight trajectory $\theta(x_i, x_j, x_k) = 0°$ for all r time-ordered triplets; and for non-straight trajectories, the maximum penalty should occur for $\theta = 180°$. Moreover, angles of triplets corresponding to distant surveys should have a larger influence in the determination of overall directionality than angles between consecutive surveys. Therefore, we suggest overall directionality of a trajectory $\mathbf{T}$ is measured using the following:

$$\text{DIR}(\mathbf{T}) = \frac{\sum w_{ijk} \times ((180 - \theta_{ijk})/180)}{\sum w_{ijk}} \qquad (3)$$

where $\theta_{ijk} = \theta(x_i, x_j, x_k)$, $w_{ikj} = d(x_i, x_j) + d(x_j, x_k)$ and summation are over all r time-ordered triplets. $0 \leq \text{DIR}(\mathbf{T}) \leq 1$ and one should expect $\text{DIR}(\mathbf{T})$ to be close to 1 if compositional dynamics are directional as in primary or secondary succession. Philippi et al. (1998) presented several ways to measure "progressive change." One of them is to conduct a Mantel test of the correlation between state dissimilarity and difference in survey times, using all pairs of states in the trajectory (i.e., a type of multivariate seriation test). This approach, however, is sensitive not only to changes in trajectory direction but also to changes in trajectory speed.

Projection of a community state onto a trajectory

Let y be the coordinates (state) in space Ω of o_A, the community that has been sampled at a given site A. Let now $\mathbf{T} = \{(x_1, t_1), \ldots, (x_n, t_n)\}$ be a trajectory in Ω representing a known dynamic pathway such as a successional sequence. Assessing the distance between y and its projection onto $\mathbf{T}$ is ecologically relevant because, if this distance is small enough, the future dynamics of o_A could follow $\mathbf{T}$. Additionally, knowing the relative position of y along $\mathbf{T}$ would allow the assessment of how far a given disturbed community is along a given pathway

known to occur under a particular model of community dynamics. We begin by considering the comparison of y with a single directed segment $s_i = \{x_i, x_{i+1}\}$. D_{ps}, the distance between a community state y and a segment s_i, is defined as (Besse et al. 2016; Fig. 2a, b)

$$D_{\mathrm{ps}}(y, s_i) = \begin{cases} d(y, \mathrm{proj}(y, s_i)) & \text{if } \mathrm{proj}(y, s_i) \in s_i \\ \min(d(y, x_i), d(y, x_{i+1})) & \text{otherwise} \end{cases}$$

$$(4)$$

where ps stands for "point-to-segment" and $\mathrm{proj}(y, s_i)$ is the orthogonal projection of y onto s_i. If the orthogonal projection of y onto s_i falls outside the segment (Fig. 2b), the second line in Eq. 4 defines the distance between y and s_i as the distance between y and the closest of the two endpoints. To calculate $d(y, \mathrm{proj}(y, s_i))$, the triplet $\{d(y, x_i), d(y, x_{i+1}), d(x_i, x_{i+1})\}$ again needs to fulfill the triangle inequality. The same triplet allows $R(y, s_i)$, the *relative position* of $\mathrm{proj}(y, s_i)$ within s_i, to be calculated ($R \in [0,1]$):

$$R(y, s_i) = \begin{cases} \frac{d(x_i, \mathrm{proj}(y, s_i))}{L(s_i)} & \text{if } \mathrm{proj}(y, s_i) \in s_i \\ 0 & \text{if } D_{\mathrm{ps}}(y, s_i) = d(y, x_i) \\ 1 & \text{if } D_{\mathrm{ps}}(y, s_i) = d(y, x_{i+1}) \end{cases} \cdot$$

$$(5)$$

The distance from a community state y to a trajectory $\mathbf{T}$, D_{pt}, is naturally defined as the minimum distance between y and the various segments of $\mathbf{T}$ (Besse et al. 2016; Fig. 2c)

$$D_{\mathrm{pt}}(y, \mathbf{T}) = \min\{D_{\mathrm{ps}}(y, s_1), \ldots, D_{\mathrm{ps}}(y, s_{n-1})\} \quad (6)$$

where pt stands for "point-to-trajectory." And the *relative position* of $\mathrm{proj}(y, s_i)$ within T is

$$R(y, \mathbf{T}) = \frac{L(s_i) \times R(y, s_i) + \sum_{k=1}^{k=i-1} L(s_k)}{L(\mathbf{T})} \quad (7)$$

where $R(y, \mathbf{T})$ is a normalized measure of the position of y when projected onto $\mathbf{T}$.

Convergence/divergence between a pair of trajectories

The temporal divergence/convergence of a set of communities synchronously sampled can be addressed by testing for between-survey differences in multivariate dispersion (Anderson 2006). We are interested here in testing the convergence or divergence between a particular *pair* of trajectories $\mathbf{T}_1 = \{(x_{11}, t_{11}), (x_{12}, t_{12}), \ldots, (x_{1n}, t_{1n})\}$ and $\mathbf{T}_2 = \{(x_{21}, t_{21}), (x_{22}, t_{22}), \ldots, (x_{2m}, t_{2m})\}$ of lengths n and m, respectively (whenever two subindices are used for x or t the first one indicates the trajectory and the second indicates the community state within it). This test can be addressed in two ways.

1) If the two sites have been surveyed synchronously, that is if $n = m$ and $t_{11} = t_{21}, t_{12} = t_{22}, \ldots, t_{1n} = t_{2n}$, one can analyze the sequence of n distances $\{d(x_{11}, x_{21}), d(x_{21}, x_{22}), \ldots, d(x_{1n}, x_{2n})\}$ (Fig. 3a).

a　$D_{ps}(y, s_i) = d(y, \mathrm{proj}(y, s_i))$
　$R(y, s_i) = d(x_i, \mathrm{proj}(y, s_i))/L(s_i)$

b　$D_{ps}(y, s_i) = d(y, x_{i+1})$
　$R(y, s_i) = 1$

c　$D_{pt}(y, \mathbf{T}) = D_{ps}(y, s_3)$
　$R(y, \mathbf{T}) = \dfrac{L(s_3) \times R(y, s_3) + L(s_1) + L(s_2)}{L(\mathbf{T})}$
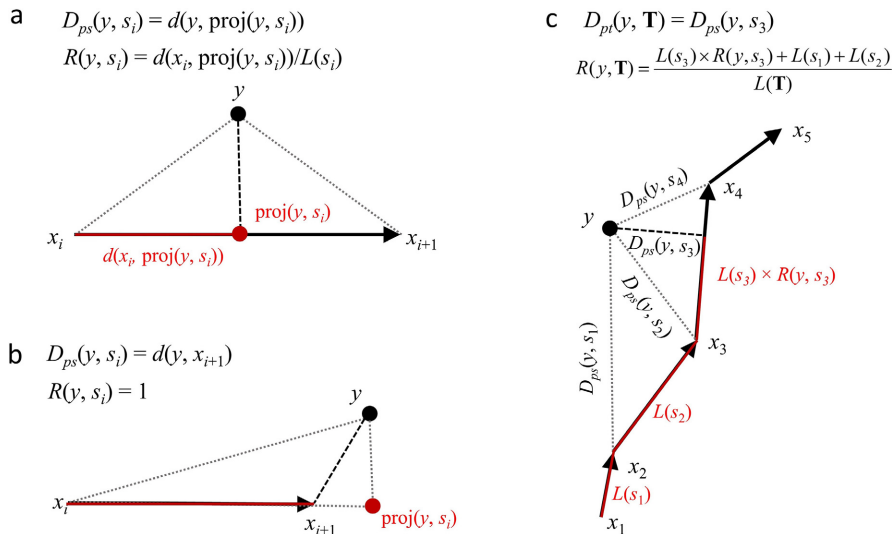
FIG. 2.　(a, b) Illustration of the calculation of the distance between a community state y and a segment s_i (dashed line; Eq. 4) and the relative position of y on s_i (Eq. 5) when the projection of y onto s_i (a) belongs to the segment and (b) when it does not. (c) Illustration of the calculation of the distance between a community state y and a trajectory $\mathbf{T} = \{s_1, s_2, s_3, s_4\}$ (Eq. 6) and the relative position of y on $\mathbf{T}$ (Eq. 7), with dotted lines representing distances between y and particular segments and the dashed line representing the minimum among them. Red lines indicate the absolute position, which leads to the relative position when compared to the total length of the segment/trajectory.
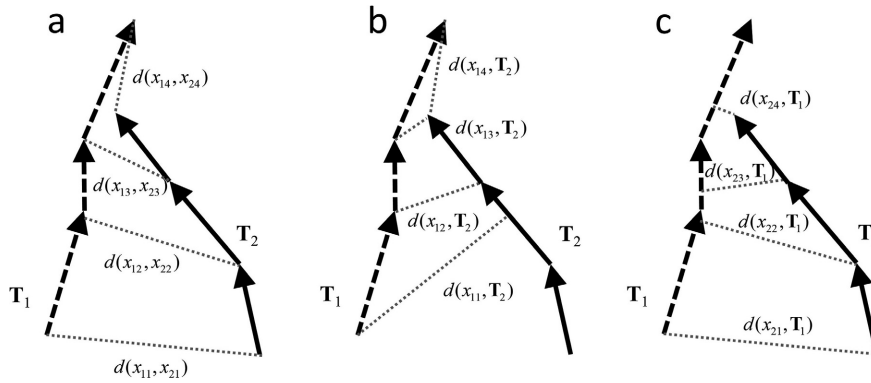
FIG. 3. Illustration of different ways to study the convergence/divergence between two trajectories T_1 (dashed segments) and T_2 (continuous segments) of sizes n and m, respectively: (a) symmetric convergence analysis by analyzing the sequence of $n = m$ distances between synchronous surveys; (b) asymmetric analysis of the n distances from states of T_1 to T_2; (c) asymmetric analysis of the m distances from states of T_2 to T_1.

2) More generally, one can analyze the sequence of n distances $\{D_{pt}(x_{11}, T_2), D_{pt}(x_{12}, T_2), \ldots, D_{pt}(x_{1n}, T_2)\}$ or the sequence of m distances $\{D_{pt}(x_{21}, T_1), D_{pt}(x_{22}, T_1), \ldots, D_{pt}(x_{2m}, T_1)\}$ (Fig. 3b, c).

The comparison of the two trajectories in approach 1 is symmetric, that is the two trajectories converge/diverge or not simultaneously. However, in approach 2, the analysis is asymmetric, so that T_1 may approach T_2 while the reverse may not be true. It may even happen that T_1 approaches T_2 while T_2 separates from T_1, for example if they are one behind the other or at right angles to one another. Approach 1 addresses the question of *whether the states of the two trajectories become closer/farther over time*, while approach 2 addresses the question of *whether the states of a given observed trajectory become closer/farther to a reference trajectory over time*. In both cases, a Mann-Kendall test (Mann 1945) can be used to test for a monotonically increasing (divergence or separation) or decreasing (convergence or approach) trend.

Geometric resemblance between a pair of trajectories

Assessing the resemblance between spatial trajectories can be done in many ways (Vlachos et al. 2002, Lee et al. 2007, Besse et al. 2016, Tripathi et al. 2016). Analogously, there are multiple ways of assessing the resemblance between community trajectories. For example, in second-stage community analysis (Clarke et al. 2006), the resemblance between a pair of trajectories is measured by calculating the nonparametric Spearman correlation between their corresponding dissimilarity matrices. In the following two subsections, we develop a geometrically based approach to trajectory resemblance that takes into account the *shape, size, direction*, and *position* of trajectories but is not sensitive to differences in *speed*. While accounting for speed would be preferable, a purely geometric approach is easier to define and may produce similar results if the time intervals between consecutive surveys are more or less constant. The last subsection deals with the issue of excluding differences in position while keeping the other components of trajectory resemblance.

Dissimilarity between two segments.—Let us first consider D_{HS}, the Hausdorff distance between two segments $s_i = \{x_i, x_{i+1}\}$ and $s_j = \{x_j, x_{j+1}\}$ (see general definition of the Hausdorff distance in Appendix S1; Besse et al. 2016; see Fig. 4a)

$$D_{HS}(s_i, s_j) = \max\{D_{ps}(x_i, s_j), D_{ps}(x_{i+1}, s_j), \\ D_{ps}(x_j, s_i), D_{ps}(x_{j+1}, s_i)\}. \tag{8}$$

An important limitation of D_{HS} for the purpose of comparing community dynamics is that it does not account for the fact that directed segments have a *direction* (from the initial endpoint to the final endpoint). Because the direction sense is important in ecology (e.g., population growth should be distinguished from mortality), D_{HS} will underestimate the difference between the two directed segments if they have opposite directions. To account for the direction of segments, Lee et al. (2007) proposed to explicitly incorporate the angle between segments in the definition of the distance. Alternatively, we define that a segment s_i has a direction *opposite* to s_j if the projection of the final endpoint of s_i (i.e., x_{i+1}) on s_j has a relative position (Eq. 5) smaller than that of the projection of the origin of s_i (i.e., x_i) on s_j, that is if $R(x_{i+1}, s_j) < R(x_i, s_j)$ (Fig. 4b). Visually, s_i has a direction *opposite* to s_j if the angle between them is larger than 90°, but this angle is not directly measured because the segments s_i and s_j have different origins. As such, we define the *directed segment* dissimilarity, D_{DS}, using a modification of Eq. 8

$$D_{DS}(s_i, s_j) = \max\{D_{ps}(x_i, s_j), D'_{ps}(x_{i+1}, s_j), \\ D_{ps}(x_j, s_i), D'_{ps}(x_{j+1}, s_i)\} \tag{9}$$
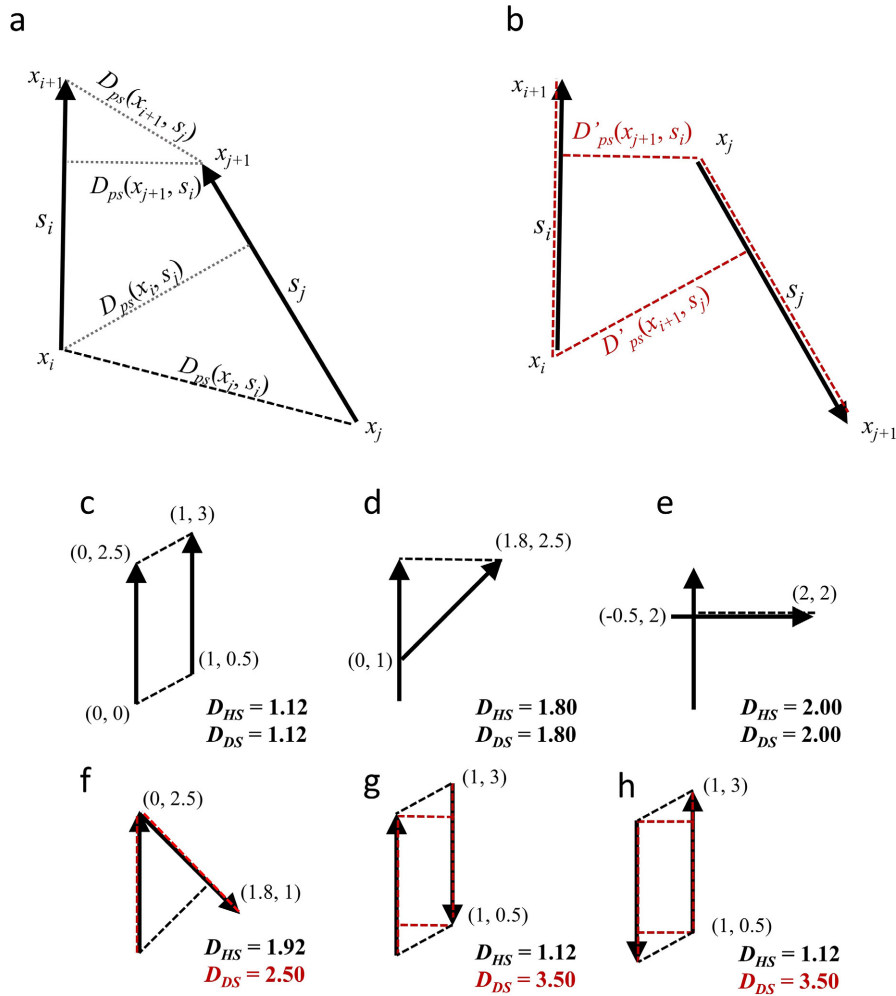
Fig. 4. (a, b) The calculation of resemblance between a pair of directed segments s_i and s_j. In panel a, D_{HS} (Eq. 8) and D_{DS} (Eq. 9) are equal and the distance value results from finding the maximum (black dashed line) among the four endpoint-to-segment distances (dotted lines). When the direction of segment s_j is reversed (b), to calculate D_{DS} (Eq. 9), the distance between the final endpoint of one segment and the other segment is modified (Eq. 10). Here, the red dashed paths indicate the modified distances $D'_{ps}(x_{i+1}, s_j)$ and $D'_{ps}(x_{j+1}, s_i)$. (c–h) Examples of D_{HS} and D_{DS} values for different pairs of directed segments (assuming Cartesian coordinates and the Euclidean distance for d). In panels c–e, dashed lines indicate the (equal) value of D_{HS} and D_{DS}; in panels f–h, black dashed lines have lengths equal to D_{HS} values, and red dashed paths have lengths equal to D_{DS} values. Note the progressive increase in D_{DS} when the second segment is tilted until its direction has opposite sense with respect to the first.

where:

$$D'_{ps}(x_{i+1}, s_j) = \begin{cases} D_{ps}(x_{i+1}, s_j) & \text{if } R(x_i, s_j) \le R(x_{i+1}, s_j) \\ D_{ps}(x_i, s_j) + L(s_i) & \text{otherwise} \end{cases}.$$

(10)

The distance between x_{i+1} and s_j is modified if $R(x_{i+1}, s_j) < R(x_i, s_j)$. The idea underlying the modification is that if s_i has a direction opposite to s_j the distance from the final endpoint of s_i (i.e., from x_{i+1}) to s_j should be larger than the distance from the initial endpoint of s_i (i.e., from x_i). If $R(x_{j+1}, s_i) < R(x_j, s_i)$, then $D'_{ps}(x_{j+1}, s_i)$ is defined analogously. Ecologically, these modifications are intended to make dissimilarity between population growth and mortality processes (or between constant composition and complete species replacement) larger than dissimilarity between different degrees of population growth (or between colonization events by different species while keeping existing ones). The red dashed paths in Fig. 4b illustrate the calculation of $D'_{ps}(x_{i+1}, s_j)$ and $D'_{ps}(x_{j+1}, s_i)$ in these cases, and Fig. 4c–h shows values of D_{HS} and D_{DS} under different situations (note the increase in D_{DS} from 4c to g).

Dissimilarity between two trajectories.—Let $\mathbf{T}_1$ and $\mathbf{T}_2$ be two trajectories to be compared. If they represent the dynamics of two sites that have been surveyed synchronously (i.e., if $n = m$ and $t_{11} = t_{21}, t_{12} = t_{22}, \ldots, t_{1n} = t_{2n}$), a straightforward way of comparing them is to calculate the mean across surveys of community

dissimilarity between the two sites (i.e., the mean of the sequence $\{d(x_{11}, x_{21}), d(x_{21}, x_{22}), \ldots, d(x_{1n}, x_{2n})\}$). However, we are interested here in the comparison of dynamics beyond this specific case. As for segments, the resemblance between trajectories $\mathbf{T}_1$ and $\mathbf{T}_2$ can be assessed using the Hausdorff distance, but this approach and others are inappropriate because segment directions are not taken into account (Appendix S1). Instead, we advocate what we call the *directed segment path dissimilarity* (D_{DSP})

$$D_{DSP}(\mathbf{T}_1, \mathbf{T}_2) = \frac{1}{n-1} \sum_{i=1}^{n-1} D_{DS}(s_{1i}, \mathbf{T}_2). \quad (11)$$

where, analogously to the distance between a community state and a trajectory (Eq. 6), the distance of segment s_{1i} to trajectory $\mathbf{T}_2$ is

$$D_{DS}(s_{1i}, \mathbf{T}_2) = \min\{D_{DS}(s_{1i}, s_{21}), \ldots, D_{DS}(s_{1i}, s_{2m})\}. \quad (12)$$

D_{DSP} (Eq. 11) takes into account the direction of trajectories because it is based on D_{DS} (Eq. 9), which explicitly accounts for the direction and sense of segments. D_{DSP} is not symmetric, which precludes the application of most multivariate analysis methods unless it is symmetrized, for example by averaging $D_{DSP}(\mathbf{T}_1, \mathbf{T}_2)$ and $D_{DSP}(\mathbf{T}_2, \mathbf{T}_1)$:

$$D_{SDSP}(\mathbf{T}_1, \mathbf{T}_2) = \frac{D_{DSP}(\mathbf{T}_1, \mathbf{T}_2) + D_{DSP}(\mathbf{T}_2, \mathbf{T}_1)}{2}. \quad (13)$$

Calculating this average results in the loss of information about asymmetry in trajectory resemblance. For example, if the pathway of $\mathbf{T}_1$ is a subset of the pathway in $\mathbf{T}_2$, then $D_{DSP}(\mathbf{T}_1, \mathbf{T}_2) = 0$ because one can perfectly map $\mathbf{T}_1$ onto $\mathbf{T}_2$, but $D_{DSP}(\mathbf{T}_2, \mathbf{T}_1) > 0$ because $\mathbf{T}_2$ includes segments that are lacking in $\mathbf{T}_1$. This distinction is lost when calculating D_{SDSP}. Although D_{SDSP} will be preferred to D_{DSP} in most cases, asymmetric resemblance between trajectories may be useful, for example, to measure the degree to which the pathway of a given monitored community is similar to the (lengthier) pathway expected under a particular model of community dynamics.

Dissimilarity between centered trajectories.—So far, we assumed that trajectory resemblance should take into account *shape*, *size*, *direction*, and *position*. However, if d measures differences in community composition, a large amount of the variation in Ω may be due to spatial compositional differences that are time invariant (i.e., species with different abundances at different sites but whose abundance is not changing through time) and that may mask temporal changes in composition (i.e., species that are decreasing or increasing). In this context, an approach excluding differences in trajectory *position* may be appropriate to assess resemblance in community

dynamics, for example to focus on the effect of local extinction and colonization. Such an approach can be achieved by *centering* all trajectories prior to the calculation of trajectory distances. Starting from the matrix $\Delta = [d]$ of compositional dissimilarities between states, trajectory centering can be conducted in a way that parallels the derivation of residual plots in PERMANOVA, by specifying a model matrix coding for trajectories and using linear algebra (see details in Anderson [2017]). The centering procedure can be applied to both Euclidean and non-Euclidean d measures, although in the second case, some distance values may be complex numbers after centering, which requires discarding the imaginary part (and hence will imply small distortions). The resulting centered dissimilarity matrix Δ^{cent} can be used to calculate distances between trajectories, as before, using D_{SDSP}. Note that centering of trajectories can be done with respect to the initial composition, rather than trajectory centroids, if the focus of CTA is on community divergence.

Spatial variation in community dynamics

Let $\{\mathbf{T}_1, \mathbf{T}_2, \ldots, \mathbf{T}_r\}$ be a set of r community trajectories and let $n_1, n_2, \ldots, n_r$ ($n_i > 1$) be the size (i.e., number of observations) of each trajectory. Then, let N_P be the total number of observations (i.e., $N_P = \sum_{i=1}^{r} n_i$) and Δ be a $N_P \times N_P$ symmetric matrix calculated by applying d to all pairs of observations. Matrix Δ defines the resemblance between all observations (community states), which may belong to the same or different trajectories. Let now $\mathbf{D}_T$ be an $r \times r$ symmetric matrix of dissimilarity values between pairs of trajectories obtained using D_{SDSP}, or another suitable measure (Clarke et al. 2006). In the following, we describe how variation in community dynamics can be assessed and summarized with well-known multivariate methods, but using $\mathbf{D}_T$ as input, instead of Δ or the usual community data tables (e.g., site-by-species compositional data). Although we require Δ to be a metric, this does not imply that $\mathbf{D}_T$ will have this property (in fact, D_{SDSP} is not a metric) and this fact must be taken into account when studying the variation in $\mathbf{D}_T$.

Overall variation in community dynamics.—Beta diversity is generally defined as the variation in community composition among sites within a geographical area (Whittaker 1960). Several frameworks have been proposed to quantify beta diversity (Anderson et al. 2011, Legendre and De Cáceres 2013), as well as to display and summarize spatial patterns of community variation (Dray et al. 2012). More recently, temporal beta diversity or temporal turnover have been defined as the change in community composition through time (Dornelas et al. 2014, Legendre and Gauthier 2014, Shimadzu et al. 2015). Being able to measure the resemblance between pairs of community trajectories allows variation in community dynamics to be

quantified. We will use the term "dynamic beta diversity" (dBD) to refer to the *overall variation of community dynamics* observed in a given area during a given period of time. While in the previous sections, we allowed trajectories to be compared without restrictions, assessing dBD requires that trajectories refer to a definite set of *r* sampling units and to a common sampling period. In what follows, we briefly repeat part of the framework of Legendre and De Cáceres (2013), adapting it to variation in community dynamics (other beta diversity frameworks could be adapted). First, the overall variation in $\mathbf{D}_T$ is found by summing the squared dissimilarities in the upper or lower half of matrix $\mathbf{D}_T$

$$SS_{Total} = \frac{1}{r} \sum_{h=1}^{r-1} \sum_{i=h+1}^{r} D_{T,hi}^2 \qquad (14)$$

where the notation SS is chosen by analogy with the sum-of-squares obtained from compositional data. SS_{Total} is not a comparable measure of variation because it depends on the number of trajectories, but forms the basis for defining an index of dBD

$$dBD_{Total} = \frac{SS_{Total}}{(r-1)}. \qquad (15)$$

Eq. 14 converts the sum-of-squares into an unbiased estimator of variance, whose values can be compared between different data sets (Legendre and De Cáceres 2013). In addition to quantifying variation in community dynamics, it is possible to split SS_{Total} into the set of *r* contributions of individual community trajectories. This requires the calculation of the Gower-centered matrix $\mathbf{G}$ (Gower 1966)

$$\mathbf{G} = \left(\mathbf{I} - \frac{\mathbf{1}\mathbf{1}'}{r}\right)\mathbf{A}\left(\mathbf{I} - \frac{\mathbf{1}\mathbf{1}'}{r}\right) \qquad (16)$$

where $\mathbf{I}$ is an identity matrix of size *r*, $\mathbf{1}$ is a vector of ones and $\mathbf{A} = [a_{hi}] = [-0.5D_{T,hi}^2]$. The set of diagonal elements of $\mathbf{G}$, that is $\text{diag}(\mathbf{G}) = \{SS_1, SS_2, \ldots, SS_r\}$, are the squared dissimilarities of the trajectories to the multivariate centroid of the principal coordinate space of $\mathbf{D}_T$. From here, the vector of *local contributions to dynamic beta diversity* (LCdBD) can be obtained as

$$[LCdBD_i] = \frac{\text{diag}(\mathbf{G})}{SS_{Total}}. \qquad (17)$$

LCdBD indices measure the degree of uniqueness of the sampling unit *i* in terms of community dynamics, in comparison with the dynamics exhibited by the other sampling units. Eqs. 14–17 can be applied directly to $\mathbf{D}_T$ regardless of Euclidean or metric properties. The interpretation of LCdBD values depends on the choices of *d* and the distance between trajectories. For example, if *d* measures differences in composition and one uses D_{SDSP}

without centering trajectories, high LCdBD values may occur in sites whose initial states are similar to others but undergo unique dynamics, or in sites whose initial states were already unique.

Displaying and summarizing variation in community dynamics.—Unconstrained ordination methods, such as PCoA (Gower 1966), metric MDS, or nonmetric MDS (Borg and Groenen 2005), are commonly used to display trajectories in the space of community resemblance (i.e., matrix $\mathbf{\Delta}$). As shown by Clarke et al. (2006), these can be complemented by ordination analyses of matrix $\mathbf{D}_T$ to display the resemblance in community dynamics. Because the matrices $\mathbf{\Delta}$ and $\mathbf{D}_T$ are related, their corresponding ordination diagrams can be displayed side by side; the former displays the similarity among community states (with trajectories joining points), and the latter displays the overall resemblance among community trajectories. Furthermore, groups of similar trajectories in the space of $\mathbf{\Delta}$ can be identified by applying clustering procedures (Everitt et al. 2011) on the space of $\mathbf{D}_T$ (or to an Euclidean representation approximating its dissimilarity values if required by the clustering algorithm), and the clustering results can be represented by colors, symbols or ellipses in the two ordination diagrams derived from $\mathbf{\Delta}$ and $\mathbf{D}_T$. Representative trajectories may be defined using centroids or medoids in the space of $\mathbf{D}_T$ and then interpreted using the original spatiotemporal community data (see section Examples of Application).

Software availability

To facilitate conducting CTA, functions to calculate length, angles, directionality, projection, and distances between segments/trajectories (Eqs. 1–13) have been included in a new version of the R package vegclust (De Cáceres et al. 2010), available in CRAN and GitHub repositories. Eqs. 14–17 can be calculated using the R package adespatial (Dray et al. 2017), and multivariate analyses can be conducted using several packages. A vignette explaining how to conduct CTA using vegclust have been added to the package.

SIMULATION STUDY

Community trajectories are meant to represent community dynamics, but their geometric properties are affected by the spatial and temporal resolution of community sampling. We conducted a simulation study to illustrate the behavior and usefulness of CTA when applied to different kinds of compositional dynamics, while assessing the variability of results derived from differences in the size of sampling units and the frequency of surveys. To model species–environment relationships, we considered a theoretical environmental space with two dimensions (gradients). Species responses were simulated using Gaussian curves on each dimension independently, and initial composition was taken as that

corresponding to the center of the environmental space. All simulations included random mortality of individuals, whereas species identity of recruits depended on the simulated scenario. Three kinds of community dynamics were simulated: (1) stabilizing selection, (2) post-disturbance recovery, and (3) directional selection. By "stabilizing selection," we mean that species recruitment probabilities were held constant and equal to the initial community composition, so that variability in transient compositional states (i.e., ecological drift) could be observed as a result of stochastic death and recruitment, but no long-term community change occurred. Under this scenario, CTA results depended on the interaction between ecological drift and sampling decisions, hence constituting a benchmark against which to compare the outcome of directional dynamics. The "disturbance recovery" scenario was similar to the previous one, but where 80% of the individuals had been randomly removed from the initial community. Under this scenario, the initial (post-disturbance) composition of small sampling units could differ due to the stochastic mortality caused by disturbance, but recruitment probabilities were the same as before, so that composition was directed toward the pre-disturbance dynamic equilibrium. Finally, "directional selection" included a progressive change toward a different dynamic equilibrium. This was obtained by setting recruitment probabilities to those corresponding to a position in the environmental space different from the one used to define the initial composition. Each simulated data set focused on one scenario of community dynamics and included the community dynamics of 16 different simulations (i.e., 16 trajectories). For stabilizing selection and disturbance recovery scenarios, all 16 simulations were conducted with the same parameters. For the directional selection, scenario-simulated community dynamics were paired. Here, each of eight pairs of initial communities was assigned a different target position of the environmental space (i.e., a different set of recruitment probabilities), located five units away from the center of the environmental space in the direction of one of the eight cardinal/intercardinal directions. Therefore, community dynamics within the same pair were independent stochastic realizations of the same directional pattern, but different pairs exhibited dynamics in different directions.

Compositional differences were measured using the percentage difference coefficient, and we calculated for each trajectory the total path length (Eq. 1), average angle between consecutive segments, and overall directionality (DIR; Eq. 3). We also calculated the dissimilarity (D_{SDSP}; Eq. 13) between the 16 trajectories in each data set and the resulting dynamic beta diversity (dBD$_{Total}$; Eq. 15). Additionally, the design of the "directional selection" scenario allowed us to evaluate (1) the power of the proposed convergence/divergence test for pairs of parallel trajectories and for diverging trajectories, (2) whether the pairs of communities

departing in different directions of the environmental space could be recovered by applying k-means clustering on the space of trajectory resemblance. The details of the simulation study are given in Appendix S3, and computer code to replicate simulations is provided in Data S1. We summarize the main results here:

1) Regardless of the type of dynamics simulated, the total path length of compositional trajectories generally decreases when the size of sampling units increases (due to a reduced effect of drift) or when the frequency of surveys increases (Fig. 5a).
2) For stabilizing selection, the average angle between consecutive segments generally increases, and DIR decreases, when the size of sampling units increases or when the frequency of surveys decreases (Fig. 5b, c). DIR is higher for disturbance recovery and directional selection than for stabilizing selection, the difference increasing with the size of sampling units.
3) The proposed convergence/divergence test allows trajectories departing in different directions to be distinguished from parallel trajectories. Factors affecting the power of this test are trajectory size (n) and the size of sampling units.
4) The proposed dissimilarity between trajectories (D_{SDSP}; Eq. 13) coupled with cluster analysis allows trajectories moving in similar directions to be identified and grouped together. However, in small sampling units, the effect of ecological drift lowered the degree of agreement between clustering results and simulated dynamics.
5) Regardless of the type of dynamics simulated, dBD$_{Total}$ tends to decrease when the size of sampling units or the frequency of surveys increases (Fig. 5d).

EXAMPLES OF APPLICATION

Compositional dynamics in the Barro Colorado Island permanent forest plot

The 50-ha (1,000 × 500 m) permanent forest plot of Barro Colorado Island (BCI; Panama; 9°9′ N, 79°51′ W) is mostly an old-growth, evergreen, tropical forest. Rainfall reaches 2,500 mm/yr but includes a four-month dry season. The plot is on a rather flat terrain, but six habitats have been distinguished by Harms et al. (2001; Fig. 6a; Table 1). Whereas a small portion of BCI consists of young forests (F), the remaining habitats correspond to old-growth forest occurring under different hydrologic regimes. The high plateau (H) is the driest area of BCI, and the andesitic cap beneath it accumulates water and creates springs along the slopes (S), which therefore have shorter duration of drought. At the other moisture extreme, the seasonal swamp (W) is inundated during the wet season, which does not occur in the low plateau (L). Finally, streamside (R) habitats occurring adjacent to seasonal streams tend to contain water into the dry season. During 1983 "El Niño," the dry
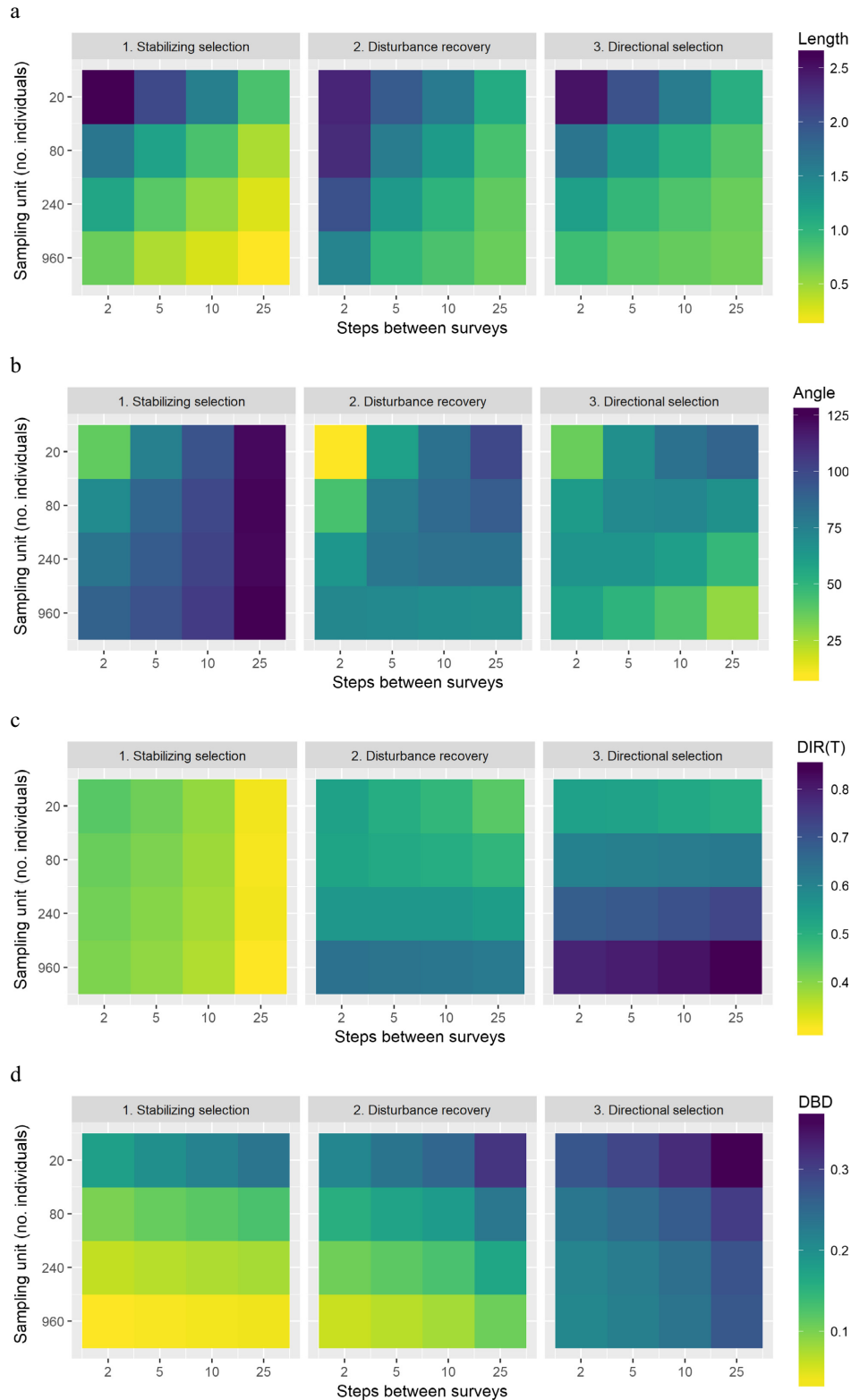
Fig. 5. Results of the simulation study for (a) mean trajectory length, (b) mean angle between consecutive segments, (c) mean trajectory directionality, and (d) dynamic beta diversity of the data set. Each panel shows the mean value of these statistics across 20 simulated data sets for combinations of the size of sampling units (in our case, the maximum number of individuals in the simulated community) and the number of steps between surveys. Columns indicate the kind of simulated dynamics.
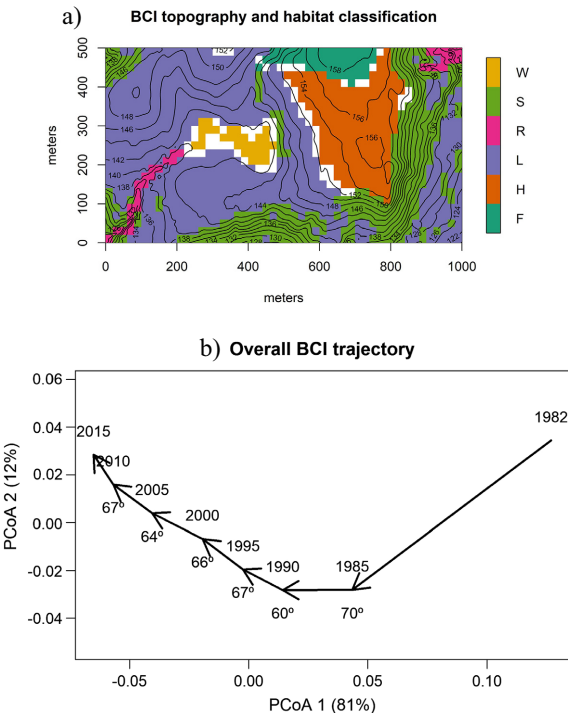
a) **BCI topography and habitat classification**

b) **Overall BCI trajectory**

Fig. 6. (a) Barro Colorado Island (BCI; Panama) forest topography and habitat classification according to Harms et al. (2001). Habitat types are W, swamp; S, slopes; R, streamside; L, low plateau; H, high plateau; F, young forest. (b) Ordination of the community states of the overall BCI compositional trajectory (PCoA). Survey years are indicated, as well as turning angles between consecutive segments.

season was unusually severe, lasting five months and leading to elevated mortality rates, particularly of drought-sensitive canopy trees (Condit et al. 1996). Indeed, mortality rates have been estimated to be around 3% per year for the 1982–1985 period but only 2% per year afterward (Condit et al. 1995), corresponding to 10% in a 5-yr period. Despite being influenced by spatially, and temporally, variable selection processes, community dynamics in BCI are strongly driven by ecological drift and immigration (Condit et al. 2012a). Eight plot surveys (1982, 1985, 1990, 1995, 2000, 2005,

2010, and 2015) have been conducted in BCI, spanning 33 years of compositional dynamics.

We applied our CTA framework on BCI tree data to (1) describe the overall tree compositional dynamics of the forest, (2) determine whether the different BCI habitats exhibit similar direction in compositional dynamics, (3) test for compositional convergence/divergence among habitats, and (4) test for differences in trajectory speed and directionality among habitats.

To study compositional dynamics at the whole forest level, we first calculated the basal area (m^2/ha) of all living stems larger than or equal to 1 cm in diameter by species and survey, obtaining a data table with 8 rows (surveys) and 328 columns (tree species). Then, we built a compositional dissimilarity matrix Δ_{BCI} using the percentage difference (Bray-Curtis) coefficient. Fig. 6b displays the trajectory of the BCI forest in this compositional space. The two longest segments correspond to changes that occurred between the first and second surveys (1982–1985) and between the second and the third (1985–1990), reflecting the above-normal mortality rates associated with the 1983 drought and subsequent recruitment. Angles between consecutive segments were all between 60° and 70°. These values can be compared with our simulations for 960 individuals and surveys every two steps (i.e., largest sample unit size and 10% individual turnover between surveys), which yielded means of 89° and 59° under stabilizing selection and directional selection, respectively. Regarding overall directionality, DIR(T) = 0.63 for BCI whereas our corresponding simulations yielded means of DIR(T) = 0.41 and DIR(T) = 0.79 under stabilizing selection and directional selection, respectively. While caution should always be taken when comparing simulations with real data sets, our comparisons are consistent with directional selection operating on the long-term compositional dynamics of the BCI forest, in addition to ecological drift. We interpret this temporal pattern as the result of species replacement derived from mortality/recruitment during the 1983 drought and the subsequent growth of the newly established individuals.

To describe and compare compositional dynamics among BCI habitats, we first calculated the basal area

TABLE 1. Number of 20 × 20 quadrats, number of trees, mean tree density per quadrat, quadrat mean trajectory length ($L_{20\times20}$) and quadrat mean directionality (DIR$_{20\times20}$) by Barro Colorado Island (BCI; Panama) forest habitats and for the whole forest.

Habitat	No. quadrats	No. trees	Density (trees/quadrat)	L_{20x20}	DIR$_{20x20}$
Young forest (F)	48	11,349	236	0.796[a]	0.635[a]
High plateau (H)	171	36,187	212	0.821[a]	0.619[a]
Low plateau (L)	619	134,360	217	0.925[b]	0.606[a]
Stream sides (R)	32	6,455	202	0.979[b]	0.602[a]
Slopes (S)	284	59,002	208	0.958[b]	0.597[a]
Swamp (W)	30	3,575	119	1.029[b]	0.557[b]
BCI forest	1,250	263,860	211	0.914	0.605

Note: Letters indicate homogeneous groups of habitats according to Tukey's honest significant differences.

by species for 20×20 m quadrats, obtaining a data table with 1,250 rows (quadrats) and 328 columns (tree species). We then averaged basal area values within each habitat, obtaining a table of 48 rows (eight surveys times six habitats). The percentage difference coefficient calculated on pairs of rows of this table led to Δ_{HAB}, the matrix of compositional dissimilarities between habitats. Like the overall BCI trajectory, habitat trajectories were fastest during the first surveys (Appendix S2: Fig. S1). Afterward, the swamp (W), the young forest (F), and stream sides (R) appeared to have faster compositional

dynamics than the other habitats, but this was an artifact of their smaller area. Angles between consecutive surveys were more variable than those of the overall BCI trajectory (Appendix S2: Fig. S1). An ordination plot of the habitat trajectories in Δ_{HAB} was not very informative because of large compositional distances between some of them (Fig. 7a). When we assessed the resemblance between habitat trajectories using D_{SDSP} (Eq. 13), the resulting dissimilarity matrix $\mathbf{D}_T$ very strongly correlated with the submatrix of Δ_{HAB} corresponding to the first BCI survey ($r = 0.975$; $P = 0.0014$ in a Mantel test). All
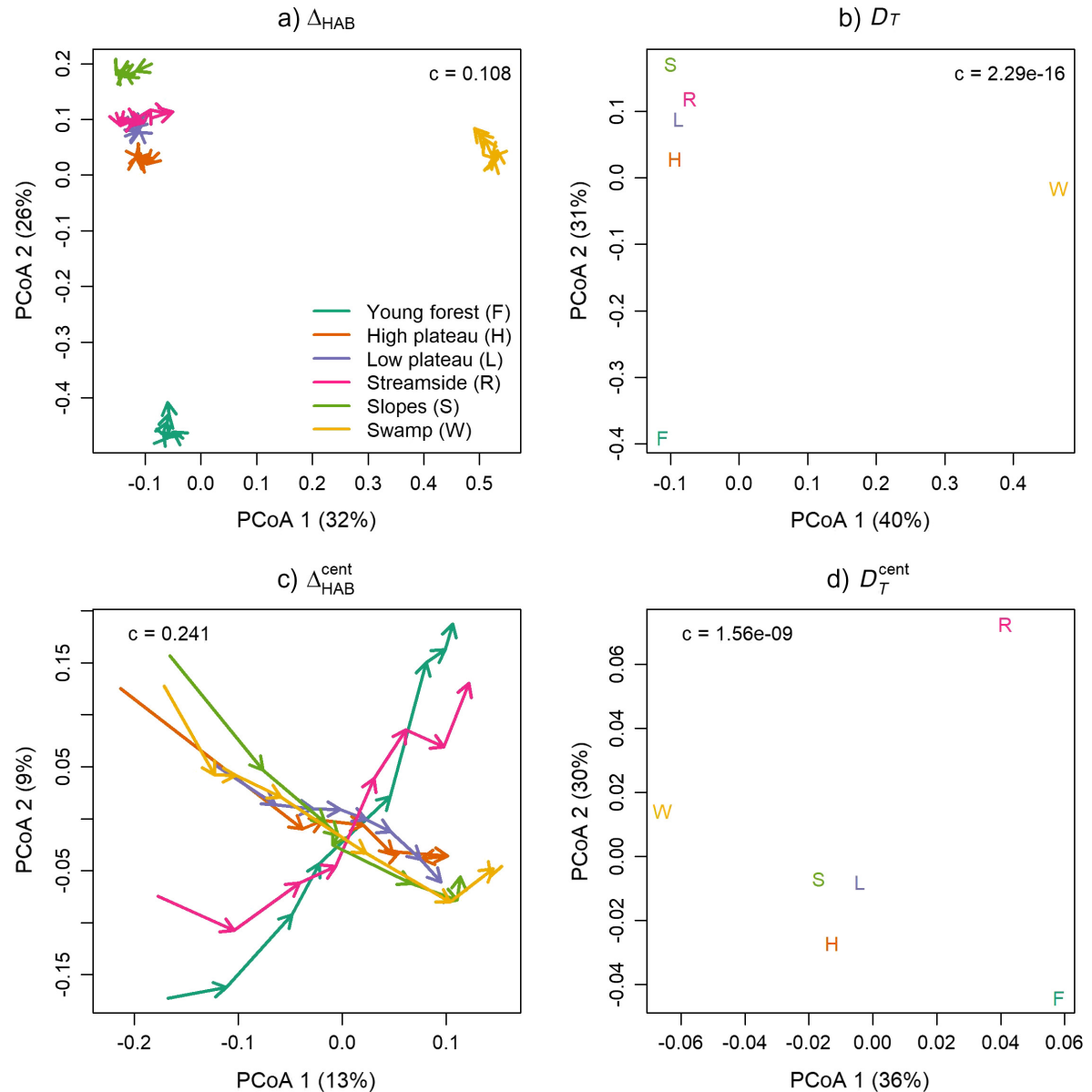
Fig. 7. Barro Colorado Island forest habitat trajectories. (a) Ordination of the habitat compositional states before centering (Δ_{HAB}), with trajectories indicated using arrows; (b) ordination of dissimilarities between trajectories (D_T) derived from Δ_{HAB}; (c) ordination of habitat compositional states after centering (Δ_{HAB}^{Cent}), with trajectories indicated using arrows; (d) ordination of dissimilarities between trajectories ($\mathbf{D}_T^{Cent}$) derived from Δ_{HAB}^{Cent}. The four ordinations were obtained by PCoA with correction factors for negative eigenvalues indicated in the top right corner of each panel.

these results indicated that the compositional variability among habitats obscured compositional dynamics and that $\mathbf{D}_T$ represented mostly time-invariant differences in composition between habitats (compare Fig. 7a and b). We thus centered trajectories to focus on compositional changes (i.e., differences in trajectory *shape*, *size*, and *direction*), obtaining matrix Δ_{HAB}^{Cent}. Compared to the (non-centered) Δ_{HAB}, matrix Δ_{HAB}^{Cent} allowed the shape and direction of habitat trajectories to be more easily displayed (Fig. 7a, c). Using Δ_{HAB}^{Cent} as input for D_{SDSP}, we recalculated the resemblance between habitat trajectories (Fig. 7d). The resulting matrix $\mathbf{D}_T^{cent}$ still correlated with the submatrix of Δ_{HAB} corresponding to the first survey to some extent ($r = 0.795$; $P = 0.026$ in a Mantel test), indicating that habitats similar in composition also tend to follow similar dynamics. The representation of $\mathbf{D}_T^{cent}$ (Fig. 7d) indicates that composition in the two plateaus (L and H) and slope (S) habitats is changing in a similar way (i.e., that the same drought-deciduous species are becoming more frequent in these habitats), whereas the swamp (W), stream sides (R), and young forest (F) are following more idiosyncratic compositional pathways (W appears to follow the same direction as L, H, and S in Fig. 7c, but the ordination plot only displays 21% of variance). The proposed asymmetric test of convergence/divergence between pairs of habitat trajectories indicated that BCI habitats had different trends (Table 2): F is compositionally approaching the old-growth BCI forest; H is becoming compositionally more similar to S; the composition in L, while moving in a similar direction to H, is at the same time slowly diverging from it; S does not show significant trends, but seems to be approaching the composition that the plateaus had in the first surveys; finally, the composition in W is approaching that of L and converging with R. Overall, convergence in composition occurred more frequently than divergence.

Lastly, we addressed compositional dynamics at the level of 20×20 m quadrats. For that we assembled a data table with 10,000 rows (1,250 quadrats times 8 surveys) and 328 columns (species) and built the corresponding dissimilarity matrix $\Delta_{20\times20}$. We estimated beta diversity (BD) following Legendre and De Cáceres (2013), for the eight submatrices of $\Delta_{20\times20}$ corresponding to the forest surveys. The resulting BD values (0.269,

0.257, 0.251, 0.249, 0.249, 0.248, 0.245, 0.244) exhibited a negative trend ($\tau = -0.929$; $P = 0.002$; Mann-Kendall test), indicating an overall compositional convergence within BCI. We used one-way ANOVA to test for differences in lengths and directionalities of quadrat trajectories among habitats. Shorter trajectory lengths (hence slower dynamics) occurred in quadrats of the high plateau and young forest ($F_{5, 1178} = 6.605$; $P < 0.0001$), which may be interpreted as drier conditions affecting tree growth rates. In contrast, trajectories were significantly more erratic (i.e., lower DIR values) in the swamp ($F_{5, 1178} = 6.712$; $P < 0.0001$), which could be an artifact of lower tree density in this area (Table 1).

Structural dynamics on permanent plots in a mountain beech forest

In this second example, we illustrate how the CTA framework can be used to summarize and characterize dynamics in forest structure, that is, changes in the distribution of tree sizes. This contrasts to the previous example that focused on dynamics in composition. The study area is mountainous and centered on the Craigieburn Range (Southern Alps), South Island, New Zealand ($43°10'$ S, $171°35'$ E; Fig. 8). *Fuscospora cliffortioides* (mountain beech) is the dominant tree species, usually forming monospecific stands. Previously, the forests consisted largely of mature stands, but a 1973 snowstorm damaged trees in 30% of stands (Harcombe et al. 1998). Profiting from woody debris created by snowfall damage and windstorms, an outbreak of the native pinhole beetle (*Platypus* spp., Coleoptera) and an associated fungal pathogen resulted in dispersed mortality of large trees and a decline in tree biomass over the following decade (Harcombe et al. 1998). This dispersed mortality meant that individual tree growth was strongly influenced by neighborhood competition as well as variation in site conditions (Coomes and Allen 2007). While forests were relatively unaffected by major disturbance events between 1983 and 1993, in 1994, an earthquake (M_w 6.7 in magnitude), with an epicenter 10 km northwest of the study area, caused substantial size-indiscriminant mortality of trees as a result of landslides (Allen et al. 1999, Hurst et al. 2011).

TABLE 2. Results of the asymmetric convergence/divergence test for combinations of habitats in Barro Colorado Island.

Habitat	F	H	L	R	S	W
Young forest (F)		−0.929**	−1.000**	−0.214[NS]	−1.000***	−0.857**
High plateau (H)	−0.071[NS]		0.333[NS]	−0.143[NS]	−0.714*	0.357[NS]
Low plateau (L)	0.071[NS]	0.810*		0.786**	0.600[NS]	−0.429[NS]
Stream sides (R)	−0.143[NS]	−0.286[NS]	0.286[NS]		−0.600[NS]	−0.929**
Slopes (S)	−0.214[NS]	−0.619†	−0.619†	−0.429[NS]		0.071[NS]
Swamp (W)	−0.571[NS]	−0.500[NS]	−0.786**	−0.643*	−0.067[NS]	

Notes: Values in the table correspond to the statistic (tau) and significance level of the Mann-Kendall test (Mann 1945) on the sequence of distances between states of the row habitat trajectory with respect to the column habitat trajectory. Negative values indicate that the row trajectory is approaching the column trajectory, and positive values indicate that it is moving away from it.
*$P \le 0.05$; **$P \le 0.01$; ***$P \le 0.001$; †$P \le 0.1$; NS, $P > 0.1$.
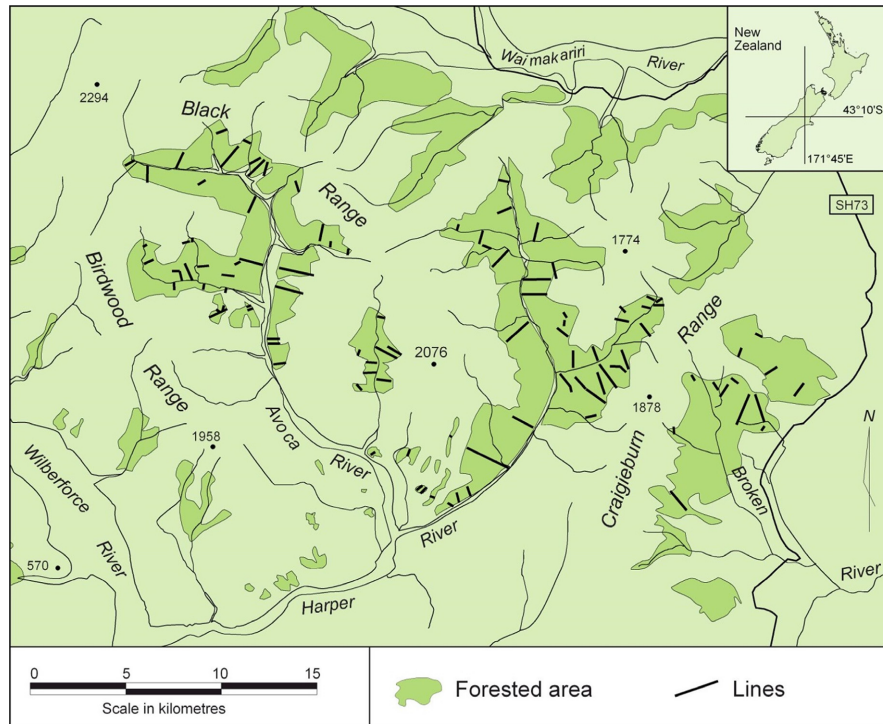
FIG. 8. Map of the study area for the analysis of structural dynamics on mountain beech forests. The location of transects (containing plots) in the study area is indicated using black lines.

The Craigieburn data set consists of 250 permanent plots sampling 9,000 ha of forest, established systematically along 98 compass lines between 1970 and 1972, with line origins located randomly along streams (Fig. 8). On each 20 × 20 m plot all tree stems with diameter at breast height >3 cm had their diameter measured and recorded; tree data from nine different surveys are used in this study: 1970–1972, 1974, 1978, 1983, 1987, 1993, 1999, 2004, and 2009. These were chosen from the full set of surveys to obtain a more or less homogeneous spacing between consecutive surveys (4–6 yr). The data set includes records of 12 tree species but 99.6% of them are from *F. cliffortioides*. Considering the 250 plots and nine surveys gives 250 × 9 = 2,250 tree community states, but three plots were completely destroyed by the 1994 earthquake, reducing the number of states to 2,241.

We used the CTA framework to summarize structural dynamics in this forested area, which result from the interplay between stand development, natural disturbances and subsequent recruitment. We calculated resemblance between plot data in terms of size structure, using the *cumulative abundance profile* (hereafter CAP) approach of De Cáceres et al. (2013). To account for differences in tree diameter, while emphasizing regeneration, we defined 19 quadratic diameter bins (in cm): {(2.25, 4], (4, 6.25], (6.25, 9] ... (110.25, 121]}; species composition was also taken into account (albeit not important for this example, given the dominance of *F. cliffortioides*) and we

took the number of stems to represent abundance. To reduce the influence of large numbers of individuals, CAPs were log-transformed before community resemblance calculations. Distance values between pairs of tree community states were calculated using a generalization of the Manhattan metric (De Cáceres et al. 2013) resulting in a 2,241 × 2,241 matrix $\boldsymbol{\Delta}$.

As before, we used PCoA on matrix $\boldsymbol{\Delta}$ to display community trajectories in the 250 forest plots (Fig. 9a). We then assessed the resemblance between the 250 plot trajectories using D_{SDSP}, and PCoA was again employed to display the resemblance matrix $\mathbf{D}_T$ (Fig. 9b). The dBD_{Total} was 4.17, and LCdBD values are displayed as point sizes in Fig. 7b. We conducted k-means clustering on the PCoA space of $\mathbf{D}_T$ to summarize trajectories of forest plots. For this, we used the R function cascadeKM in package vegan, where the SSI criterion suggested a clustering structure into 5, 13, or 19 groups. We chose the smallest value ($k = 5$) to facilitate interpretability. To characterize forest structural dynamics in each group we inspected the temporal changes in the median diameter-size class distribution of *F. cliffortioides* trees (Fig. 10); we also calculated the median basal area for each group by survey, the median segment length/speed for each group by segment, and the median angle for each group by consecutive segments (see Appendix S2: Fig. S2; Rydgren et al. 2004). Groups 1, 2, and 4 were characterized by tree growth (i.e., shifts over time toward larger diameter classes) and self-thinning (i.e., progressive decrease in
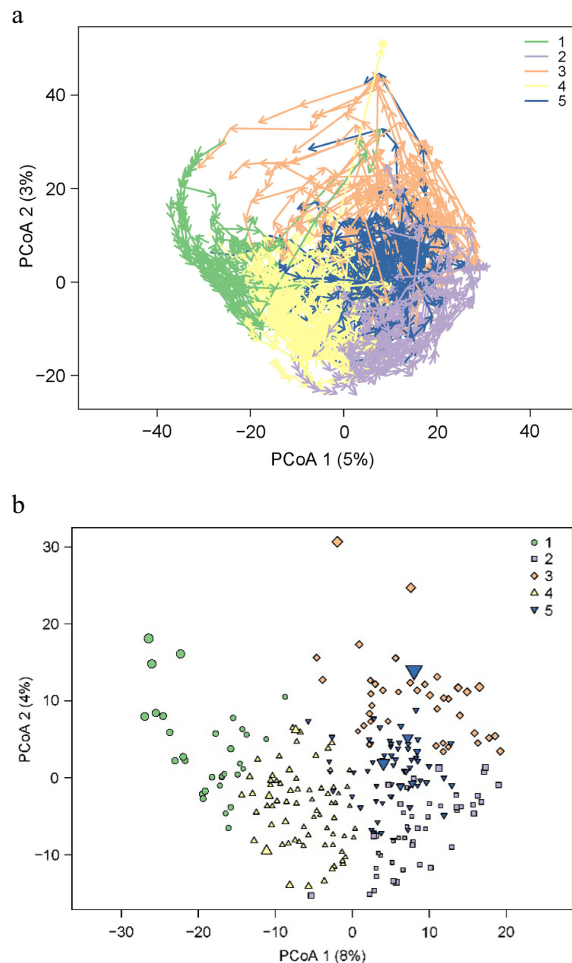
a

b

Fig. 9. (a) Ordination of tree community states (PCoA on Δ, after applying Cailliez's [1983] correction for negative eigenvalues; $c = 88.51$), with trajectories indicated using arrows; (b) ordination of plot trajectories in the space of structural dynamics (PCoA on D_T, after applying correction for negative eigenvalues; $c = 42.86$). Arrow colors in panel a and colored symbols in panel b indicate the five groups obtained using k means. Point sizes in panel b are proportional to LCdBD (local contributions to dynamic beta diversity) values.

the number of smaller, presumably shaded, individuals) as part of stand development. However, group 1 had higher basal area and exhibited a much narrower diameter distribution than group 4 (Fig. 10 and Appendix S2: Fig. S2), whereas group 2 was characterized by low numbers of trees spread across size classes. Like the former three, group 5 dynamics reflected growth and self-thinning, but the last two surveys showed the incorporation of new recruits. Finally, group 3 exhibited the fastest structural dynamics, including decreasing numbers of large individuals between 1970 and 1978, and a high level of recruitment since 1987 (Fig. 10 and Appendix S2: Fig. S2). Although we focused on structural dynamics in this example, we believe CTA is a useful framework to summarize the dynamic patterns of a set of sites by grouping them according to similarity in

dynamics in any kind of community property, provided the user makes an informed choice of d.

DISCUSSION

Advantages, applications, and limitations of the proposed framework

We presented here a framework to describe and analyze community dynamics based on geometric analysis of trajectories in a chosen space of community dissimilarity. Unlike approaches that focus on testing for particular kinds of community dynamics (Bagchi et al. 2017, Kalyuzhny and Shnerb 2017), or approaches that allow the response of communities to specific treatments to be examined with respect to dynamics under control conditions (Van Den Brink and Ter Braak 1999), CTA provides a more general way to describe and compare trajectories. Among other advantages, CTA allows the relationships in community dynamics to be summarized in dissimilarity values. As a result, any dissimilarity-based statistical tool designed for the (static) analysis of community variation can be also used to address questions of variation in community dynamics. Therefore, the goals of CTA are very different to those of recent statistical frameworks that pursue an explicit modelling and prediction of species interactions and community dynamics (Hampton et al. 2013, Thorson et al. 2016).

A key element of our contribution is the ability to evaluate resemblance in community dynamics between sampling units through the geometric comparison of their representation as trajectories, which leads to the definition of a space of trajectory resemblance. Clarke et al. (2006) developed a similar idea in their "second-stage" community analysis, a term coined by Somerfield and Clarke (1995). In their approach, the resemblance between a pair of trajectories $\mathbf{T}_1$ and $\mathbf{T}_2$ is measured by calculating the nonparametric Spearman correlation between the two dissimilarity matrices that describe each trajectory. This requires synchronous surveys of the sampling units, because one needs to match the elements of the two dissimilarity matrices (following our notation, one needs to define $\{d(x_{1i}, x_{1j}), d(x_{2i}, x_{2j})\}$ pairs, for all i and j surveys). As it discards the information contained in dissimilarities between states of one trajectory and states of the other (i.e., $d(x_{1i}, x_{2j})$ values), the Mantel correlation is appropriate to compare the *shape* of the two trajectories, but fails include several other geometric aspects. Moreover, the correlation will be -1 if the pathway of $\mathbf{T}_1$ goes from x to y and the pathway of $\mathbf{T}_2$ follows an exact inverse pathway from y to x, but it will be $+1$ if both trajectories have the same exact overall shape but $\mathbf{T}_1$ goes from x to y and $\mathbf{T}_2$ goes from x to z. Therefore, the *sense* is the only aspect of trajectory direction to which this approach is sensitive. Our framework extends the approach of Clarke et al. (2006) in at least three ways. First, CTA can be applied when trajectories differ in the number of surveys, or when survey times do not
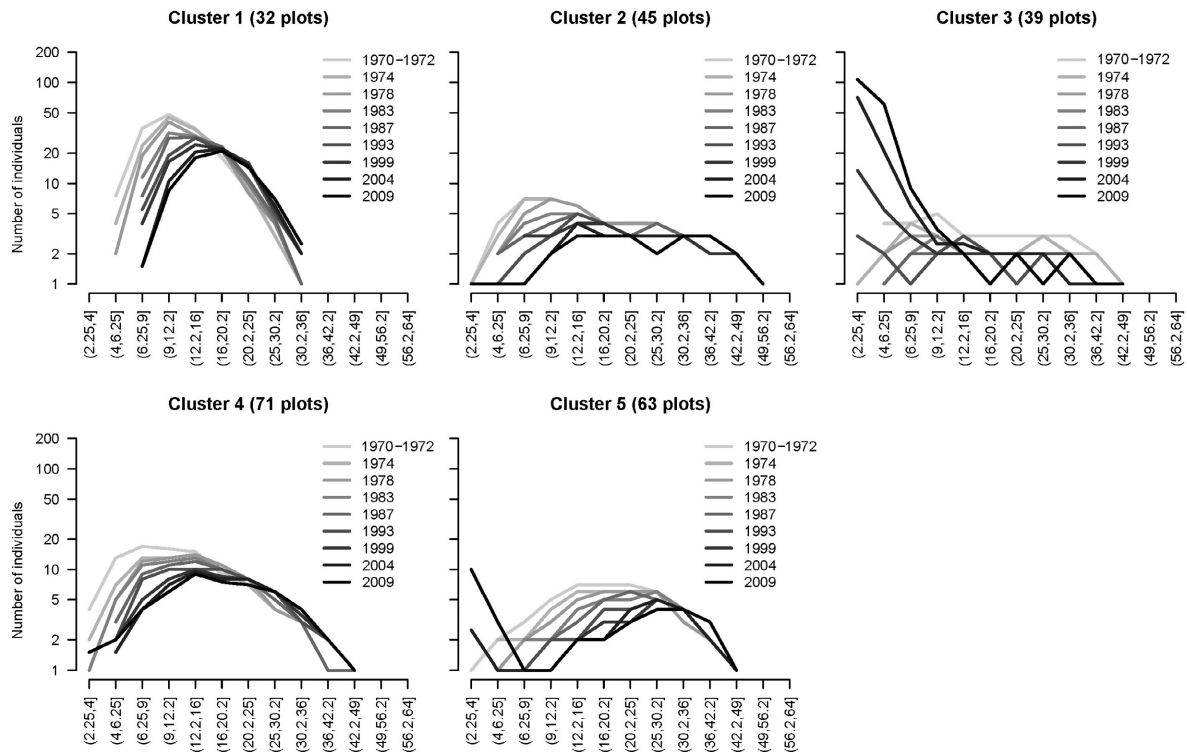
FIG. 10. Temporal changes in diameter class distribution for different forest plot groups. For each diameter class and group, the median across plots belonging to that group is shown. Different surveys are indicated using gray tones. Note log scale of y-axes.

match. Second, it allows differences in geometric aspects beyond *shape* (i.e., *position*, *size*, and *direction*) to be taken into account. Finally, in CTA, we showed how to compare compositional trajectories after centering them (i.e., after removing differences in trajectory *position*), which allows static spatial variation to be discarded to focus on the spatiotemporal component of community resemblance. Clarke et al. (2006) claimed that their approach allowed this, but in our opinion, it suffers from the limitation of not appropriately accounting for differences in *direction*, and neglecting differences in *size*, aspects of geometry that may be relevant to study spatiotemporal interaction.

There are multiple applications of CTA. (1) Basic geometric properties of community trajectories (length, speed, or overall directionality) can be compared across sites or regressed against explanatory factors. For example, it is increasingly accepted that the order and timing of species arrival during community assembly (i.e., priority effects) alters ecosystem structure and functioning (Tan et al. 2012), but it is much less clear how variation in community assembly history influences the rate at which trajectories diverge or the speed of community changes. (2) Projection of community states onto trajectories, asymmetric convergence/divergence tests, and trajectory asymmetric resemblance assessments are three tools that can be used to study the dynamics observed at particular sampling units *with respect to pathways*

defined a priori from known models of community dynamics, such as models of primary/secondary succession or seasonal cyclical dynamics. (3) Classification and ordination analyses on (symmetric) matrix $\mathbf{D}_T$ can be used to display and summarize patterns of community dynamics in a set of monitored sites (or the output of simulation models), as illustrated here for forest structural dynamics. (4) Instead of comparing community dynamics *among sampling units*, dBD_{Total} values (and local contributions) can be used to compare the overall variation in community dynamics *among sets of sampling units* surveyed with comparable sampling methods, analogously to the comparison of beta diversity across forests (De Cáceres et al. 2012). (5) In the same way that multivariate regression models are routinely used in ecology to hypothesize drivers of observed static patterns (Legendre and Legendre 2012), the application of distance-based regression frameworks, such as PERMANOVA (Anderson 2001, 2017), on matrix $\mathbf{D}_T$ could be used to test hypotheses about the causes of variation in community dynamics. This use of CTA was already illustrated by Clarke et al. (2006), who employed one-way ANOSIM (Clarke and Green 1988) to test for the consistency of marine benthic community dynamics across sites. The numerous papers applying their method (Caro et al. 2010, Kroeker et al. 2013, Schulz et al. 2013) demonstrate the usefulness of analyzing the variation in $\mathbf{D}_T$ against all kinds of factors. Our contribution

provides a more robust basis for defining $\mathbf{D}_T$ and a broader framework within which to naturally embed such analyses. When analyzing the variation in $\mathbf{D}_T$, it is important to remember that the beginning and end of a trajectory is dictated by sampling and is often arbitrary (Dornelas et al. 2012). This means the influence of the starting point must be considered when interpreting community dynamics in terms of potential predictors.

Regardless of the application, users of CTA should also be aware of the multiple factors that may affect the geometry of community trajectories, most importantly the choice of d but also the different artifacts derived from spatiotemporal sampling decisions and their interaction with ecological processes. To be able to compare CTA results among different data sets, trajectory analyses should be conducted after homogenizing as much as possible the size of sampling units (e.g., by aggregation of data in data sets whose sampling unit size is smallest) and the frequency of surveys (e.g., by selecting surveys of data sets with the highest temporal resolution), so that the effects of temporal and spatial resolution affect all data sets in the same way. Null models of community dynamics coupled with of sampling decisions may also be helpful as benchmarks to provide context for CTA results in real data sets, as we did in our simulation study.

Improvements and extensions of the current framework

When developing our framework, we evaluated the degree to which procedures from spatial movement analysis or spatial trajectory data mining should be modified to be used in the multidimensional spaces typical of community ecology. This led us to propose new definitions of trajectory angles and overall directionality and to define indices to compare trajectories while incorporating the direction of segments D_{DS} and D_{SDSP}. The behavior of DIR and D_{SDSP} was satisfactory in our study of spatiotemporal dynamics using simulations and in the two examples above, but better alternatives may be defined. Hence, additional work will be needed to decide which properties are desirable for the appropriate comparison of the dynamics of ecological communities. In the same way, we calculated D_{SDSP} after centering of trajectories to discard differences in trajectory *position*; methods related to Procrustes analysis could be tested and added to our framework to compare trajectories solely in terms of their *shape* (Adams and Collyer 2009).

A potential extension of this framework concerns the analysis of matrix $\mathbf{D}_S$, containing dissimilarities between segments. Instead of clustering trajectories, we could have chosen to group directed segments, from the same or different trajectories, to define what could be called "consensus community trajectories." In trajectory data mining, spatial consensus trajectories have been obtained using density-based cluster analysis applied to $\mathbf{D}_S$ (Lee et al. 2007, Tripathi et al. 2016). An example of consensus community trajectory in forest structural dynamics is the one representing stand development.

As in other interdisciplinary tools such as cluster analysis, literature on methods of trajectory analysis lies dispersed across different disciplines (mathematics, engineering, movement ecology, evolution, etc.). Therefore, trajectory analysis methods of which we are unaware may be adapted and added to the CTA toolbox. Searching the literature of these related fields may be a fruitful endeavor to enlarge and improve the present framework. Further, although we presented our framework as focused on community ecology, CTA can readily be used to study trajectories in other fields where temporal series are multivariate or where spatial transects result in trajectories in multivariate spaces (Adams and Collyer 2009, Lohman et al. 2017).

Literature Cited

Adams, D. C., and M. L. Collyer. 2009. A general framework for the analysis of phenotypic trajectories in evolutionary studies. Evolution 63:1143–1154.

Allen, R. B., P. J. Bellingham, and S. K. Wiser. 1999. Immediate damage by an earthquake to a temperate montane forest. Ecology 80:708–714.

Allison, G. 2004. The influence of species diversity and stress intensity on community resistance and resilience. Ecological Monographs 74:117–134.

Anderson, M. J. 2001. Permutation tests for univariate or multivariate analysis of variance and regression. Canadian Journal of Fisheries and Aquatic Sciences 58:626–639.

Anderson, M. J. 2006. Distance-based tests for homogeneity of multivariate dispersions. Biometrics 62:245–253.

Anderson, M. J. 2017. Permutational multivariate analysis of variance (PERMANOVA). Wiley StatsRef: Statistics Reference Online. Pages 1–15. https://doi.org/10.1002/9781118445112.stat07841

Anderson, M. J., et al. 2011. Navigating the multiple meanings of β diversity: a roadmap for the practicing ecologist. Ecology Letters 14:19–28.

Angeler, D. G., O. Viedma, and J. M. Moreno. 2009. Statistical performance and information content of time lag analysis and redundancy analysis in time series modeling. Ecology 90:3245–3257.

Austin, M. P. 1977. Use of ordination and other multivariate descriptive methods to study succession. Plant Ecology 35:165–175.

Bagchi, S., N. J. Singh, D. D. Briske, B. T. Bestelmeyer, M. P. McClaran, and K. Murthy. 2017. Quantifying long-term plant community dynamics with movement models: implications for ecological resilience. Ecological Applications 27:1514–1528.

Besse, P., B. Guillouet, J.-M. Loubes, and R. François. 2016. Review and perspective for distance based trajectory clustering. IEEE Transactions on Intelligent Transportation Systems 17:3306–3317.

Borg, I., and P. J. F. Groenen. 2005. Modern multidimensional scaling: theory and applications. Statistical modelling. Springer, New York, New York, USA.

Bray, R. J., and J. T. Curtis. 1957. An ordination of the upland forest communities of southern Wisconsin. Ecological Monographs 27:325–349.

Buchin, K., M. Buchin, M. Van Kreveld, M. Löffler, R. I. Silveira, C. Wenk, and L. Wiratma. 2013. Median trajectories. Algorithmica 66:595–614.

Cailliez, F. 1983. The analytical solution of the additive constant problem. Psychometrika 48:305–308.

Caro, A. U., S. A. Navarrete, and J. C. Castilla. 2010. Ecological convergence in a rocky intertidal shore metacommunity despite high spatial variability in recruitment regimes. Proceedings of the National Academy of Sciences USA 107:18528–18532.

Chase, J. M. 2003. Community assembly: When should history matter? Oecologia 136:489–498.

Clarke, K. R., and R. H. Green. 1988. Statistical design and analysis for a "biological effects" study. Marine Ecology 46:213–226.

Clarke, K. R., P. J. Somerfield, L. Airoldi, and R. M. Warwick. 2006. Exploring interactions by second-stage community analyses. Journal of Experimental Marine Biology and Ecology 338:179–192.

Collins, S. L., F. Micheli, and L. Hartt. 2000. A method to determine rates and patterns of variability in ecological communities. Oikos 91:285–293.

Condit, R., S. Hubbell, and R. B. Foster. 1995. Mortality rate of 205 Neotropical tree and shrub species and the impact of a severe drought. Ecological Monographs 65:419–439.

Condit, R., S. P. Hubbel, and R. B. Foster. 1996. Changes in tree species abundance in a Neotropical forest: impact of climate change. Journal of Tropical Ecology 12:231–256.

Condit, R., R. A. Chisholm, and S. P. Hubbell. 2012a. Thirty years of forest census at Barro Colorado and the importance of immigration in maintaining diversity. PLoS ONE 7:1–6.

Condit, R., S. Lao, R. Pérez, S. B. Dolins, R. B. Foster, and S. P. Hubbell. 2012b. Barro Colorado Forest Census Plot Data, 2012 Version. https://doi.org/10.5479/data.bci.20130603

Coomes, D. A., and R. B. Allen. 2007. Mortality and tree-size distributions in natural mixed-age forests. Journal of Ecology 95:27–40.

De Cáceres, M., X. Font, and F. Oliva. 2010. The management of vegetation classifications with fuzzy clustering. Journal of Vegetation Science 21:1138–1151.

De Cáceres, M., P. Legendre, and F. He. 2013. Dissimilarity measurements and the size structure of ecological communities. Methods in Ecology and Evolution 4:1167–1177.

De Cáceres, M., et al. 2012. The variation of tree beta diversity across a global network of forest plots. Global Ecology and Biogeography 21:1191–1202.

Dornelas, M., N. J. Gotelli, B. J. McGill, H. Shimadzu, F. Moyes, C. Sievers, and A. E. Magurran. 2014. Assemblage time series reveal biodiversity change but not systematic loss. Science 344:296–300.

Dornelas, M., et al. 2012. Quantifying temporal change in biodiversity: challenges and opportunities. Proceedings of the Royal Society B 280:20121931.

Dray, S., G. Blanchet, D. Borcard, T. Guenard, G. Jombart, G. Larocque, P. Legendre, N. Madi, and H. H. Wagner. 2017. adespatial: multivariate multiscale spatial analysis. R package version 0.3-2. https://CRAN.R-project.org/package=adespatial

Dray, S., et al. 2012. Community ecology in the age of multivariate multiscale spatial analysis. Ecological Monographs 82:257–275.

Everitt, B. S., S. Landau, M. Leese, and D. Stahl. 2011. Cluster analysis. Fifth edition. John Wiley & Sons, Chichester, UK.

Foster, B., and D. Tilman. 2000. Dynamic and static views of succession: testing the descriptive power of the chronosequence approach. Plant Ecology 146:1–10.

Fukami, T., T. M. Bezemer, S. R. Mortimer, and W. H. Van Der Putten. 2005. Species divergence and trait convergence in experimental plant community assembly. Ecology Letters 8:1283–1290.

Gower, J. C. 1966. Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika 53:325–338.

Gower, J. C., and P. Legendre. 1986. Metric and Euclidean properties of dissimilarity coefficients. Journal of Classification 3:5–48.

Halpern, C. B. 1989. Early successional patterns of forest species: interactions of life history traits and disturbance. Ecology 70:704–720.

Hampton, S. E., E. H. Holmes, L. P. Scheef, M. D. Scheuerell, S. L. Katz, D. E. Pendleton, and E. J. Ward. 2013. Quantifying effects of abiotic and biotic drivers on community dynamics with multivariate autoregressive (MAR) models. Ecology 94:1–4.

Harcombe, P. A., R. B. Allen, J. A. Wardle, and K. H. Platt. 1998. Spatial and temporal patterns in stand structure, biomass, growth and mortality in a monospecific Nothofagus solandri var. cliffortioides (Hook. f.) Poole forest in New Zealand. Journal of Sustainable Forestry 6:313–343.

Harms, K. E., R. Condit, S. P. Hubbell, and R. B. Foster. 2001. Habitat associations of trees and shrubs in a 50-ha neotropical forest plot. Journal of Ecology 89:947–959.

Hooten, M. B., D. S. Johnson, B. T. McClintock, and J. M. Morales. 2017. Animal movement. Statistical models for telemetry data. CRC Press, Boca Raton, Florida, USA.

Hughes, J. M. R. 1990. Lotic vegetation dynamics following disturbance along the Swan and Apsley Rivers, Tasmania, Australia. Journal of Biogeography 17:291–306.

Hurst, J. M., R. B. Allen, D. A. Coomes, and R. P. Duncan. 2011. Size-specific tree mortality varies with neighbourhood crowding and disturbance in a montane Nothofagus forest. PLoS ONE 6:e26670.

Kalyuzhny, M., and N. M. Shnerb. 2017. Dissimilarity-overlap analysis of community dynamics: opportunities and pitfalls. Methods in Ecology and Evolution 2017:1764–1773.

Kroeker, K. J., M. C. Gambi, and F. Micheli. 2013. Community dynamics and ecosystem simplification in a high-CO2 ocean. Proceedings of the National Academy of Sciences USA 110:12721–12726.

Lee, J.-G., J. Han, and K.-Y. Whang. 2007. Trajectory clustering: a partition-and-group framework. Pages 593–604 in Proceedings of the 2007 ACM SIGMOD international conference on Management of data—SIGMOD '07. ACM, New York, New York, USA.

Legendre, P., and M. J. Anderson. 1999. Distance-based redundancy analysis: testing multispecies responses in multifactorial ecological experiments. Ecological Monographs 69:1–24.

Legendre, P., and M. De Cáceres. 2013. Beta diversity as the variance of community data: dissimilarity coefficients and partitioning. Ecology Letters 16:951–963.

Legendre, P., M. De Cáceres, and D. Borcard. 2010. Community surveys through space and time: testing the space-time interaction in the absence of replication. Ecology 91:262–272.

Legendre, P., and O. Gauthier. 2014. Statistical methods for temporal and space–time analysis of community composition data. Proceedings of the Royal Society B 281:20132728.

Legendre, P., and L. Legendre. 2012. Numerical ecology. Third edition. Elsevier Science BV., Amsterdam, The Netherlands.

Lingoes, J. C. 1971. Some boundary conditions for a monotone analysis of symmetric matrices. Psychometrika 36:195–203.

Lohman, B. K., D. Berner, and D. I. Bolnick. 2017. Clines arc through multivariate morphospace. American Naturalist 189:354–367.

Magalhães, M. F., P. Beja, I. J. Schlosser, and M. J. Collares-Pereira. 2007. Effects of multi-year droughts on fish assemblages of seasonally drying Mediterranean streams. Freshwater Biology 52:1494–1510.

Mann, H. 1945. Nonparametric tests against trend. Econometrica 13:245–259.

Matthews, W. J., E. Marsh-Matthews, R. C. Cashner, and F. Gelwick. 2013. Disturbance and trajectory of change in a stream fish community over four decades. Oecologia 173:955–969.

McEwan, R. W., J. M. Dyer, and N. Pederson. 2011. Multiple interacting ecosystem drivers: toward an encompassing hypothesis of oak forest dynamics across eastern North America. Ecography 34:244–256.

Ovaskainen, O., G. Tikhonov, D. Dunson, V. Grøtan, S. Engen, B.-E. Sæther, and N. Abrego. 2017. How are species interactions structured in species-rich communities? A new method for analysing time-series data. Proceedings of the Royal Society B 284:20170768.

Pewsey, A., M. Neuhäuser, and G. D. Ruxton. 2013. Circular statistics in R. Oxford University Press, Oxford, UK.

Philippi, T., P. Dixon, and B. Taylor. 1998. Detecting trends in species composition. Ecological Applications 8:300–308.

Pickett, S. T. A., S. L. Collins, and J. J. Armesto. 1987. Models, mechanisms and pathways of succession. Botanical Review 53:335–371.

Pickett, S. T. A., and P. S. White. 1985. The ecology of natural disturbance and patch dynamics. Academic Press, Orlando, Florida, USA.

Reyer, C. P. O., et al. 2015. Forest resilience and tipping points at different spatio-temporal scales: approaches and challenges. Journal of Ecology 103:5–15.

Rydgren, A. K., R. H. Økland, and G. Hestmark. 2004. Disturbance severity and community resilience in a boreal forest. Ecology 85:1906–1915.

Schaefer, J., K. Gido, and M. Smith. 2005. A test for community change using a null model approach. Ecological Applications 15:1761–1771.

Schulz, K. G., et al. 2013. Temporal biomass dynamics of an Arctic plankton bloom in response to increasing levels of atmospheric carbon dioxide. Biogeosciences 10:161–180.

Shimadzu, H., M. Dornelas, and A. E. Magurran. 2015. Measuring temporal turnover in ecological communities. Methods in Ecology and Evolution 6:1384–1394.

Smith, S. D. P. 2012. Identifying and evaluating causes of alternative community states in wetland plant communities. Oikos 121:675–686.

Somerfield, P. J., and K. R. Clarke. 1995. Taxonomic levels in marine communities revisited. Marine Ecology Progress Series 127:113–119.

Takeuchi, Y., H. Ochi, M. Kohda, D. Sinyinza, and M. A. Hori. 2010. A 20-year census of a rocky littoral fish community in Lake Tanganyika. Ecology of Freshwater Fish 19:239–248.

Tan, J., Z. Pu, W. A. Ryberg, and L. Jiang. 2012. Species phylogenetic relatedness, priority effects, and ecosystem functioning. Ecology 93:1164–1172.

Thibault, K., E. White, S. Ernest, and D. R. Community. 2004. Temporal dynamics in the structure and composition of a desert rodent community. Ecology 85:2649–2655.

Thorson, J. T., J. N. Ianelli, E. A. Larsen, L. Ries, M. D. Scheuerell, C. Szuwalski, and E. F. Zipkin. 2016. Joint dynamic species distribution models: a tool for community ordination and spatio-temporal monitoring. Global Ecology and Biogeography 25:1144–1158.

Trexler, J. C., W. F. Loftus, and S. Perry. 2005. Disturbance frequency and community structure in a twenty-five year intervention study. Oecologia 145:140–152.

Tripathi, P. K., M. Debnath, and R. Elmasri. 2016. A direction based framework for trajectory data analysis. Pages 1–8 in Proceedings of the 9th ACM International Conference on PErvasive Technologies Related to Assistive Environments – PETRA '16. ACM, New York, New York, USA.

Turchin, P. 1998. Quantitative analysis of movement. Sinauer Associates, Sunderland, Massachusetts, USA.

Van Den Brink, P. J., and C. J. F. Ter Braak. 1999. Principal response curves: analysis of time-dependent multivariate responses of biological community to stress. Environmental Toxicology and Chemistry 18:138–148.

Vellend, M. 2016. The theory of ecological communities. Princeton University Press, Princeton, New Jersey, USA.

Vlachos, M., G. Kollios, and D. Gunopulos. 2002. Discovering similar multidimensional trajectories. Pages 673–684 in Proceedings 18th International Conference on Data Engineering. IEEE, San Jose, California, USA.

Walters, K., and L. D. Coen. 2006. A comparison of statistical approaches to analyzing community convergence between natural and constructed oyster reefs. Journal of Experimental Marine Biology and Ecology 330:81–95.

Whittaker, R. H. 1960. Vegetation of the Siskiyou mountains, Oregon and California. Ecological Monographs 30:279–338.

Zheng, Y. U. 2015. Trajectory data mining: an overview. ACM Transactions on Intelligent Systems and Technology 6:1–41.

Supporting Information

Additional supporting information may be found online at: http://onlinelibrary.wiley.com/doi/10.1002/ecm.1350/full

Data Availability

The Craigieburn forest data are archived in the New Zealand National Vegetation Databank and can be located using the search facility available at http://nvs.landcareresearch.co.nz by searching for "HARPER/AVOCA FOREST." Data corresponding to the first seven surveys of BCI are available on the DSpace Repository at https://doi.org/10.5479/data.bci.20130603 (Condit et al. 2012b).

Appendix S1. Hausdorff and segment path distances between two trajectories

The Hausdorff distance D_H (Hausdorff 1914) is a metric that measures the distance between two subsets of a metric space. Informally, two sets are close in the Hausdorff distance if every point of either set is close to some point of the other set. For every point in one set, the infimum distance to any point of the other set is computed, and the Hausdorff distance is the supremum of all these distances. Given X and Y subsets, the general definition is:

$$D_H(X,Y) = \max\{\sup_{x\in X}\inf_{y\in Y} d(x,y), \sup_{y\in Y}\inf_{x\in X} d(x,y)\} \ . \qquad \text{(Eq. S1)}$$

The Hausdorff distance is hard to calculate but in the case of polygonal curves, like trajectories or segments, calculations are simplified (Besse et al. 2016). The Hausdorff distance between two segments is given in eq. 8 of the main text. Let $T_1 = \{(x_{11},t_{11}),(x_{12},t_{12}),\ldots,(x_{1n},t_{1n})\}$ and $T_2 = \{(x_{21},t_{21}),(x_{22},t_{22}),\ldots,(x_{2m},t_{2m})\}$ be two trajectories of lengths n and m, to be compared. The Hausdorff distance between two trajectories, D_{HT}, is the maximum among D_{ps} values between points of one trajectory and segments of the other (Besse et al. 2016):
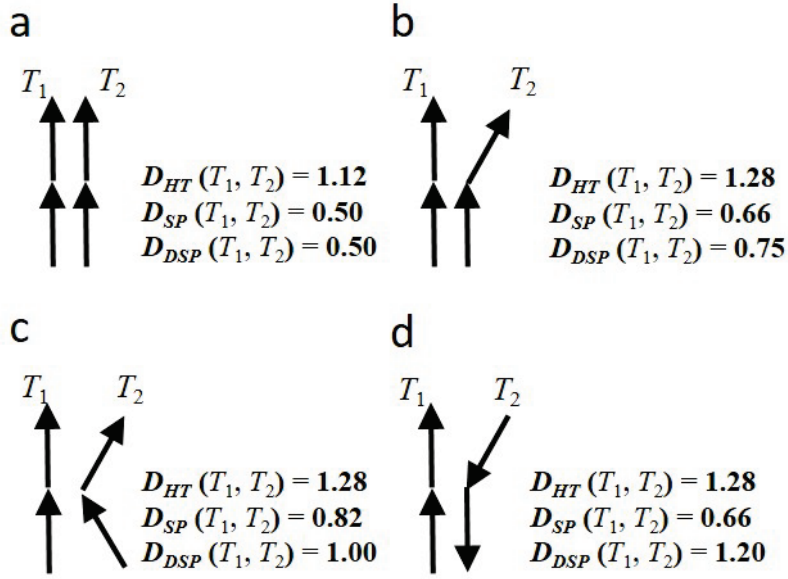
$$D_{HT}(T_1,T_2) = \max\{\max_{\substack{i\in\{1\ldots n\}\\ j\in\{1\ldots m-1\}}}\{D_{ps}(x_{1i},s_{2j})\}, \max_{\substack{i\in\{1\ldots n-1\}\\ j\in\{1\ldots m\}}}\{D_{ps}(x_{2j},s_{1i})\}\} \ . \qquad \text{(Eq. S2)}$$

D_{HT} will always be equal to the largest D_{ps} value, regardless of the relationship between other parts of the trajectories (Fig. S1.1). Other indices have been proposed to avoid this limitation (Lin and Su 2005). For example, the *segment path distance*, D_{SP}, from trajectory T_1 to trajectory T_2 is the average of all distances from points composing T_1 to trajectory T_2 (Besse et al. 2016):

$$D_{SP}(T_1,T_2) = \frac{1}{n}\sum_{i=1}^{n} D_{pt}(x_{1i},T_2) \ . \qquad \text{(Eq. S3)}$$

While D_{SP} is sensitive to changes in the position of every point (Fig. S1.1), neither D_{HT} or D_{SP} take into account the direction of segments. The lack of D_{HT} and D_{SP} responsiveness to trajectory direction makes them inappropriate to assess the resemblance in community dynamics (Fig. S1).

Fig. S1. Examples of values of D_{HT}, D_{SP} and D_{DSP} (eq. 11 in main text) for different pairs of trajectories, assuming Cartesian coordinates and the Euclidean distance for d.


a

$D_{HT}(T_1, T_2) = 1.12$
$D_{SP}(T_1, T_2) = 0.50$
$D_{DSP}(T_1, T_2) = 0.50$

b

$D_{HT}(T_1, T_2) = 1.28$
$D_{SP}(T_1, T_2) = 0.66$
$D_{DSP}(T_1, T_2) = 0.75$

c

$D_{HT}(T_1, T_2) = 1.28$
$D_{SP}(T_1, T_2) = 0.82$
$D_{DSP}(T_1, T_2) = 1.00$

d

$D_{HT}(T_1, T_2) = 1.28$
$D_{SP}(T_1, T_2) = 0.66$
$D_{DSP}(T_1, T_2) = 1.20$

References

Besse, P., B. Guillouet, J.-M. Loubes, and R. François. 2016. Review and perspective for distance based trajectory clustering. IEEE Transactions on Intelligent Transportation Systems 17:3306–3317.

Hausdorff, F. 1914. Grundzüge der Mengenlehre. Veit, Leipzig.

Lin, B., and J. Su. 2005. Shapes based trajectory queries for moving objects. Proceedings of the 2005 International Workshop on Geographic Information Systems (GIS '05):21–30.

Appendix S2. Additional figures
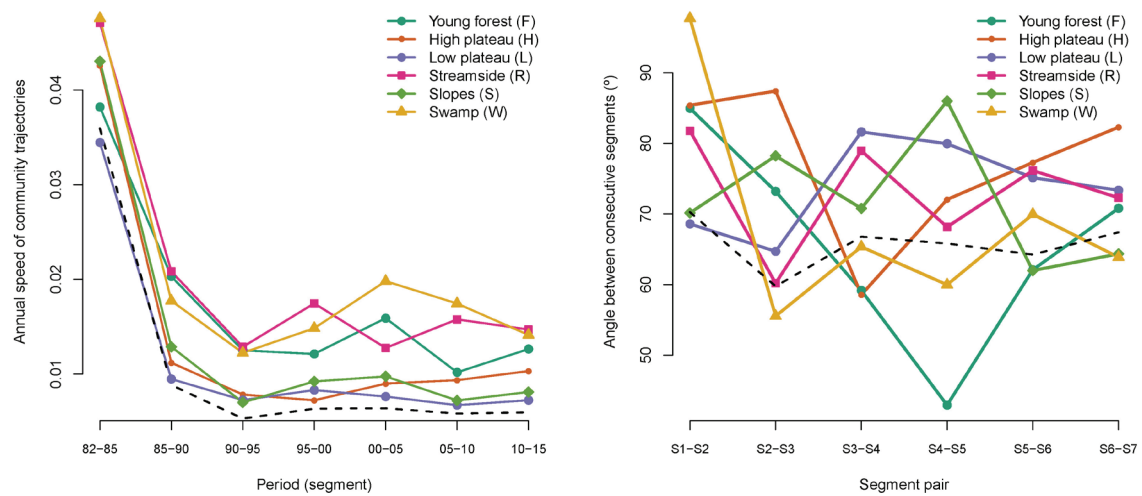
Fig. S1 Characteristics of habitat trajectories in BCI: Annual trajectory speed per habitat and segment (left) and angle between consecutive segments per habitat (right). Black dashed lines indicate the characteristics of overall BCI forest trajectory.
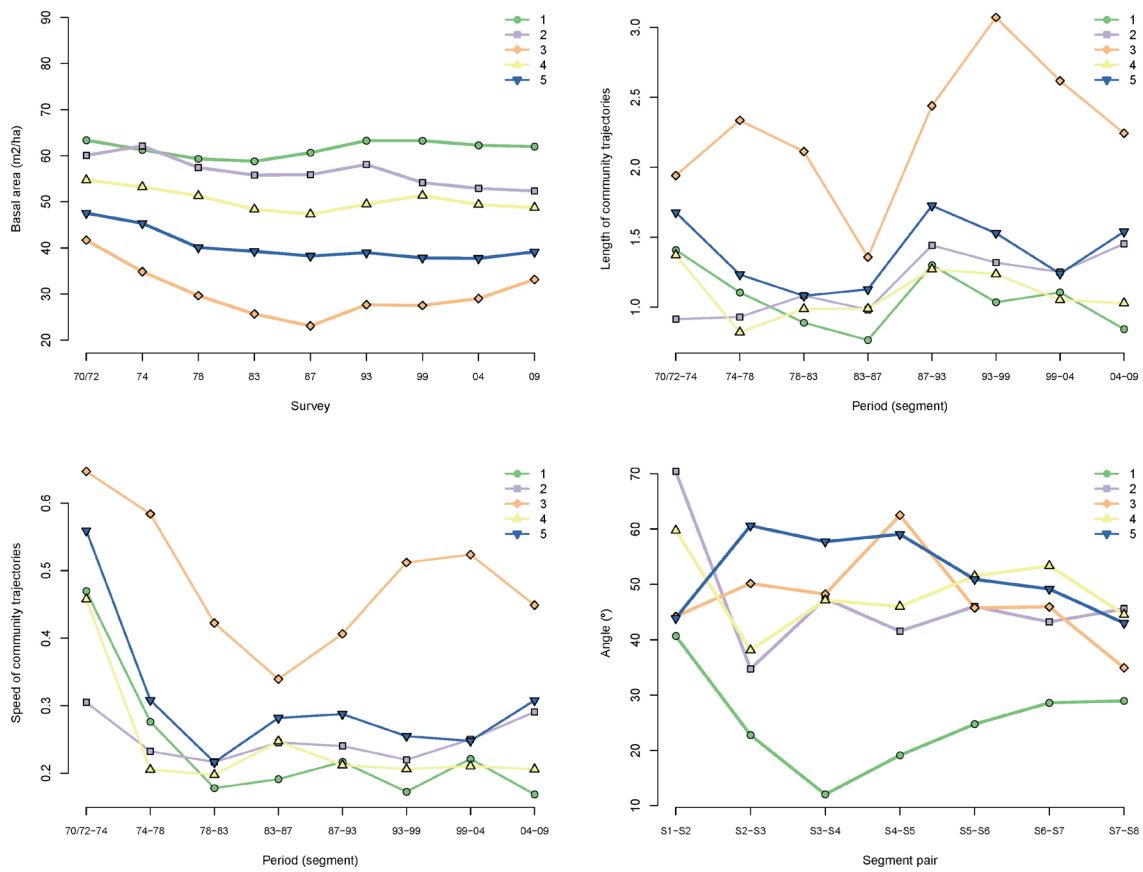
Fig. S2 Characterization of structural trajectories in mountain beech forests of Cragieburn range, New Zealand. Median basal area ($m^2 \cdot ha^{-1}$) per group and survey (upper left), median length and speed of structural trajectories (upper right and lower left, respectively) per group and segment, and median angle per group and consecutive segment (lower right).

Appendix S3. Details of the simulation study

Section S1 Simulated community dynamics

We used simulated data to illustrate the behavior of community trajectory analysis (CTA) under different kinds of compositional dynamics, while addressing the variation of results depending on sampling decisions, i.e. the size of sampling units and the frequency of surveys. To model community composition, we considered a theoretical environmental space of two dimensions (gradients). Species responses along this space were simulated using the R package 'coenocliner' (Simpson 2018), assuming that gradients were uncorrelated and species response followed Gaussian curves (see Fig. S1). The gradients had boundaries between -5 and 5 units and species optimum positions were distributed uniformly along them. Tolerance parameters were set to two units for all species, and optimum abundances followed a lognormal distribution with mean three (in the logarithm space).

Fig. S1 Example of simulated response (Gaussian shape) of four random species along the two environmental axes.
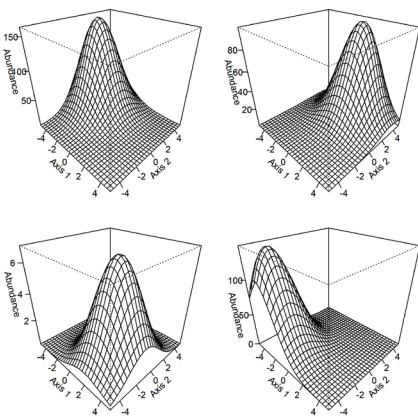
We considered simulations of either 20 or 50 species but results were qualitatively similar and we report the results on spaces with 50 species only. To determine the initial community state, we used the mean species response (i.e. without adding a source of error

for counts, such as Poisson or Negative Binomial distributions) at the center of the environmental space (i.e. at coordinates (0,0)). Species abundances were rescaled to obtain a total number of individuals equal to a fixed carrying capacity (CC) of the community (i.e. the maximum number of individuals that can occupy the sampled area). In our simulations CC can be also interpreted as the size of sampling units, because a larger sampled area implies that more individuals will be found. We considered four levels of CC: **20**, **80**, **240** and **960** individuals.

To simulate dynamics, we considered **mortality** and **recruitment** (including both local birth and immigration) as population processes (Vellend 2016). Assuming a constant rate of replacement of individuals, we defined the temporal step of simulated dynamics as that corresponding to the replacement of 5% of individuals in the community (i.e. 1, 4, 12 and 12 individuals for CC = 20, 80, 240 and 960, respectively). Mortality followed a multinomial distribution with species probabilities being proportional to the extant individuals at any given step, hence all individuals had the same probability of dying (i.e. mortality was simulated as in a neutral community). However, recruitment was also modelled using a multinomial distribution, but with probabilities depending on species identity. Recruitment probabilities depended on the type of community dynamics considered:

1. *Stabilizing selection*: Recruitment with probabilities proportional to the response of species at the center of the gradient (0, 0). At each step, 5% of the individuals of the community were replaced (i.e. 5% of deaths and the same number of recruitment events). Note that the resulting dynamics cannot be called *neutral*, because species identity was taken into account for recruitment, according to the simulated species response in the environmental space. Nevertheless, community dynamics include drift around the attractor point determined by recruitment probabilities.
2. *Post-disturbance recovery*: Before starting the simulation we eliminated 80% of the community, and subsequent steps were designed to recover the (pre-disturbance) state. To allow a recovery in the number of individuals, post-disturbance death rate was proportionally reduced according to the ratio between community size and the CC. Recruitment probabilities were simulated as in the previous case.
3. *Directional selection*: Progressive change of the community from the species composition at the center of the gradient (0, 0) towards a different community state, enforced by recruitment of different probabilities. As in the stabilizing selection case, at each step 5% of the individuals of the community were replaced. Mortality was modelled as before, but recruitment was biased towards a different target community state. Recruitment probabilities were taken from the response of species in a position of the environmental space lying 5 units away from the center of the gradient.

For each type of community dynamics, we simulated datasets of **sixteen** communities. All sixteen community dynamics were simulated in the same way for stabilizing selection and
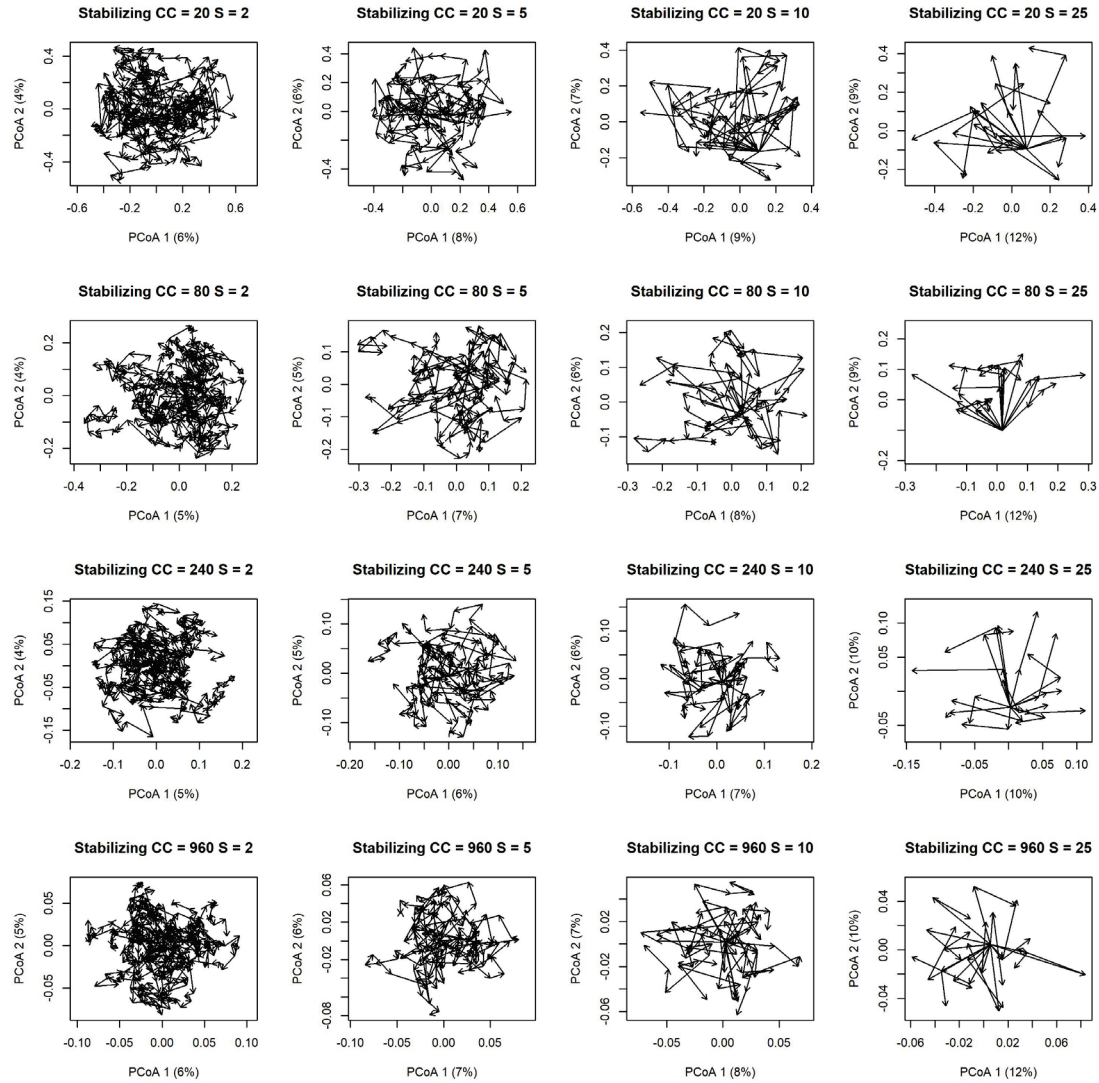
disturbance-recovery dynamics. However, in the case of directional dynamics *pairs of trajectories* were set to point to different target positions of the environmental space, in the direction of the eight cardinal points (N, NE, E, SE, S, SW, W and NW).

We simulated community dynamics for **50 steps**, which equals to potentially replacing the whole population more than twice. **20 replicates** of the sixteen trajectories were simulated for each combination of dynamics' type and CC. Then we applied community surveys of different frequency, with surveys occurring every **2**, **5**, **10**, or **25** steps (S). These frequencies correspond to a replacement of 10%, 25%, 50% and 100% of individuals in the community, and to trajectories of 25, 10, 5 and 2 segments, respectively.

Section S2 Trajectories in the space of community resemblance

We used the **percentage difference** (*alias* Bray-Curtis dissimilarity) (Bray and Curtis 1957) as index to measure dissimilarity between community states of each dataset (i.e. metric d used to define the space Ω). Figures S2, S3 and S4 below show (for different values of CC and S, and for the three types of community dynamics, respectively) ordination plots resulting from the application of Principal Coordinates Analysis (PCoA) on the resulting dissimilarity matrices (including corrections for negative eigenvalues). Trajectory segments are shown with arrows to display trajectories. Community trajectories corresponding to stabilizing selection appear random as expected (Fig. S2), although when the number of segments is small some trajectories may appear to temporarily follow particular directions:
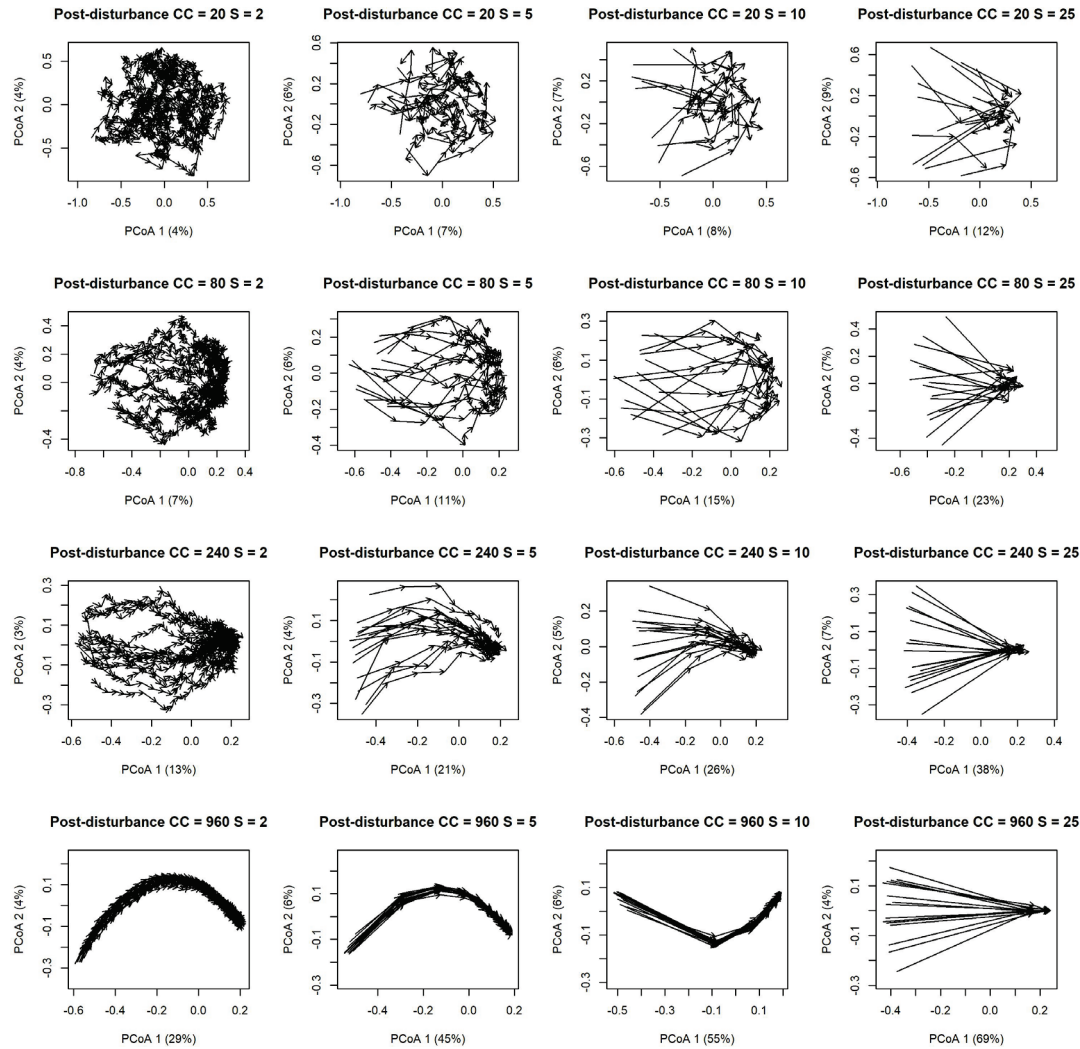
Fig. S2 Examples of simulated dynamics in the 'stabilizing selection' scenario and for different values of carrying capacity (CC, in rows). For each simulation, different trajectories are displayed corresponding to different number of steps (S) between surveys.

Ordination plots of datasets corresponding to post-disturbance recovery dynamics clearly show the return to the pre-disturbance state (Fig S3). If the carrying capacity is low (CC = 25) and the surveys are frequent (S=5) the ecological drift is also apparent. For large values of the carrying capacity (CC = 960), the stochastic effect in the initial community composition is small and trajectories are very similar.

Fig. S3 Examples of simulated dynamics in 'post-disturbance recovery' scenario and for different values of carrying capacity (CC, in rows). For each simulation, different trajectories are displayed corresponding to different number of steps (S) between surveys.

In the case of directional selection, trajectories depart in different directions of the ordination diagram (Fig. S4), as expected (colors are used to group pairs of communities simulated with the same target community state). Still, when carrying capacity (CC) is low and survey frequency is high trajectories contain some degree of ecological drift. However, an increase in CC leads to trajectories straightly departing different directions.

Fig. S4 Examples of simulated dynamics in the 'directional selection' scenario and for different values of carrying capacity (CC, in rows). For each simulation, different trajectories are displayed corresponding to different number of steps (S) between surveys.
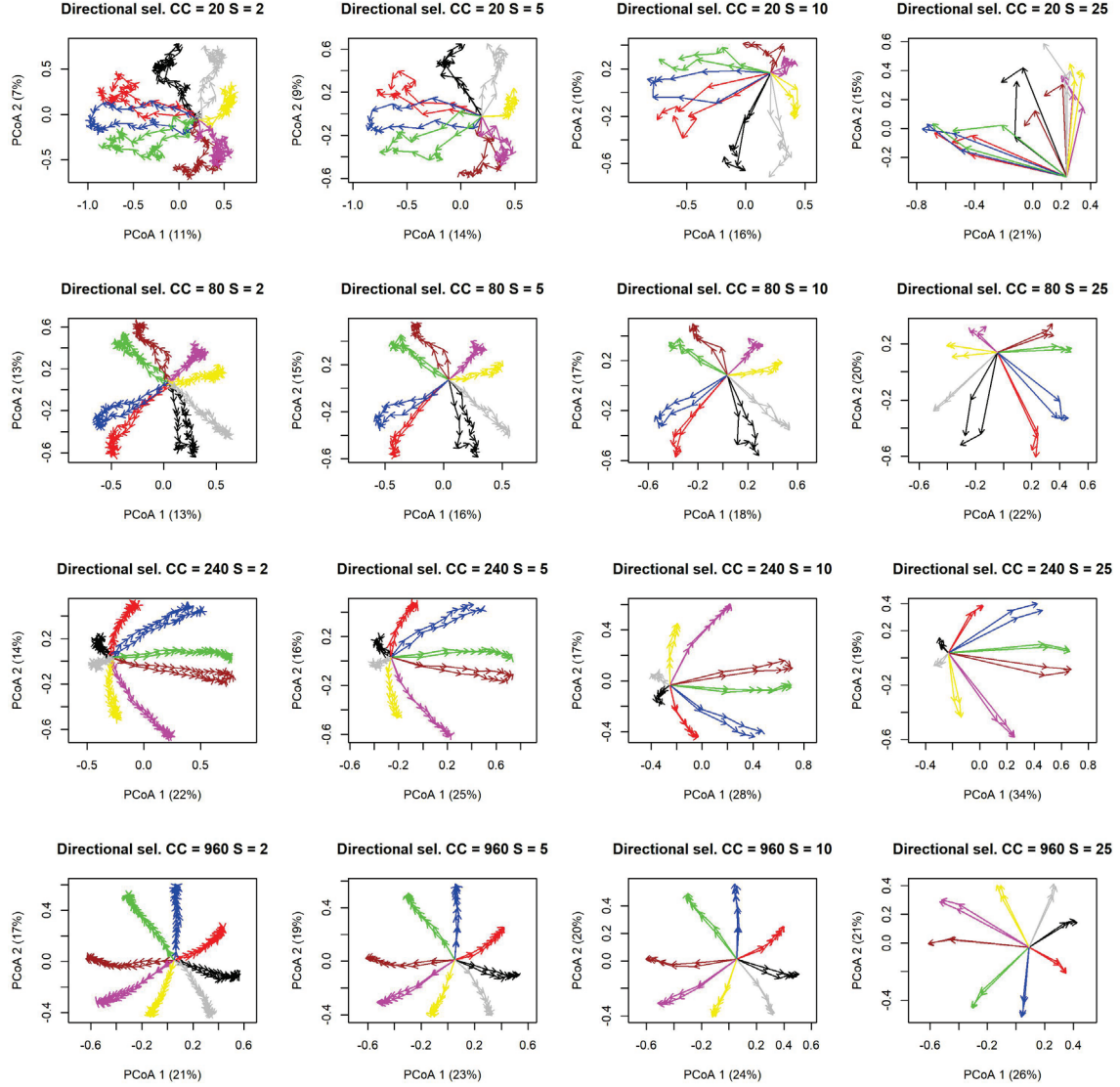

Section S3. Trajectory length

We calculated total path lengths of trajectories using eq. 1 of the main manuscript. Figure S5 below shows, for each type of community dynamics, the average lengths across the sixteen trajectories in a dataset. Since all simulations contained some degree of ecological drift, either increasing the carrying capacity or decreasing the frequency of surveys lead to a decrease in trajectory length regardless of the type of dynamics. Trajectory lengths in

post-disturbance recovery or directional selection dynamics were often similar in length and speed to those of stabilizing selection.

Fig. S5 Trajectory lengths for different types of community dynamics: **Sta** – Stabilizing selection; **Dst** – Post-disturbance recovery; **Dir** – Directional selection. Boxplots show distribution across replicates.
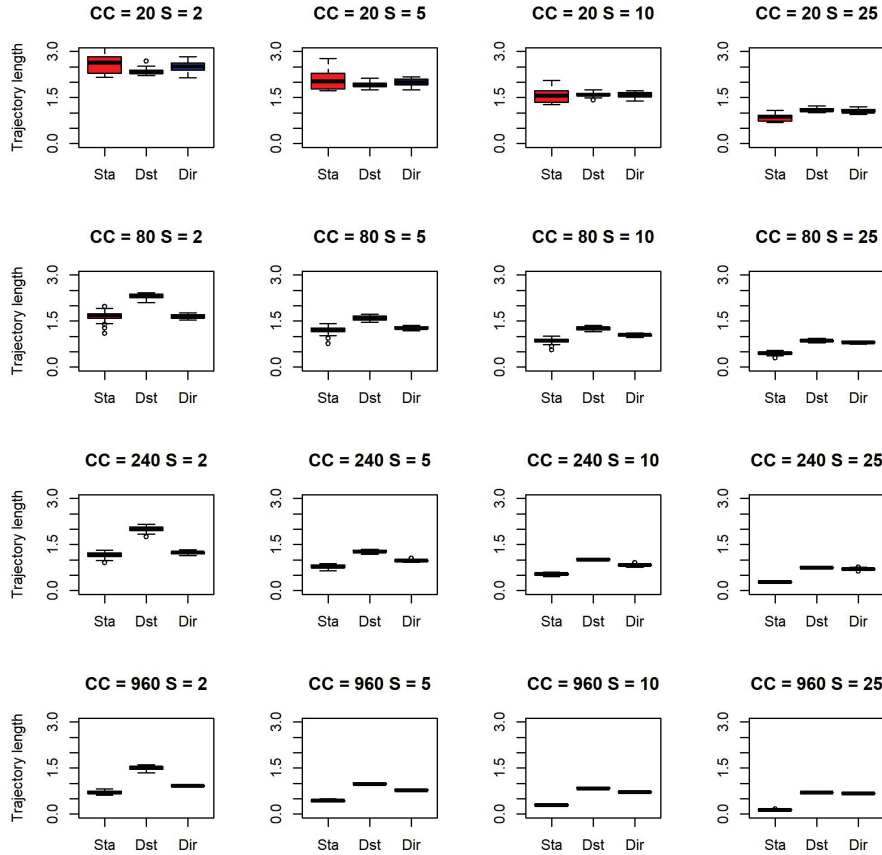
Section S4 Trajectory direction

Trajectory angles averaged across pairs of consecutive segments were relatively small for stabilizing selection if carrying capacity was small and surveys were frequent (CC = 20, S =2; Fig. S6), they increased when increasing carrying capacity or when decreasing the frequency of surveys. The average angles for either disturbance recovery or directional selection were smaller than for stabilizing selection. However, differences were less apparent for small values of CC and frequent surveys.

Fig. S6 Average trajectory angles for different type of community dynamics: **Sta** – Stabilizing selection; **Dst** - Post-disturbance recovery; **Dir** – Directional selection. Boxplots show distribution across replicates.
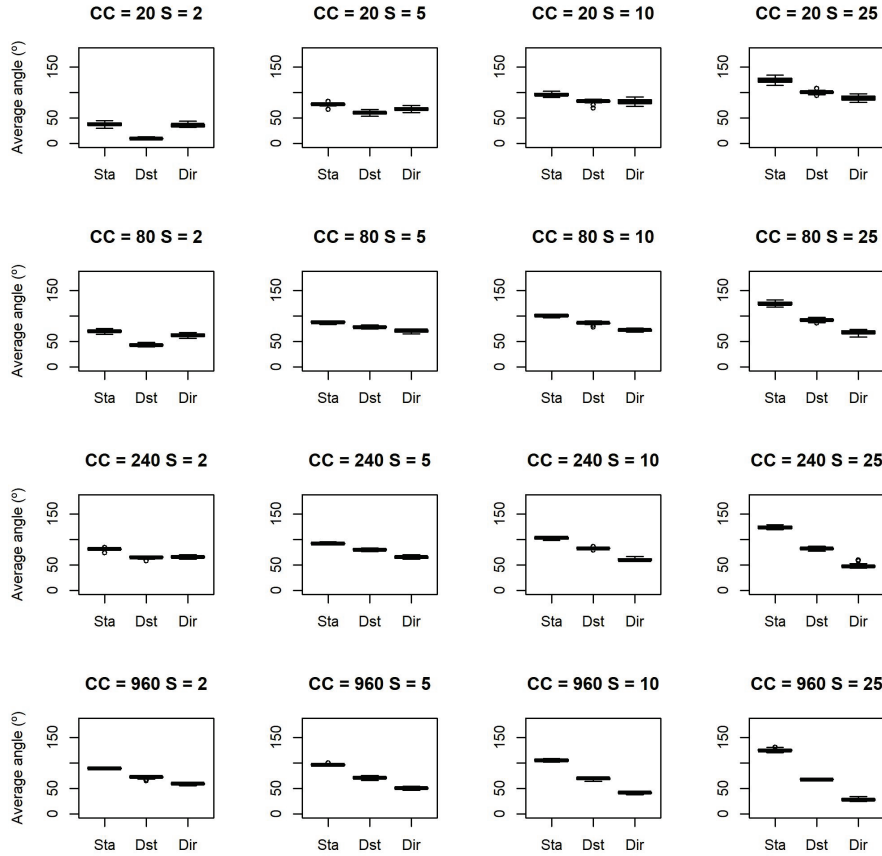
Non-directional and directional dynamics were easier to tease apart when calculating the directionality index (DIR; eq. 3 of the main text). Trajectories under stabilizing selection obtained DIR values around 0.4 for small populations (CC=20) and highly frequent surveys (S=2). Directionality of stabilizing selection trajectories slightly decreased with the increase of CC and S. Values of DIR for trajectories derived from disturbance recovery or directional selection were always larger than those of stabilizing selection. Furthermore, the difference in DIR values increased with carrying capacity.

Fig. S7 Directionality index (DIR; eq. 3 of the main text) for different type of community dynamics: **Sta** – Stabilizing selection; **Dst** - Post-disturbance recovery; **Dir** – Directional selection. Boxplots show distribution across replicates.
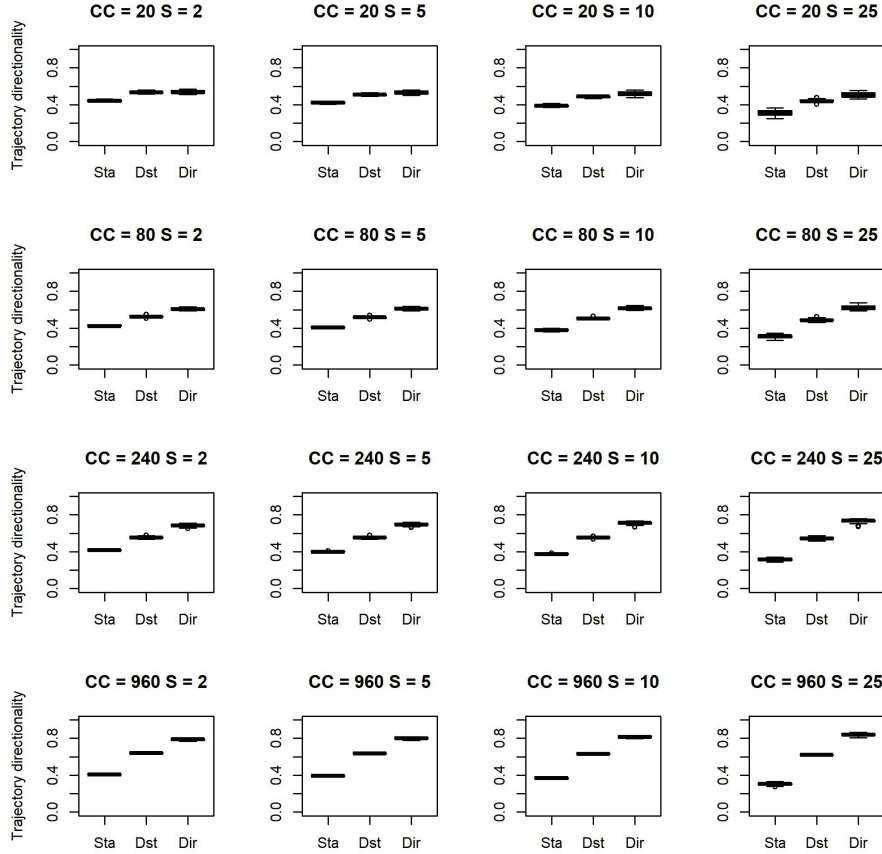
Section S5 Trajectory divergence

We studied trajectory divergence in the simulations of directional selection dynamics. Specifically, we conducted the symmetric convergence/divergence test for pairs of trajectories, distinguishing between pairs that followed the same direction according to the simulation parameters and pairs that did not. Table S1 shows average *p*-values of the Mann-Kendall test in these situations. For trajectories following the same direction, *p*-values were almost always non-significant (left values), whereas the reverse happened for trajectories following different directions (right values). The power of the test in this second case increased with carrying capacity, as a consequence of a milder drift effect. As expected, decreasing the number of surveys had the effect of decreasing statistical power (i.e. increasing average *p*-values) due to the test being performed with smaller sample sizes.

	S = 2	S = 5	S = 10	S = 25
CC = 20	0.221/0.014	0.332/0.032	0.486/0.094	1.0/1.0
CC = 80	0.232/0.000	0.349/0.003	0.496/0.025	1.0/1.0
CC = 240	0.172/0.000	0.282/0.000	0.445/0.013	1.0/1.0
CC = 960	0.217/0.000	0.301/0.000	0.448/0.009	1.0/1.0

Table S1. Average *p*-values of the symmetric convergence/divergence test (i.e. Mann-Kendall test on the sequence of distances between trajectory states) for parallel trajectories (left values) and diverging trajectories (right values) under different combinations of carrying capacity (CC) and number of steps between surveys (S).

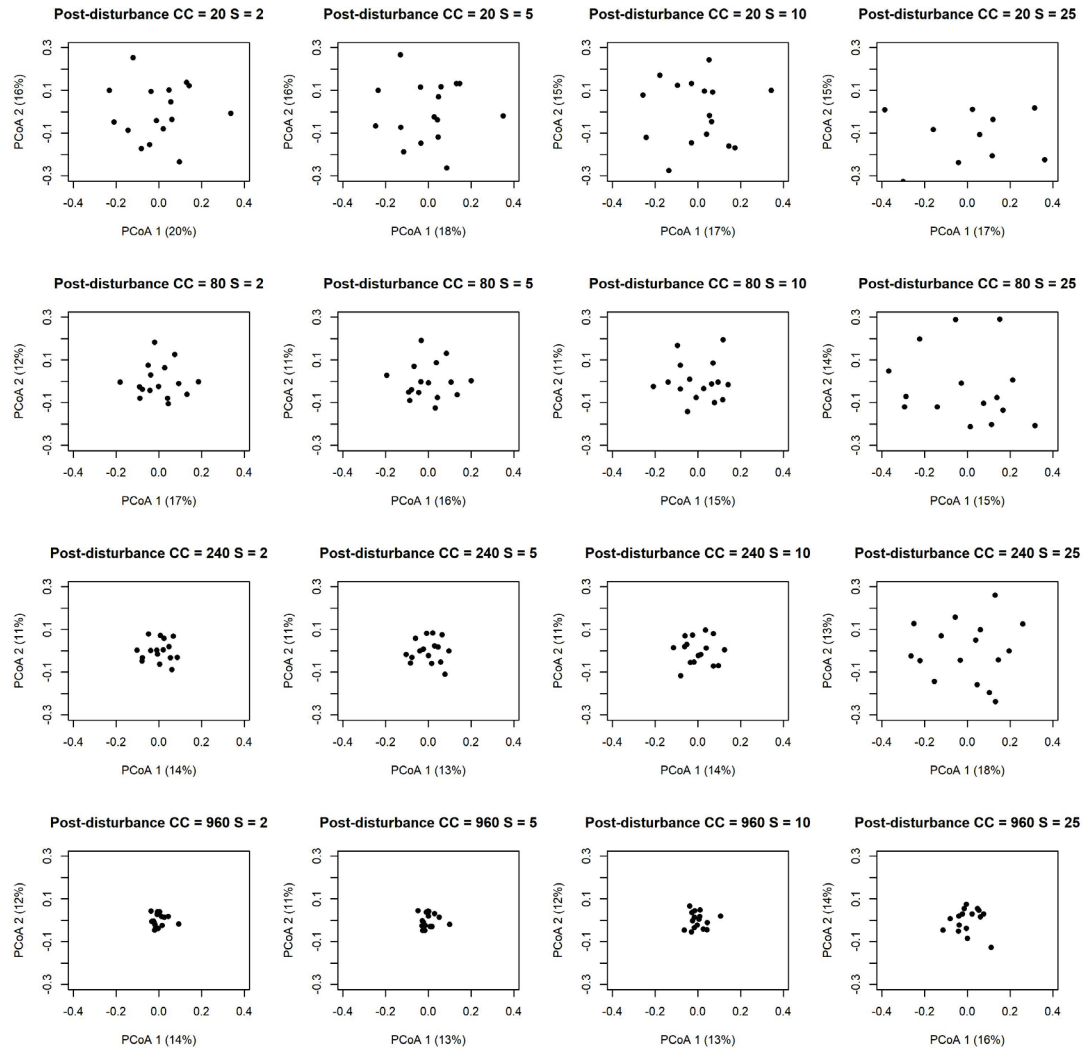*Section S6 Dissimilarity between trajectories*

We used D_{SDSP} (eq. 13 of the main text) to calculate the dissimilarity between the sixteen trajectories of each dataset. Then, we conducted PCoA to display the resulting 16×16 dissimilarity matrix in an ordination space (while correcting for negative eigenvalues). Dissimilarity values between stabilizing selection trajectories were rather small in all cases (Fig. S8). However, increasing the carrying capacity (i.e. decreasing the role of drift) decreased the dissimilarity between trajectories.

Fig. S8 Examples of ordination (PCoA) of dissimilarities between trajectories in the case of stabilizing selection. Plotted dissimilarities correspond to the trajectories shown in Fig. S2.
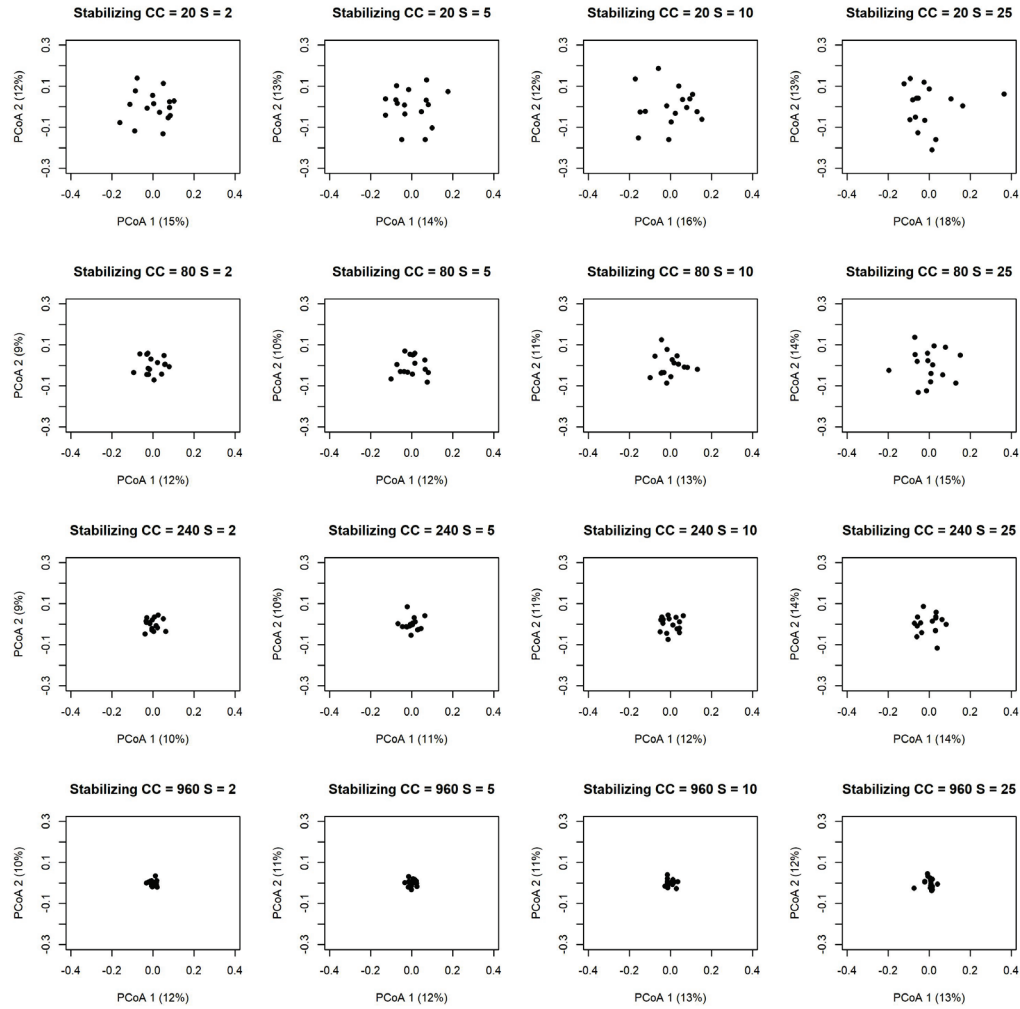
Dissimilarity values between pairs of post-disturbance recovery trajectories (Fig. S9) were slightly larger than between pairs stabilizing selection trajectories. As before, however, the same sampling effects were observed.

Fig. S9 Examples of ordination (PCoA) of dissimilarities between trajectories in the case of post-disturbance recovery. Plotted dissimilarities correspond to the trajectories shown in Fig. S3.

Finally, increasing carrying capacity lead to larger dissimilarity between directional selection trajectories corresponding to different directions (compare Fig. S4 and Fig. S10).

Fig. S10 Ordination diagram (PCoA) of dissimilarities between trajectories in the case of directional dynamics. Plotted dissimilarities correspond to the trajectories shown in Fig. S4.
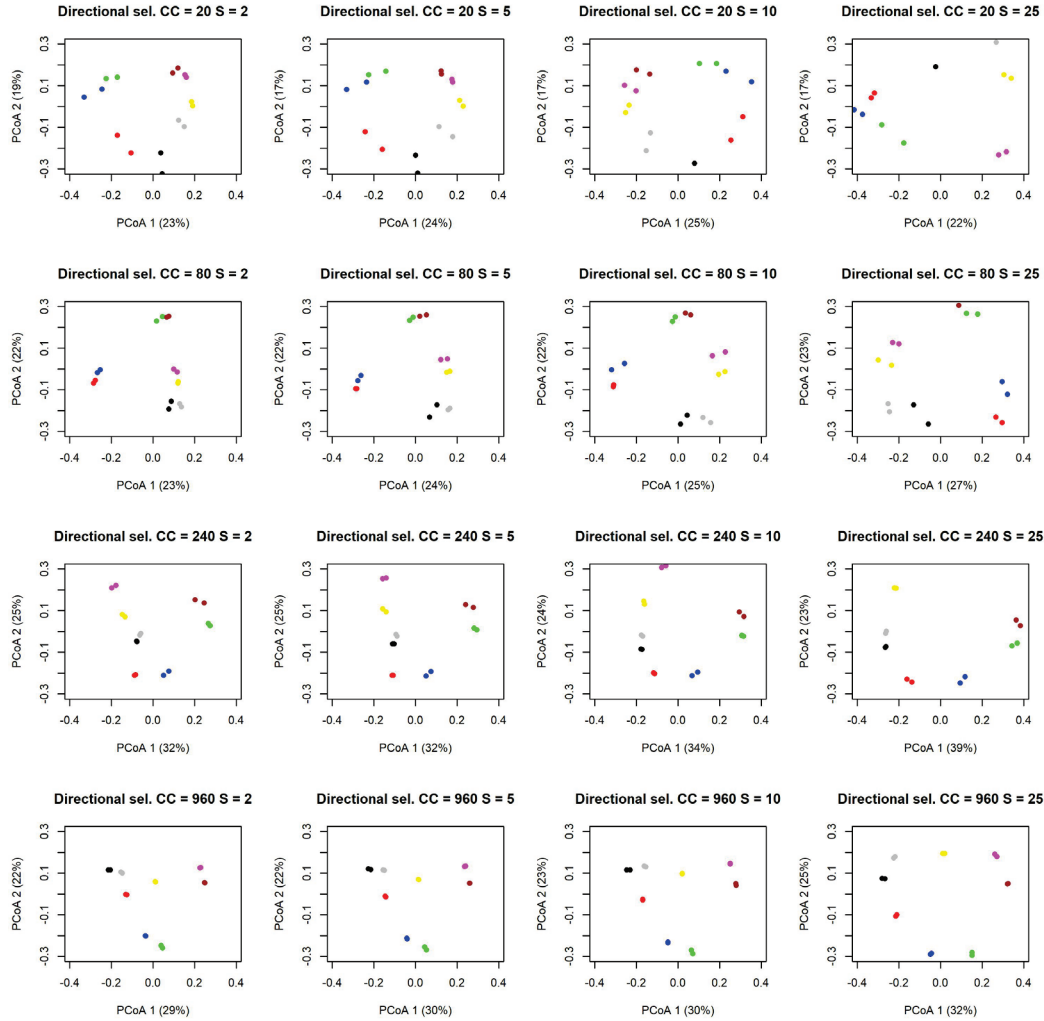
We evaluated to which extent trajectory dissimilarity allowed distinguishing between trajectories following the same direction according to the simulation parameters. Thus, for each simulated data set we conducted k-means clustering on the PCoA space of resemblance between trajectories, looking for 8 clusters. Then, we calculated the adjusted Rand index (Hubert and Arabie 1985) between the known groups and the eight data-driven clusters (Rand values equal to zero indicate random agreement whereas Rand values equal to one indicate perfect agreement). For small values of CC the average value of the Rand index was lower than 1, indicating that drift caused distances to not differentiate completely between directions (Table S2). The adjusted Rand index was also slightly lower when decreasing the frequency of surveys. Nevertheless, for large values of CC trajectory distances allowed a correct classification of the dynamic pathways most of the times, regardless of survey frequency.

	S = 2	S = 5	S = 10	S = 25
CC = 20	0.934	0.938	0.887	0.825
CC = 80	1.000	1.000	1.000	0.967
CC = 240	1.000	1.000	1.000	1.000
CC = 960	1.000	1.000	1.000	1.000

Table S2. Average values of the adjusted Rand index (across simulated datasets) under different combinations of carrying capacity (CC) and number of steps between surveys (S).

Section S7 Overall variation in community dynamics

For each simulated dataset, we used eq. 14 of the main manuscript to calculate variation in community dynamics across the sixteen simulated trajectories (i.e. dBD - dynamic beta diversity). In accordance with Fig. S8, dBD was small for datasets in the case of stabilizing selection, decreased with an increase in carrying capacity (Fig. S11). However, dBD increased slightly when decreasing the frequency of surveys. dBD increased remarkably for either post-disturbance recovery and directional selection dynamics, but the same sampling effects were observed on these cases.

Fig. S11 Dynamic beta diversity for different type of community dynamics: **Sta** – Stabilizing selection; **Dst** - Post-disturbance recovery; **Dir** – Directional *s*election. Boxplots show distribution across replicates.
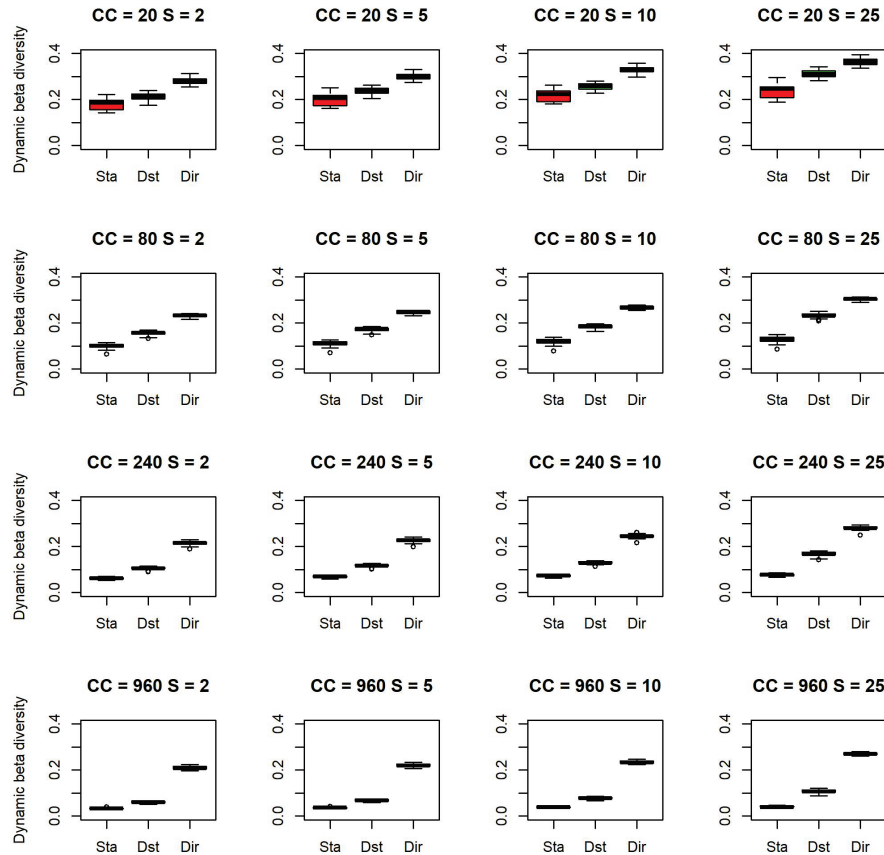
References

Bray, R. J., and J. T. Curtis. 1957. An ordination of the upland forest communities of southern Wisconsin. Ecological Monographs 27:325–349.

Harms, K. E., R. Condit, S. P. Hubbell, and R. B. Foster. 2001. Habitat associations of trees and shrubs in a 50-ha neotropical forest plot. Journal of Ecology 89:947–959.

Hubert, L., and P. Arabie. 1985. Comparing partitions. Journal of Classification 2:193–218.

Simpson, G. L. 2018. coenocliner: A coenocline simulation package for R (ver. 0.2-2).

Vellend, M. 2016. The theory of ecological communities. Princeton University Press.