Assessing the relative importance of neutral stochasticity in ecological communities


M. Vellend (mark.vellend@usherbrooke.ca), C. D. Brown and I. H. Myers-Smith, Dépt de Biologie, Univ. de Sherbrooke, Sherbrooke, QC, J1K 2R1, Canada. – D. S. Srivastava, K. M. Anderson, J. E. Jankowski, E. J. Kleyhans, N. J. B. Kraft, A. D. Letaw, A. A. M. Macdonald, J. E. Maclean, Dept of Zoology and Biodiversity Research Centre, Univ. of British Columbia, Vancouver, BC, V6T 1Z4, Canada. – A. R. Norris, Dept of Forest Sciences, Univ. of British Columbia, Vancouver, BC, V6T 1Z4, Canada. – X. Xue, Dept of Botany and Biodiversity Research Centre, Univ. of British Columbia, Vancouver, BC, V6T 1Z4, Canada.

A central current debate in community ecology concerns the relative importance of deterministic versus stochastic processes underlying community structure. However, the concept of stochasticity presents several profound philosophical, theoretical and empirical challenges, which we address here. The philosophical argument that nothing in nature is truly stochastic can be met with the following operational concept of neutral stochasticity in community ecology: change in the composition of a community (i.e. community dynamics) is neutrally stochastic to the degree that individual demographic events – birth, death, immigration, emigration – which cause such changes occur at random with respect to species identities. Empirical methods for identifying the stochastic component of community dynamics or structure include null models and multivariate statistics on observational species-by-site data (with or without environmental or trait data), and experimental manipulations of ‘stochastic’ species colonization order or relative densities and frequencies of competing species. We identify the fundamental limitations of each method with respect to its ability to allow inferences about stochastic community processes. Critical future needs include greater precision in articulating the link between results and ecological inferences, a comprehensive theoretical assessment of the interpretation of statistical analyses of observational data, and experiments focusing on community size and on natural variation in species colonization order.
the goal is to place real communities on a continuum between the two extremes (Gewin 2006). General knowledge may then be advanced by understanding the factors (e.g., scale, habitat) that determine the position of natural communities on this continuum (Chase and Myers 2011). However, Clark and colleagues (Clark et al. 2007, Clark 2009) threw down the gauntlet on the continuum consensus, arguing that stochasticity is an attribute only of models and not of nature: “contrary to the emerging consensus, while models do indeed represent a continuum, there is no evidence for such a continuum in the underlying causes. Moreover, the continuum in models is one of knowledge, not cause” (Clark et al. 2007). This is an unsettling possibility for ecologists employing the continuum approach as it implies that posing the question “what is the relative importance of stochastic processes in this community?” is a logical fallacy.

Reconciling these perspectives is challenged by the many meanings of the term stochasticity in the literature. For example, one of the most common forms of stochasticity in the ecological literature – environmental stochasticity – involves unpredictable environmental fluctuations (including disturbance events), and is thus modeled as stochastic (Morris and Doak 2002), but the effect of environmental fluctuations on biotic outcomes might be entirely deterministic. Here we do not consider environmental stochasticity, but rather we focus on recent studies that have used the term ‘stochasticity’ to mean neutral stochasticity, which we define more precisely in the next section. Empirical studies have inferred the action of neutrally stochastic processes in several ways: when variation in community or trait composition is no different than patterns generated by a null model (Gotelli and Graves 1996, Ellwood et al. 2009, Chase and Myers 2011), when this variation is explained by spatial rather than environmental variables (Cottenie 2005), or when species colonization order has an impact on composition (Fukami 2004, Chase 2010). Synthesizing these types of studies to provide broader insights into ecological communities will require a clearer understanding of how the concept of stochasticity is operationalized in different studies.

Here we present an analysis of the concept of neutral stochasticity in community ecology on philosophical, theoretical, and empirical grounds. We first tackle philosophical issues related to stochasticity, and provide an operational definition that reconciles opposing viewpoints. Next, we conceptually relate the fundamental processes underlying community dynamics to the stochastic–deterministic dichotomy. Finally, we systematically compare empirical approaches to characterizing the neutrally stochastic component of community structure or dynamics, and outline a program of future research to maximize progress on this topic.

**Philosophical issues**

By placing stochasticity at the center of a key debate, community ecology is grappling with an issue that has occupied philosophers for centuries: does anything in nature really happen by chance? (Cohen 1976, Gigerenzer et al. 1989). According to the dictionary definition, a process is considered stochastic if outcomes have “a random probability distribution or pattern that may be analysed statistically but may not be predicted precisely” (Soanes and Stevenson 2008). Clark’s (2009) point of view that stochastic model terms represent only human ignorance of underlying processes can be traced back to at least the 18th century (Gigerenzer et al. 1989). Oft-cited examples include coin flipping and dice tossing, which we model as stochastic, despite deterministic physical processes that decide the outcome. An analogous ecological example is the survival (or death) of a single organism over a specified time frame. Knowledge of the organism’s traits and the environment in which it lives help improve our estimate of the survival probability (Clark 2009), but most often the estimate is not zero or one, and so the process is considered (partly) stochastic.

Of the various branches of science that have grappled with the concept of stochasticity, ecology’s sister discipline – evolutionary biology – has the most direct lessons to offer ecologists (Lenormand et al. 2009). Community ecology is closely analogous to population genetics, in that both fields are primarily concerned with understanding why different numbers and types of biological variants – species and alleles, respectively – are found at different places and times (Vellend 2010). Mutation and genetic drift are widely accepted as stochastic processes contributing to evolutionary change (Lenormand et al. 2009). The fact that each mutation may have an underlying deterministic cause (e.g., the action of a chemical mutagen) does not alter the view of mutation as stochastic, given that mutation involves what Sewall Wright (1964) referred to as “practically irreducible probabilities like those in the fall of dice” (see also Cohen 1976). A similar argument can be made concerning drift, which occurs in small populations when individual survival and reproduction are random with respect to allelic states at particular loci.

It is our view that Clark’s (2009) argument that “stochastic elements stand in for unknown processes” is technically valid (except perhaps at the level of quantum mechanics), but that there is a probabilistic component to ecological dynamics that is, in the words of Wright (1964), “practically irreducible”. Although the operational distinction between deterministic and stochastic can appear arbitrary (Denny and Gaines 2000), it can be made explicit, and therefore potentially comparable among studies. This is done by specifying, for outcome Y, that it is stochastic (or random) with respect to factor X. McShea and Brandon (2010) offer the analogy of two people flipping coins. Even if each coin flip has, technically, a deterministic outcome, each sequence of heads and tails is entirely random with respect to the other. Similarly, each death or reproductive episode in a population might have deterministic causes, but these events may be completely random with respect to the alleles at a particular locus; genetic drift is the result, and is usefully considered stochastic. In a community context, if demographic events occur randomly with respect to species identities, stochastic community drift is the result. McShea and Brandon (2010) refer to this as stochasticity in the “with-respect-to sense”, and we argue that this conception of the term stochasticity may have great utility in ecology.

We offer the following conception of neutral stochasticity in community ecology: change in the composition of a community (i.e. community dynamics) is neutrally stochastic to the degree that individual demographic events – birth, death, immigration, emigration – which cause such changes occur at random with respect to species identities. In practical
terms, if we consider a vector of species abundances, A, to be the object of prediction, then community dynamics can be considered stochastic to the extent that predicting changes in A requires terms that are probabilistic with respect to species identities. Thus, it is possible to articulate an operational and useful definition of neutrally stochastic processes underlying community dynamics, even if the events of interest (birth, death, etc.) have deterministic underpinnings with respect to processes outside the scope of ecology (e.g. the deterministic path of a hailstone that kills a seedling).

**Fundamental processes in community ecology**

**Drift, selection and speciation**

Four fundamental processes underlie community dynamics: drift, selection, speciation and dispersal (Vellend 2010). Drift – random changes in local species relative abundances – happens when birth and death events in a community occur at random with respect to species identity (Hubbell 2001). As such, community drift is unambiguously neutrally stochastic; indeed ‘demographic stochasticity’ is often considered a synonym for drift. Ecological selection involves deterministic fitness differences between individuals of different species, and as such, is unambiguously not stochastic. Even if fluctuations in selection, including disturbance events, are unpredictable (typically modeled as ‘environmental stochasticity’), we consider the resulting community dynamics to be deterministic, as they depend on species differences in demographic responses to environment. Processes lumped under the umbrella of ‘niche’ also fall under the umbrella of selection, although in some influential frameworks (Chesson 2000), niche differences include only those that stabilize community dynamics, and not those involving fitness differences leading to competitive exclusion, which also represent a form of selection.

Over large spatial and temporal scales, speciation is a key factor shaping regional species pools, which, in turn, may influence community patterns over shorter spatial and temporal scales. For example, whether species arise via “ecological speciation” (sensu Rundle and Nosil 2005) or via other modes of speciation (e.g. drift between isolated gene pools or sexual selection) can have an important impact on local patterns of species diversity (McPeek 2007). However, while the role of stochastic and deterministic factors underpinning speciation is an important topic within evolutionary biology (Rundle and Nosil 2005), we consider the process of speciation as largely outside the scope of the present paper. Studies conducted in the context of the stochasticity–determinism debate within community ecology typically take as a given a regional pool of species, and the central question concerns how local communities are assembled from that pool. We do not consider speciation any further until the conclusions of this paper.

**Dispersal**

Dispersal presents by far the most ambiguities concerning inferences about stochastic versus deterministic underpinnings of community dynamics (Lowe and McPeek 2014). The movement of an individual or propagule from one place to another is a dispersal event, and as such dispersal is no more stochastic or deterministic, in and of itself, than other events like birth or death. For simplicity, we consider the magnitude of dispersal as the quantity of individuals moving between habitat patches, setting aside the issue of dispersal distance. Even if all individuals in a community, regardless of species, have the same probability of dispersal (i.e. dispersal itself is stochastic with respect to species identity), there need not be any influence of this stochastic dispersal on community dynamics (a key part of our definition of stochasticity in community ecology). This is because, unlike for birth and death, the movement of an organism is not necessarily related to its fitness (although clearly it can be). For the same reason, interspecific differences in dispersal do not necessarily equate with selection. However, dispersal can strongly influence the importance of local drift and selection, either opposing or reinforcing the changes caused by these processes (Vellend 2010).

A useful heuristic for evaluating the community-level consequences of dispersal is to distinguish between dispersal into ‘empty’ habitats (i.e. colonization) and dispersal between communities where species have established and local selection has already had time to operate (‘established’ communities). For both cases, we can further identify two key components of dispersal: 1) the mean across all species and 2) the variance among species in the rate of movement to a new habitat or between established habitat patches. Many metacommunity models address the consequences of variation in 1) but not of 2) (Lowe and McPeek 2014); here we present some predictions based on first principles and analogous models in population genetics (Hartl and Clark 1997). Low mean dispersal means that few propagules of any species move between sites, while low variance indicates that all species have similar dispersal.

If A is the vector of species abundances, the case of colonization (Fig. 1A) represents a special case in which the initial state of A is null (i.e. all zeros). In this case, a low mean dispersal rate (relative to local population growth rates) and low variance among species can lead to high stochastic variation (i.e. unpredictability) in initial species composition among sites (Chase 2003, Fukami 2010). This phenomenon is akin to founder effects in population genetics, whereby multiple founding populations that are small (analogous to low dispersal) and that represent random samples from a regional population (analogous to low variance with respect to allelic identity) will show large genetic differences among them (Hartl and Clark 1997). High mean dispersal indicates that many propagules arrive per unit time, and high variance means that there are many more propagules of some species than others – that is, the dispersal process itself involves selection favoring certain species. This high–high combination should lead to high predictability in initial species composition (i.e. low stochasticity). Combinations of low mean and high variance, or high mean and low variance, are expected to lead to intermediate degrees of the predictability of initial species composition.

For the case of dispersal between established communities, we present one example in which we additionally assume contrasting environments with divergent selection pressures (Fig. 1B). In this case, high mean dispersal can counter both local drift and selection. With variance among species in dispersal, the effects in established communities depend on the relationship between species’ dispersal ability and local
fitness. With a positive correlation, dispersal effectively reinforces local selection. A negative correlation directly counters local selection and may create frequency-dependent selection at the level of the metacommunity via what is commonly referred to as a competition-colonization tradeoff (Tilman 1994; Fig. 1B).

A huge number of alternative situations can be envisioned, but these deliberately simplified dispersal scenarios reveal some important general lessons. First, colonization at a low rate that is random with respect to species identity can stochastically create differences in composition among otherwise identical communities (Chase 2003). As described above, these are essentially ‘founder effects’, which have the same effect as bottlenecks (i.e. a period of low population/community size, which increases drift) (Hartl and Clark 1997). Dispersal among established communities can strongly influence the local importance of stochastic (drift) or deterministic (selection) processes (Vellend 2010). An interesting corollary is that the influence of dispersal depends on the local environment, such that while the distance between sites might be a good predictor of exchange via dispersal, it might be a poor reflection of the ultimate consequences of dispersal (i.e. the patterns one finds in observational data). As such, it is difficult, if not impossible, to unambiguously classify the consequences of dispersal as stochastic or deterministic in the same sense as drift or selection. Variance in dispersal ability among species may contribute to selection – depending on additional details such as whether there is a tradeoff between dispersal ability and other traits – but it need not. In a homogeneous environment, for example, dispersal may have no effect on fitness and therefore no selective consequence. Likewise, an absence of interspecific dispersal variance does not necessarily make community dynamics any more or less stochastic. For example, low mean dispersal (often described as ‘dispersal limitation’) may permit both local selection and local drift to have a greater influence on community dynamics (Fig. 1B, low variance and low mean dispersal). Finally, the consequences of dispersal may depend strongly on the spatial scale of observation. In a local patch, a high rate of incoming dispersal may counter selection favoring a particular species, whereas at a regional level, a tradeoff between dispersal ability and competitive ability maintains coexistence deterministically via frequency-dependent selection (Tilman 1994). In sum, while dispersal is clearly a key process that can determine community dynamics, it does not appear useful to classify its effects as stochastic or deterministic per se.

An example: historical contingency

These four fundamental processes of community ecology – drift, selection, speciation and dispersal – interact to determine the degree of stochasticity in community dynamics, and therefore the predictability of community composition. We elaborate on this point by exploring the concept of historical contingency, which refers to the importance of previous or initial species abundances in determining the dynamics or structure of current ecological communities (Fukami 2010). Some degree of correlation between past and present abundance (i.e. a form of historical contingency) is necessary given that constrained rates of mortality and reproduction prevent immediate attainment of equilibria. While such inertia is
theoretically trivial, it can have important consequences for the communities we observe in the field (Körner et al. 2008) and transient states can persist for very long periods of time (Fukami and Nakajima 2011, Olszewski 2011). As such, research in this area has focused in large part on situations in which initial species abundances alter the selective regime (Chase 2003), or in which small community size leads to an important role for drift (Fukami 2004). ‘Priority effects’ occur when the order that species colonize habitats affects final community composition (reviewed by Fukami 2010).

One consequence of historical contingency is that identical replicates of a habitat (in terms of environment, spatial context, etc.) may contain different species assemblages, and the observation of high compositional variation (i.e. beta diversity) in the absence of environmental heterogeneity is often interpreted as evidence of stochastic processes (Chase 2003, 2007, 2010, Fukami 2004, 2010, Chase et al. 2009). However, this pattern can arise from several different combinations of stochastic and deterministic processes (as noted in many of these papers). If we consider colonization of new habitats, initial similarity in species composition across replicate patches depends on dispersal parameters (Fig. 1A). Assuming some degree of stochastic variation in A generated by dispersal, subsequent selection and drift may reduce or amplify this variation. Constant selection in the form of an environmental stress can erase initial differences in A, causing convergence across patches. Priority effects, or any situation that favors multiple stable equilibria, effectively represent divergent selection that can amplify initial differences in A. Drift can also increase beta diversity over time, but likely at a slower rate (Clark and McLachlan 2003). Therefore, persistent variation in species composition among environmentally similar patches may not be due only to stochastic colonization, but also to a mixture of stochastic and deterministic processes that act on this initial variation (Chase 2003).

**Empirical approaches to inferring stochastic community dynamics**

Many studies attempt to quantify the relative importance of stochasticity in community dynamics either in a single (meta)community, or by comparing local communities with different properties (e.g. disturbance, productivity). Evaluating different approaches in such studies is complicated by inconsistent terminology used to describe very similar ideas. Many studies are described as testing for the presence of neutral processes in communities, without explicitly using the term stochastic(ity). But since dynamics in Hubbell’s (2001) neutral model are driven entirely by stochastic processes, we consider such studies to address the issue of stochasticity. However, in our analysis of methods, we do not include studies that only compare observed versus predicted patterns of species’ abundances, as these are very weak tests of neutral theory (Clark 2012), nor do we include studies focused on rejecting neutral theory.

| Table 1. An evaluation of methods used to characterize stochastic components of community structure and dynamics. |

<table>
<thead>
<tr>
<th>Method</th>
<th>Result with potential relevance to stochasticity</th>
<th>Inference</th>
<th>Key assumptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Null models</td>
<td>Pattern in species × site matrix (possibly with trait or phylogenetic information) not significantly different from null model</td>
<td>Species distributions stochastic with respect to one another</td>
<td>Null model constraints are not themselves influenced by the observed covariances in species distributions</td>
</tr>
<tr>
<td>Variance partitioning</td>
<td>Variance in community composition explained by spatial variables, independent of environmental variables</td>
<td>Community composition spatially structured, stochastic with respect to environment</td>
<td>All relevant environmental variables have been measured. Spatial patterns of community composition not created by biotic interactions.</td>
</tr>
<tr>
<td>Community assembly via trait selection analysis</td>
<td>Local species abundances resemble regional abundances, independent of traits</td>
<td>Community composition is determined by dispersal, stochastic with respect to traits</td>
<td>Per capita dispersal into local sites is equal across species. Traits combine linearly to influence abundance.</td>
</tr>
<tr>
<td>Dissimilarity at zero distance</td>
<td>Local species abundances not predicted by traits or regional abundances</td>
<td>Community composition stochastic with respect to traits and dispersal</td>
<td>Compositional differences due to drift among distant communities not considered stochastic.</td>
</tr>
<tr>
<td>Priority effects experiments</td>
<td>Colonization order (initial species composition) influences final species composition</td>
<td>Colonization order – stochastic with respect to species identity – can create variation among sites in equilibrium community composition</td>
<td>In applying inference to nature: colonization order varies stochastically in nature as in experiment</td>
</tr>
<tr>
<td>Species density/ frequency experiments</td>
<td>Intraspecific density dependence = interspecific density dependence</td>
<td>Demography stochastic with respect to species identity</td>
<td>Failure to reject null hypothesis (of no difference) is support for null hypothesis</td>
</tr>
</tbody>
</table>
(Adler et al. 2006, Dornelas et al. 2006, Harpole and Suding 2007) without also attempting to quantify or characterize a stochastic component of community dynamics or structure. We focus on neutral stochasticity, meaning stochastic with respect to species identity, and so do not consider studies that examine the importance of demographic stochasticity in predicting the abundances of each species without assessing whether this differs between species (Martorell and Freckleton 2014). Since the number of relevant studies is immense, here we describe broad categories of observational and experimental approaches (Table 1), rather than attempting a comprehensive review. For each approach, we describe the ’with respect to’ sense in which stochasticity is assessed (Table 1), but we do not specify this when describing conclusions from particular studies, given that authors have typically not made this specification.

**Observational approaches: null models**

The simplest null model approach most often begins with presence–absence data for a given set of species across multiple sites, and asks whether there are non-random patterns in the data with reference to some alternative way the data might be structured – the null model (Table 1). One can also incorporate data on the traits or phylogenetic relationships among species, again asking whether the pattern of trait or phylogenetic differences among locally co-occurring species differs from the patterns produced by a null model (Cavender-Bares et al. 2009, Cornwell and Ackerly 2009, Bernard-Verdier et al. 2012). As described by Gotelli and Graves (1996, p. 7), “Null models emphasize the potential importance of stochastic mechanisms in producing natural patterns”. The optimistic view is that one can specify a null model that takes the original data and then removes the influence of a particular process of interest, most commonly interspecific competition. If observed patterns of species co-occurrence or trait structure are not significantly different than those found in a large number of realizations of the null model, one concludes that community dynamics are largely stochastic with respect to the excluded process (Chase and Myers 2011). Community patterns most often deviate from those expected under null models, although the magnitude of such deviations varies greatly (Gotelli and McCabe 2002). The controversy surrounding null models, and indeed all observational approaches, stems from the great difficulty in inferring process from pattern (Weiher and Keddy 2001). Many different processes (e.g. competition, differences in species responses to the environment) can create non-random patterns, although the specific deterministic processes at play are not necessarily relevant to identifying a stochastic component to community structure. More importantly, the influence of an important ecological process might be contained in the very constraints included in a null model (Harvey et al. 1983). For example, null models typically take as a given (i.e. hold constant) some property of the regional community such as total abundances or the shape of a trait distribution (Gotelli and Graves 1996), although such properties may well themselves result from deterministic processes (Harvey et al. 1983). As a result, while null models have great utility in characterizing a pattern of interest (e.g. species richness) while controlling for some other pattern of interest (e.g. sampling intensity), we consider failure to reject a null model as fairly weak evidence in the specific case of stochastic community dynamics.

**Observational approaches: multivariate analysis**

Incorporating environmental variation directly into community analyses is typically done using multivariate statistics (Legendre and Gauthier 2014). Using the “raw data” approach (Legendre et al. 2005), a vector $A$ of species abundances (or incidences) in a given site represents a multivariate response to potential predictor variables. An ordination-based analysis (e.g. canonical correspondence analysis) can then partition site-to-site variation in community composition into components explained by environmental variables (typically representing several important gradients structuring species distributions), spatial variables (based on the locations of sites), environment and space jointly (because environmental conditions are frequently spatially autocorrelated), and unexplained variation (Legendre et al. 2005; Table 1). The ‘distance approach’ first calculates an index of community dissimilarity between each pair of sites as a response, which may then be predicted by pairwise geographic distances or environmental dissimilarities (Tuomisto and Ruokolainen 2006). Although there are some important technical distinctions between these two types of analysis (Tuomisto and Ruokolainen 2006, Gilbert and Bennett 2010, Anderson et al. 2011), they are frequently aimed at testing the same qualitative hypotheses. Legendre et al. (2005) list three main hypotheses: community composition is 1) uniform across sites, 2) spatially autocorrelated but otherwise random with respect to the environment, or 3) related to environmental variables. Legendre et al. (2005) equate hypothesis 2 directly with neutral theory, and therefore with stochastic community dynamics (see also Cottenie 2005).

Using these approaches, a statistical effect of ‘space’ on community composition, independent of environmental variables, has been interpreted as support for neutral theory (Gilbert and Lechowicz 2004, Cottenie 2005), sometimes with specific reference to ‘stochastic’ processes (Barber and Marquis 2010). In a meta-analysis using the raw data approach, Cottenie (2005) found that environmental variables were most important in explaining variation in species composition, but that spatial variables frequently explained significant variance. The conclusion was that “disregarding neutral dispersal processes would result in missing important patterns in 37% of the studied communities” (Cottenie 2005). This type of interpretation is fraught with difficulty, because of the way ‘space’ is represented, and because a spatial effect is open to alternative interpretations (Jacobson and Peres-Neto 2010, Logue et al. 2011). The distance approach uses Euclidean geographic distances between sites, which should be monotonically related to the amount of dispersal between sites (to a first approximation), but the raw-data approach represents space using functions of each site’s spatial x–y coordinates, most often polynomials (Borcard et al. 1992) or sine waves (Borcard and Legendre 2002, Dray et al. 2006), with no clear theoretical link to dispersal or any other specific process (Jacobson and Peres-Neto 2010, Peres-Neto et al. 2012). Either way, it is possible that the ‘space only’ component of variation is actually driven by...
unmeasured, spatially structured environmental variables, or by environment-independent biotic interactions (Diamond 1975). As such, the inference of stochastic community dynamics based on spatially structured community composition rests on very shaky ground (Gilbert and Bennett 2010, Smith and Lundholm 2010).

The ordination-based approach effectively models species distributions as a function of environment and space, and one can also interpret unexplained variance as representing processes that are stochastic with respect to environment and space. In an example in which each species was modeled independently, Soininen et al. (2013) constructed “niche” models from environmental and biotic information, and assumed that the unexplained portion represented stochasticity. However, this interpretation is open to Clark et al.’s (2007) critique that the unexplained variance is just that – unexplained – rather than unexplainable and therefore not truly stochastic.

**Observational approaches: community trait structure**

Shipley et al. (2012) recently proposed a method of combining data on traits and abundances across plots to quantify the relative roles of selection (trait-based filtering of species), dispersal (ressemblance of local to regional abundances, independent of traits), and demographic stochasticity (unexplained variance in species abundances) (Table 1). The method was applied to tropical tree communities in French Guiana. Using species mean trait values (rather than site-level trait values), the majority of variance was explained by dispersal and demographic stochasticity, whose importance increased and decreased, respectively, with increasing plot size (Shipley et al. 2012). The mathematical details of this approach are beyond the scope of the present paper, but several points can be made concerning underlying processes. On the positive side, distinguishing effects of dispersal and drift (i.e. demographic stochasticity) is an important advance on previous methods. However, the attribution of unexplained variance to stochasticity assumes that the relevant traits have been measured and that the measured traits combine linearly to determine fitness – very difficult assumptions to verify across many species. Finally, the analysis of Shipley et al. (2012) formally incorporates traits, but not plot-level environmental conditions, which is exactly the opposite of the multivariate variance partitioning approach described above, pointing to an obvious avenue for future integration.

**Observational approaches: dissimilarity at zero distance**

The raw-data approach to multivariate community analysis with only spatial distance as a predictor variable is akin to spatial autocorrelation analysis. Using raw distances, the y-intercept can be estimated, and is referred to as the “nugget” (Legendre and Fortin 1989), representing the expected dissimilarity between plots at zero distance apart (i.e. perfect replicates). Brownstein et al. (2012) proposed using the nugget as a direct measure of the stochastic component of community structure. From a statistical point of view, the nugget does appear to represent truly irreducible variance in community composition, and thus provides a promising addition to the toolbox. However, variance in community composition among plots at non-zero distances apart may also be due to stochastic drift, such that the nugget represents a minimum estimate (Table 1). In analyses of empirical data, Brownstein et al. (2012) reported no relationship between this measure of stochasticity and local richness, but a positive correlation with regional richness. Given that correlations between local richness, regional richness and beta diversity can result from statistical inevitabilities (Kraft et al. 2011), the sensitivity of this metric to other attributes of a data set (potentially unrelated to stochastic processes) needs to be evaluated.

**Experimental approaches**

Several studies have imposed stochastic variation in species colonization order (or initial composition) in experimental communities, and asked whether the degree of subsequent differentiation among replicate communities depends on initial site conditions (“Priority effects experiments” in Table 1). If so, it is inferred that the stochastic process related to the order of site colonization is most important under conditions in which community differentiation was greatest. Using this approach, differentiation among freshwater communities (producers and small animals) was found to increase with greater productivity (Chase 2010) and in the absence of drought (Chase 2007). Although the results of these studies were couched in terms of the relative importance of stochastic processes, interpretations clearly implicated deterministic processes (frequency-dependent selection) in amplifying initial community differences. Fukami (2010) reviewed similar types of lab, field, and theoretical studies, concluding that community dynamics are more stochastic in smaller, more isolated, and less heterogeneous communities (Fukami 2010). These results are consistent with a purely stochastic explanation, closely paralleling the result from theoretical population genetics that drift is most important in small, isolated populations in homogeneous environments (Hartl and Clark 1997). These studies typically test the outcome of manipulated variance in colonization order, rather than actual variance observed in the field.

A different kind of experiment aims to directly test the assumption of neutral theory that individuals of different species are equivalent in a particular setting (“Species density/frequency experiments” in Table 1). If individuals of two species compete for resources, but there is no selective advantage of one species over the other, regardless of their relative frequencies, community change can only happen via drift (ignoring dispersal and speciation). Siepielski et al. (2010) created mixtures of two species of larval damsel flies, finding that per capita fitness of both species was sensitive to total density but not relative density of the two species. Combined with observations that species relative abundances were uncorrelated with environmental variables among lakes, the conclusion was that stochastic processes play a prominent role in driving community dynamics (Siepielski et al. 2010). This experiment recalls studies with Tribolium flour beetles, where under certain combinations of temperature and humidity, two competing species appeared perfectly matched, such that in different replicates either one could drift to dominance or extinction (Mertz et al. 1976). These are compelling cases in which the inference of stochastic community dynamics appears quite strong. Unfortunately, many systems are not amenable to experimentation of this sort, due, for example, to long generation times.
The way forward

Move beyond the stochastic versus deterministic dichotomy

Past studies of stochastic vs. deterministic underpinnings of community dynamics and structure provide a strong foundation for further progress, and our review suggests some important guidelines and novel avenues for future research. First, on a conceptual level, it seems counterproductive to attempt to categorize all ecological processes as stochastic or deterministic. In a local, closed community, this might be appropriate, with drift (stochastic) and selection (deterministic) being the key processes driving community dynamics. In a metacommunity context, however, neatly classifying the consequences of dispersal as stochastic or deterministic is fraught with difficulties (Lowe and McPeek 2014). Dispersal is clearly key to a general understanding of ecological communities, but its effects are most often best characterized based on interactions with selection and drift, rather than lumped under the same umbrella as one or the other of these other processes (see also Bell 2005, Shipley et al. 2012). Likewise, to the extent that speciation—via its effects on the regional species pool—influences (meta)community patterns, there is no unambiguous way to classify such effects as stochastic or deterministic. Such ‘species pool’ effects (which also invoke large-scale dispersal) are often revealed when comparing communities in regions with partly independent evolutionary histories, which effectively redirects the key questions to a larger spatial scale, rather than arbitrating between stochastic versus deterministic models.

Inferring process from pattern: synthesizing different statistical approaches

There is a clear need for synthesis and integration among the different statistical approaches aimed at revealing the processes underlying community dynamics. Since the stochastic component of community dynamics is often inferred from unexplained variation, accounting for the deterministic sources of variation is critical. The null model approach is often applied without environmental data, aiming to reveal the action of interspecific competition, although a pattern of non-random species co-occurrences can just as easily be produced by environmental filtering (spatially-variable selection). While the exact cause of deviations from a null model might seem unimportant for quantifying the stochastic component of community structure, incorporating different constraints (e.g., those expected due to competition or environment) can certainly influence the overall perceived influence of deterministic versus stochastic processes. Multivariate variance partitioning typically incorporates environmental and spatial data, but is not aimed at species interactions. Traits can offer considerable insights into the nature of selection, as can phylogenetic relationships among species as a proxy for trait differences (Cavender-Bares et al. 2009), but Shipley et al.’s (2012) method does not formally incorporate environmental or spatial data. All of these methods start with the same exact type of species × site data, incorporate (or not) environmental, spatial, or trait data, and lead authors to inferences concerning the importance of drift and different kinds of selection or dispersal.

New methods are needed that incorporate all sources of relevant data on characteristics of species (traits, phylogenetic relationships) and sites (environment, spatial context). Indeed there has been a recent push to develop such methods (Dray and Legendre 2008, Leibold et al. 2010, Peres-Neto et al. 2012), although they have yet to be broadly applied. There is also a pressing need to evaluate the ability of statistical methods to permit inferences about processes from patterns. For example, Gilbert and Bennett (2010) demonstrated that both distance-based and raw-data approaches to variance partitioning have generally low precision in revealing underlying environmental and spatial control on community composition, but that large changes in underlying processes are reflected in such analyses. More comprehensive studies are now needed in which community dynamics are simulated under different assumptions, such as variable community size, interspecific competitive interactions, dispersal scenarios, etc. (Rangel et al. 2007, Gotelli et al. 2009, Smith and Lundholm 2010), with the resulting patterns analyzed using the range of methods discussed here. Such studies will deepen our understanding of how to interpret the results of statistical analyses of observational data, which are certain to maintain their currently prominent role in community ecology, given that they offer the greatest scope for comparisons across widely different systems. As a cautionary note, researchers should avoid the pitfall of simulating data using the exact same type of model upon which particular statistical tests are based (Legendre et al. 2005), which biases results towards an unrealistically favorable view of the ability of such analyses to recover process from pattern.

Regardless of the statistical methods used, sampling protocols can be designed to eliminate particular confounding explanations for observed patterns. For example, surprisingly few studies have followed Gilbert and Lechowicz’s (2004) approach of selecting plots specifically to minimize variation in community composition that cannot be uniquely attributed to environmental or spatial variables. In addition, observational studies comparing different systems can sharpen our ability to infer greater or lesser importance of stochastic processes when they focus on systems that are as similar as possible apart from key variables of interest. Examples include studies of multiple sets of ponds (Chase 2007, 2010) or forest patches (Vellend et al. 2007) that vary in a single key variable of interest (e.g. high versus low predation ponds, old versus young forests). Similarly, one can study different subsets of the same exact communities, distinguished either taxonomically (Leibold et al. 2010), or based on traits such as body size (Farjalla et al. 2012). In such studies, explanations for any difference in, for example, unexplained variance in community composition, cannot be attributed to the use of different environmental variables or community types, although one can never rule out the possibility that different results would obtain with different environmental variables or traits.

Match observational studies with conventional and novel experiments

Several new types of experimental study can greatly advance our understanding of stochastic community dynamics. First, there has been a huge number of observational studies, and for some systems experiments can allow targeted tests of the
changing other important variables, such as patch size (and manipulating community size without simultaneously changing other important variables, such as patch size (and therefore perimeter-to-core ratio, or the presence of a top predator) or resource availability per individual (Blakely and Didham 2010). One solution would be to create different ‘effective’ community sizes in aquatic microcosms by manipulating the volume of inoculum used during serial transfers—that is, effectively putting communities through different sized demographic bottlenecks.

Finally, a number of experiments have manipulated colonization order as a putatively stochastic driver of initial community differences, but without relating such manipulations to naturally occurring variation in colonization history. The next generation of such experiments should attempt to mimic natural variation in colonization history or initial community composition (Fig. 2). In addition, if each colonization sequence is replicated (Kreyling et al. 2011), researchers can begin to assess the roles of drift (beta diversity among exact replicates), deterministic divergence via frequency-dependent selection (beta diversity among different colonization histories), and convergence via ‘directional’ selection (very low beta among all replicates) in creating overall community compositional variation among patches in the metacommunity (Fig. 2). In addition, manipulating dispersal among communities after an initial period of assembly can further help distinguish divergence via drift or via selection towards multiple stable states; a high rate of dispersal should counter the effects of drift but have little influence on multiple stable states.

Conclusions

Figure 2. Experimentally distinguishing convergent selection, divergent selection (multiple stable states or attractors), and drift in community dynamics. β = beta diversity, within ((βW) or among ((βs) treatments.

The key variable that should influence the importance of drift is community size (the number of individuals summed across species), but very few studies have manipulated community size (but see Fukami 2004, Blakely and Didham 2010). This may be due, in part, to the great difficulty of manipulating community size without simultaneously changing other important variables, such as patch size (and inferences (i.e. hypotheses) generated by such studies. For example, if neutral stochasticity appears more important in one system (e.g. tropical grasslands) than another (e.g. temperate grasslands), species in the former should show reduced sensitivity of performance to experimental manipulations of relative abundances of other species. Very few studies have combined observations and experiments in this way (but see Siepielski et al. 2010). Another prediction would be that experimental replicates under identical environmental conditions show greater community differentiation over time via drift in systems where neutral stochasticity is predicted to be especially important (Otrrock and Watling 2010, Segura et al. 2011).

Although doubts have been raised concerning the validity of treating stochastic aspects of community dynamics as real (Clark et al. 2007, Clark 2009), it seems clear that causes of the demographic events underlying community dynamics include components that are irreducibly probabilistic—that is, stochastic with respect to species identity. In principle, such effects are clearly distinguished from processes that are deterministic with respect to species identity, although detection remains a considerable practical challenge.

Progress in our general understanding of the role of stochasticity in community ecology can be facilitated by greater precision and care in drawing inferences about processes from various kinds of observations and experiments (Table 1). Although experiments are typically thought of as the only way to unambiguously reveal underlying processes, as often as not, the outcome of an experiment (e.g. manipulating X had a particular effect on Y) leaves much unknown about how and why X influenced Y, which may include a mixture of both stochastic and deterministic processes. We hope our analysis will promote more explicit and precise articulations of how particular results relate to underlying processes, and how these processes map (or not) onto the stochastic–deterministic dichotomy. As for other questions in ecology, there will be no one critical test (Pickett et al. 2007) for the presence or importance of stochasticity underlying community dynamics or structure. Taking a multi-pronged approach seems most fruitful: observations and experiments of various kinds on the same set of communities (Siepielski et al. 2010), and closely comparable methods applied to many different systems (Cottenie 2005). In this way, we can collectively advance the goal of understanding...
the circumstances under which different ecological processes are of greatest importance in structuring ecological communities in nature.

Acknowledgements – MV and DS were supported by grants from the Natural Sciences and Engineering Research Council, Canada. Much of the work on this paper was conducted as part of a graduate course in the Biodiversity Research Centre, Univ. of British Columbia. Pedro Peres-Neto, Jeremy Fox and Emily Grman provided valuable comments on an earlier version of the paper.

References


