

# Severe limitations of the FEve metric of functional evenness and some alternative metrics

Evsey Kosman<sup>1</sup>, Samuel Scheiner<sup>2</sup>, and Hans-Rolf Gregorius<sup>3</sup>

<sup>1</sup>Tel Aviv University

<sup>2</sup>National Science Foundation

<sup>3</sup>Affiliation not available

August 14, 2020

## Abstract

1. The metric of functional evenness FEve is an example of how approaches to conceptualizing and measuring functional variability may go astray. 2. The index of functional evenness FEve has critical conceptual and practical drawbacks: a) Different values of the FEve index for the same community can be obtained if the species have unequal species abundances; this result is highly likely if most of the traits are categorical. b) Very minor differences in even one pairwise distance can result in very different values of FEve. c) FEve uses only a fraction of the information contained in the matrix of species distances. Counterintuitively, this can cause very similar FEve scores for communities with substantially different patterns of species dispersal in trait space. d) FEve is a valid metric only if all species have exactly the same abundances. However, the meaning of FEve in such an instance is unclear as the purpose of the metric is to measure the variability of abundances in trait space. 3. We recommend not using FEve metric in studies of functional variability. Given the wide usage of FEve index over the last decade, the validity of the conclusions based on those estimates are in question. 4. Instead, we suggest three alternative metrics that combines variability in species distances in trait space with abundance in various ways, and more broadly recommend that researchers think about which community properties (e.g., trait-distances of a focus species to the nearest neighbor or all other species, variability of pairwise interactions between species) they want to measure and pick from among the appropriate metrics.

## Abstract

1. The metric of functional evenness FEve is an example of how approaches to conceptualizing and measuring functional variability may go astray.
2. The index of functional evenness FEve has critical conceptual and practical drawbacks:
  1. Different values of the FEve index for the same community can be obtained if the species have unequal species abundances; this result is highly likely if most of the traits are categorical.
  2. Very minor differences in even one pairwise distance can result in very different values of FEve.
  3. FEve uses only a fraction of the information contained in the matrix of species distances. Counterintuitively, this can cause very similar FEve scores for communities with substantially different patterns of species dispersal in trait space.
  4. FEve is a valid metric only if all species have exactly the same abundances. However, the meaning of FEve in such an instance is unclear as the purpose of the metric is to measure the variability of abundances in trait space.
1. We recommend not using FEve metric in studies of functional variability. Given the wide usage of FEve index over the last decade, the validity of the conclusions based on those estimates are in question.

2. Instead, we suggest three alternative metrics that combines variability in species distances in trait space with abundance in various ways, and more broadly recommend that researchers think about which community properties (e.g., trait-distances of a focus species to the nearest neighbor or all other species, variability of pairwise interactions between species) they want to measure and pick from among the appropriate metrics.

Functional trait variability is a component of biodiversity that for the species within a community measures variability in the traits that are assumed to play a role in organismal or ecosystem functions. Many aspects of ecosystem processes depend on the nature, distribution, and variation of organismal traits. Therefore, a proper assessment of functional trait variability is important, and numerous metrics and approaches have been developed since the 1990s to measure this key community attribute (many of them are listed in Scheiner, 2019).

Functional trait variability is a complex concept. To describe its different facets, Villéger, Mason, and Moullot (2008) suggested three separate metrics: functional richness (FRic), functional evenness (FEve), and functional divergence (FDiv), which measure, respectively, the amount of trait space filled by the community, the evenness of species abundances as they are distributed in trait space, and how abundances are spread across trait space. Among the few commonly used approaches, these three metrics are some of the most complicated, but they are assumed to provide an exhaustive measure of functional variability within a community. Despite some criticisms of these indices, mainly focused on functional evenness (e.g. Ricotta, Bacaro, & Moretti, 2014; Legras & Gaertner, 2018), their usage has continually grown in recent years from 134 citations in 2015 to 288 in 2019, with a current total of over 1500 citations. In this paper, we demonstrate that functional evenness (FEve) has severe limitations in its applicability and interpretation. We concentrate on FEve as an example of how approaches to conceptualizing and measuring functional evenness may go astray.

A community can be characterized by its species and their abundances. Using additional information about those species, relationships among the species can be expressed in terms of pairwise distances, that in turn can be used to measure overall community variation. In particular, if each species is described by the same set of  $T$  traits (standardized trait values are assumed), a community of  $S$  species can be represented by  $S$  points in a  $T$ -dimensional trait space. While distances can be estimated with different metrics, relationships are completely predetermined by the species' dispersion in the trait space. Functional trait diversity can be measured in a variety of ways; the differences in trait space among species can be measured using all pairwise distances, the mean distance of a given species from other species, or nearest-neighbor distances (Scheiner, 2019). Those distances can then be further weighted by the species abundances to provide a measure of abundance-weighted functional trait variation within this multi-trait space. FEve measures functional evenness based on abundance-weighted nearest-neighbor distances, so this metric might be relevant if the primary interactions within a community are among species that are most similar in trait values. While such types of interactions occur in many circumstances, there are many circumstances when this is not true for either species or types of interactions. However, the FEve metric has been widely used to analyze functional variation without consideration of the types of processes and entities being considered. We return to this issue in the final section of the paper when we discuss alternative measures of functional variation.

## Conceptual problems

Functional evenness (FEve) is based on a minimum spanning tree (MST) of a complete, undirected network of  $S$  vertices (species) with edges weighted by distance. An MST links all vertices through  $S - 1$  edges such that there are no cycles, i.e., there is only one pathway between any two species. For  $S$  vertices there are  $S^{S-2}$  possible spanning trees. The MST is the tree with the minimum possible total sum of the distances between all pairs of connected vertices (species). Importantly, several MSTs with the same minimum total distance may exist for a given network, if there are edges with the same distances. Such equal distances are highly likely if most of the traits are categorical or meristic (counts). At the extreme, if all edges are of equal distance, there are  $S^{S-2}$  MSTs.

If FEve is to be used as a measure of some property of biodiversity, its conceptual basis needs to be described and justified. In particular, what is the reasoning for the use of MST-edges in combination with abundances as a functional characteristic? Which functional characteristic is addressed by this combination? In what sense is it a measure of evenness? In addressing these questions, we uncover two conceptual problems: (1) the possibility of non-uniqueness of MSTs and, (2) its use as an index of evenness.

Given a particular MST with  $S$  nodes (species  $s_i, i = 1, 2, \dots, S$ ), FEve is calculated as follows. First, each edge linking species  $s_i$  and  $s_j$  with functional distance  $d_{ij} = \text{dist}(s_i, s_j)$  between them is weighted by the sum of their abundances ( $w_i$  and  $w_j$ ):

$$EW_{ij} = \frac{\text{dist}(s_i, s_j)}{w_i + w_j} = \frac{d_{ij}}{w_i + w_j}. \quad (1)$$

Second, those weighted edges are normalized by the sum of the  $EW_{ij}$  values for the corresponding MST:

$$PEW_{ij} = \frac{EW_{ij}}{\sum_{(i,j)=1}^{S-1} EW_{ij}}, \quad (2)$$

where  $(i, j)$  designates an edge between species  $s_i$  and  $s_j$ . (Because of this normalization, either relative or absolute abundances can be used.) Finally, FEve is calculated as:

$$FEve = \frac{\left[ \sum_{(i,j)=1}^{S-1} \min(PEW_{ij}, \frac{1}{S-1}) - \frac{1}{S-1} \right]}{\left[ 1 - \frac{1}{S-1} \right]}, \quad (3)$$

which takes values between 0 and 1 (the denominator is the theoretically possible maximum value of the numerator). According to Villéger et al. (2008), "our new functional evenness index measures both the regularity of branch lengths in the MST and evenness in species abundances." From context, it is also clear that the authors intended the MST branches (edges) to connect nearest neighbors.

The authors do not explicitly identify the characteristics and objects that are the focus of their metric. We do so as follows. The combination of an edge plus abundances ( $PEW_{ij}$  values) serves as the characteristic of interest, with pairs of species being the objects (eq. 1). Evenness of these objects (eq. 2) is the focus of the metric. Evenness is quantified as a deviation of the relative representations from their associated uniform distribution (the numerator in eq. 3).

This approach has several conceptual problems. First, abundance-edge combinations ( $EW_{ij}$  values) do not necessarily represent evenness relationships between nearest neighbors. The internal nodes of an MST have at least two connecting edges. If one edge is smaller than another (e.g.,  $d_{ij} < d_{ik}$ ), its abundance-weighted representation can be larger than that of the second edge ( $EW_{ij} > EW_{ik}$ ) when the sum of abundances of the corresponding species is sufficiently larger ( $w_i + w_k \gg w_i + w_j$ ). Depending on how species abundances are distributed along MST nodes, it is possible that none of those abundance-edge pairs on the MST represent nearest neighbors. Therefore, estimation of functional evenness with FEve does not really mirror a concept of measuring functional variability based on the functionally nearest types (species).

Second, the authors state that "to transform species distribution in a T-dimensional functional space to a distribution on a single axis, we choose to use the minimum spanning tree". No reasoning is given for why such a transformation is required. Nor is it explained in what way an MST can be considered as yielding a distribution on a single axis, given that nodes can connect to more than two others.

Third, the functional relevance of combining MST-distances (edge values) with abundances is simply assumed. The use of abundances assumes that any and all functional traits have a similar per capita functional effect.

Fourth, the potential for a single set of points to have multiple MSTs is ignored. While one can demonstrate that the distribution of edge values is the same for all alternative MSTs, this property is lost when they are combined with species abundances. As we show in the next section, such combinations can lead to more than one value of evenness for the same data set.

## Inferences from constructed examples

Multiple MSTs can result in multiple, different values of FEve index for the same community if the species have unequal species abundances, which severely limits the utility of the metric. The following example demonstrates such a situation. Let the community consist of three equally distant species ( $s_1, s_2$ , and  $s_3$ ) in a given trait spaces (i.e.  $\text{dist}(s_1, s_2) = d_{12} = d_{13} = d_{23} = d$ ) with abundances  $w_1 = 1$ ,  $w_2 = 2$  and  $w_3 = 3$ , respectively (Fig. 1, community network). There are three MSTs with the same minimum total distance ( $2d$ ): MST<sub>1</sub> with one edge connecting  $s_1$  and  $s_2$ , and one edge connecting  $s_2$  and  $s_3$  (Fig. 1, MST<sub>1</sub>); MST<sub>2</sub> with edges connecting  $s_1$  and  $s_2$ , and  $s_1$  and  $s_3$  (Fig. 1, MST<sub>2</sub>); and MST<sub>3</sub> with edges connecting  $s_1$  and  $s_3$ , and  $s_2$  and  $s_3$  (Fig. 1, MST<sub>3</sub>). The three trees result in different estimates for FEve (Fig. 1). Thus, for any community there is a likelihood for multiple FEve estimates making interpretation of any estimates suspect.

This problem does not arise if all distances for a given network are different; then there will be only one, unique MST. Such differences in distances are likely if all or most of the traits are quantitative. However, such unique values for FEve do not solve the underlying conceptual problems.

Now consider the three MSTs in Fig. 1 to be three different communities and the distances between the species no longer identical, but just very, very slightly different so that each MST is unique for that community (e.g., for MST<sub>1</sub>  $d_{12} = d_{23} = 1$  and  $d_{13} = 1.0001$ ; for MST<sub>2</sub>  $d_{12} = d_{13} = 1$  and  $d_{23} = 1.0001$ ; for MST<sub>3</sub>  $d_{13} = d_{23} = 1$  and  $d_{12} = 1.0001$ ). Intuition says that the three communities have nearly the same evenness, and yet they have very different values of FEve.

Additional, hidden pitfalls come about from how FEve is often calculated. Rather than using the original matrix of pairwise distances, PCoA or Multidimensional Scaling (MDS) is used first to transform the distance matrix, and then only the first two or three axes of the transformed space are considered when calculating species' distances (e.g., Mouillot, Villéger, Scherer-Lorenzen, & Mason, 2011; Taudiere, & Violle, 2016). This transformation generally results in a distribution of nodes with no equal distances so that the corresponding MST is unique. However, because of the dimensional reduction, the new pairwise distances are only approximations of the original ones, and the corresponding FEve estimate depends on accuracy of PCoA performance (goodness of fit of the approximations to the original distances). While one could argue that the problems with FEve can be solved by always using untransformed distances, doing so does not guarantee a solution to the other problems listed above.

There is one circumstance that non-unique MSTs result in the same FEve values. This can happen if all species have exactly the same abundances. This equality occurs because any two MSTs of a given network have the same distribution of the edge weights. However, the meaning of FEve in such an instance is unclear as the purpose of the metric is to measure variability of abundances in trait space.

A central reason for the problems raised above is that FEve uses only a fraction of the information contained in the matrix of species distances. Only  $S - 1$  of the  $\frac{(S-1) \times S}{2}$  pairwise distances are used in the calculation of FEve; the much larger portion of the distances are simply ignored. This can cause the same FEve scores for communities with different patterns of species dispersal in trait space (Fig. 2). In our example, this result occurs because the distance between species 1 and 3 is ignored. In addition, it is possible to have a community where FEve = 1 even when neither species abundances nor distances between species are evenly distributed (Fig. 3). In general, complete evenness (FEve = 1) is realized if and only if all PEW<sub>ij</sub> values are equal (eqs. 2, 3), which does not necessarily imply that all distances or all abundances are equal. This behaviour contradicts the claim of Villéger, Mason, and Mouillot (2008; p. 2293) that, "FEve decreases either when abundance is less evenly distributed among species or when functional distances among species are less regular." Their claim is correct as an absolute statement only if the other factor (abundances or distances) are held constant, which will not occur when comparing actual communities.

## Examples from data

Our constructed examples demonstrate the potential problems with the FEve metric. Here we show how the problem of multiple estimates from a single dataset emerges with actual data. Importantly, there is no way to know in advance the number or range of different FEve estimates for a given dataset. Our first example is the traditional type of data used for functional trait analyses: bats and feeding traits. The other three

examples are from less commonly used data: genetic profiles where the traits are the presence and absence of different genes. These examples demonstrate the problem of multiple MSTs that arises with non-continuous traits. For the two examples that lack actual abundance data, we show how a single distance matrix can result in multiple, disparate FEve estimates with simulated abundances. For the other two examples, we show analyses with both actual abundances and two sets of simulated abundances to show how different types of abundance distributions can result in highly variable FEve estimates.

Bats and feeding traits .

The first example consists of a set of five bat species (*Carollia manu* , *Chiroderma salvini* , *Dermanura glauca* , *Enchisthenes hartii* , and *Micronycteris megalotis* ) in the Manu Biosphere Reserve located on the eastern slopes of the Andes in southeastern Peru. Our analysis was based on species characterization with 16 binary categorical traits (Table S3 in Scheiner, Kosman, Presley, & Willig 2017) that were separated into three groups: diet (fruit, nectar, invertebrates, vertebrates, fish, blood), foraging location (open areas, over water, above canopy, canopy, subcanopy, understory), and foraging strategy (aerial, gleaning, hovering, other). To determine the functional distance between species, Jaccard dissimilarity was calculated for each group of binary traits, and then the combined distance between species was determined by an equal-weight averaging of the three group-specific dissimilarities (Table 1). Because the distance matrix contains many equal values, three different MSTs can be generated (Table 1). Because abundance data were not available, we provided two different sets of simulated values. For each set of simulated abundances, the multiple MSTs resulted in FEve estimates that varied 16% and 28%, respectively, between the smallest and largest values (0.374 to 0.480; and 0.676 to 0.785).

Bryozoan genotypes .

*Cristatella mucedo* is a diploid freshwater bryozoan. We used data on eight microsatellite loci (Table 2 in Kosman & Jokela, 2019) for ten genetically separate individuals from bryozoan colonies in Lake Aegery, Switzerland. The distance between the genotypes was calculated by assuming a stepwise mutation model of microsatellite evolution with variable rates of mutations at different loci (SMMv; Kosman & Jokela, 2019). The corresponding matrix of pairwise distances is presented in Table 2. Abundance data were not available, so we provided simulated values. Again, multiple MSTs can be generated based on the distance matrix that result in four different FEve estimates (Table 2) that ranged from 0.533 to 0.635.

*Wheat fungal pathogen (Puccinia graminis f. sp. tritici) genotypes* .

The data consisted of eleven virulence phenotypes of *P. graminis* isolates collected from bread wheat in the Novosibirsk region of Russia. The binary phenotypes (virulence/avirulence) were determined with a set of twenty North American wheat differential lines (Skolotneva et al., 2020). The distance between the phenotypes was calculated using simple mismatch dissimilarity; the corresponding matrix of pairwise distances are presented in Table 3. Twenty-four different MSTs can be generated (Table 3). For the actual abundances, ten different FEve estimates ranged from 0.659 to 0.737 (Fig. 4). Even minor changes in abundances resulted in substantial changes in number and values of different FEve estimates: for the Y-modification, twenty-four values ranged from 0.708 to 0.793; for the Z-modification, eighteen values ranged from 0.573 to 0.695 (Fig. 4).

*Wheat fungal pathogen (Puccinia triticina Erikss) genotypes* .

The data consist of eleven genotypes of single-uredinal isolates of *P. triticina* (a dikaryotic fungus) collected from durum wheat in Russia using eleven microsatellite markers (Table 3 in Kosman & Jokela, 2019; Gulyaeva et al., 2017). The distance between the microsatellite genotypes was calculated assuming an infinite alleles model (IAM; Kosman & Leonard, 2005), and the corresponding matrix of pairwise distances is presented in Table 4A. Three different MSTs can be generated based on the distance matrix (Table 4B). We compared the FEve estimates for the actual abundances with simulated values for three scenarios: (1) two dominant and nine rare types (simulation P), nine dominant and two rare types (simulation R), and all types equally abundant (simulation E). For the real abundances, FEve values ranged from 0.612 to 0.651

(about 7%). For simulation P, the values have a wider range (0.711 – 0.801, around 13%). For simulation R, the values have a very wide range, from 0.234 to 0.828 (about 354%), which shows the outsized influence of differences in MSTs when the node has a high abundance. For simulate E, as expected, equally abundant types resulted in the same value of 0.88 for all MSTs, despite their variation.

## Summary of FEve issues

In constructing species networks, it is assumed that trait values are measured without error and that there is no variation within species, two assumptions that we know are false. This issue could be addressed by a procedure that would estimate the mean and variability of relevant estimates over all closely related networks. We do not know of any attempt to study that matter for any diversity metric. Nevertheless, this problem seems much more acute for FEve comparing with many other metrics because of the potential for non-unique MSTs in addition to errors in distance measures. Multiple estimates of a diversity metric obtained for different, closely-related networks are natural. However, it is conceptually incorrect to assume that functional evenness has multiple values for a community represented by a single network. Given the wide usage of FEve index over the last decade, the validity of the conclusions from those studies are now in question. Our examples show that a single dataset can result in considerable variability in FEve estimates, especially when the data include rare types. The combination of functional relationships and abundances (species distance divided by sum of their abundances) into a single assessment of evenness results in a metric that fails to distinguish between distance evenness and abundance evenness (Gregorius, 1990).

This entire paper has been about FEve, but we would be remiss if we do not mention PEve – phylogenetic evenness – which was defined by Dehling et al. (2014) to be identical to FEve, but substituting nearest-neighbor phylogenetic distances for distances in functional-trait space. We discussed the non-uniqueness problem with FEve that occurs when you have two species which have identical nearest-neighbor distances to a third, but differ in their abundances. This problem is most likely for categorical traits or those based on counts with just a few possible values and so many not occur that often. However, this problem is highly likely for phylogenetic data. It will occur any time you have a pair of sister species that are equally distant from a third and that differ in their abundances. PEve has been used much less frequently than FEve, but should also be abandoned. As with functional traits, there are alternative for phylogenetic evenness that can measure the same properties while avoiding the uniqueness problem (Tucker et al. 2017, Scheiner 2019).

## Next steps

We have shown that FEve has critical conceptual and practical drawbacks, and therefore, we recommend not using this index in studies of functional variability. However, it is still possible to measure evenness of functional traits combined with information about abundances using alternative methods that do not have the limitations of FEve. An alternative metric based on Hill numbers that combines nearest-neighbor distances with abundances is the evenness derivative of the diversity metric of Scheiner (2012):

$${}^qD(AT_N) = \left( \sum_{i=1}^S \left( \frac{n_i d_{i \min}}{\sum_{j=1}^S n_j d_{j \min}} \right)^q \right)^{\frac{1}{(1-q)}} \quad (4)$$

$${}^qE(AT_N) = \frac{{}^qD(AT_N)}{S}, \quad (5)$$

where  ${}^qD(AT_N)$  is effective number of distinct species that equally contribute to functional interaction and variability within a community based on nearest-neighbor distances ( $n_i d_{i \min} = n_j d_{j \min}$  for all  $i \neq j$ ),  $S$  is the number of species,  $n_i$  is the number of individuals of species  $i$ ,  $d_{i \min}$  is the nearest-neighbor distance of species  $i$ , and  $q$  is the exponent of the Hill function. [The metrics here and below follow the symbol convention of Scheiner (2019).] This metric measures the evenness of the joint distribution of abundances and nearest-neighbor distances. Because each species has a unique nearest-neighbor distance, the resulting metric always has a single value and small deviations of those values will result in only small changes in the metric, eliminating the problems that we outlined above for FEve. It should be used in conjunction with an examination of the separate evennesses of abundances [ ${}^qE(A)$ ] and nearest-neighbor distances [ ${}^qE(T_N)$ ]. For example, it is possible that neither parameter is evenly distributed singly, but that the joint distribution

has an even distribution, which can occur if they are strongly negatively correlated. Such a combination of values would then point to the potential importance of processes that jointly affect traits and abundances (e.g., competitive exclusion).

An alternative approach for combining trait distance and abundance information is the use of the abundance-weighted distance of species  $i$  from all of the other  $S - 1$  species:

$$d_i = \sum_{\substack{k=1 \\ k \neq i}}^S d_{ik} \left( \frac{n_k}{N} \right), \quad (6)$$

where  $N = \sum_{j=1}^S n_j$  is the total number of individuals in the assemblage. Then functional diversity can be estimated in terms of Hill numbers as:

$${}^q D(AT_T) = \left( \sum_{i=1}^S \left( \frac{n_i d_i}{\sum_{j=1}^S n_j d_j} \right)^q \right)^{\frac{1}{1-q}}, \quad (7)$$

which is the effective number of distinct species that equally contribute to functional interaction and variability within a community based on abundances and weighted distances of every species from all other species ( $n_i d_i = n_j d_j$  for all  $i \neq j$ ). From this, we can obtain an evenness measure as:

$${}^q E(AT_T) = \frac{{}^q D(AT_T)}{S}. \quad (8)$$

This measure of evenness would be appropriate if a given species interacts with all of the other species in a community in a way that ‘averages’ over all of those interactions (e.g., in a system with diffuse competition).

The evenness metrics given in eqs. 5 and 8 are based on the individual properties of each species. An alternative approach is to measure functional variation based on pairs of species:

$${}^q H(AT_P) = \left( \sum_{i=1}^S \sum_{j=1}^S \left( \frac{n_i n_j d_{ij}}{\sum_{k=1}^S \sum_{l=1}^S n_k n_l d_{kl}} \right)^q \right)^{\frac{1}{1-q}}, \quad (9)$$

which measures the effective number of equally interacting pairs of species (equal values of  $n_i n_j d_{ij}$  for all  $i, j = 1, 2, \dots, S$ ,  $i \neq j$ ) (see eq. A23 in Scheiner et al., 2017), so that the number of equally interacting species is determined as follows:

$$\frac{{}^q D(AT_P) = \left( 1 + \sqrt{1 + 4^q H(AT_P)} \right)}{2}, \quad (10)$$

(eqs. 4 and A4, Scheiner et al., 2017). The corresponding metric of functional evenness is then:

$${}^q E(AT_P) = {}^q D(AT_P) / S. \quad (11).$$

This measure of evenness would be appropriate if the pairwise interactions are important and those interactions occur with all of the other species in the community (e.g., scramble competition for a spectrum of resources). The metrics presented here (eqs. 4 – 11), as well as FEve itself, assume that all individuals within a species are identical; somewhat different forms are necessary to capture within-species variation.

More general concepts (Gregorius & Kosman 2017, 2018) and a large variety of metrics (Scheiner 2019) exist for measuring functional variation, and can be used as alternative for FEve. We caution, though, that many of them have not yet been critically evaluated. The metrics suggested here (eqs. 5, 8, and 11) are all based on a concept of diversity of the dispersion of an effective number of types. Division of this effective number by the actual number of types turns these into metrics of functional evenness. While there is no single best way to measure functional-trait evenness or its combination with abundance, there are metrics, such as FEve, that should be avoided.

## Authors’ contributions

E.K. conceived the idea of the research and wrote the first draft. S.M.S. and H-R.G. contributed to the development and interpretation of examples. All authors developed the manuscript, and S.M.S. prepared the final version.

## Acknowledgments

This manuscript is based on work done by S.M.S. while serving at the U.S. National Science Foundation. The views expressed in this paper do not necessarily reflect those of the National Science Foundation or the United States Government.

## References

- Dehling, D. M., Fritz, S. A., Töpfer, T., Päckert, M., Estler, P., Böhning-Gaese, K., & Schleuning, M. (2014) Functional and phylogenetic diversity and assemblage structure of frugivorous birds along an elevational gradient in the tropical Andes. *Ecography*, 37, 1047-1055.
- Gregorius, H-R. (1990) A diversity-independent measure of evenness. *American Naturalist*, 136, 701-711.
- Gregorius H.-R., & Kosman E. (2017) On the notion of dispersion – from dispersion to diversity. *Methods in Ecology and Evolution*, 8, 278–287.
- Gregorius H.-R., & Kosman E. (2018) Structural type diversity: measuring structuredness of communities by type diversity. *Theoretical Ecology*, 11, 383-394.
- Gulyaeva, E. I., Aristova, M. K., Shaidayuk, E. L., Mironenko, N. V., Kazartsev, I. A., Akhmetova, A., & Kosman, E. (2017) Genetic differentiation of *Puccinia triticina* Erikss. in Russia. *Russian Journal of Genetics*, 53, 1053–1060.
- Kosman, E., & Jokela, J. (2019) Dissimilarity of individual microsatellite profiles under different mutation models – empirical approach. *Ecology and Evolution*, 9, 4038–4054.
- Kosman, E., & Leonard, K. J. (2005) Similarity coefficients for molecular markers in studies of genetic relationships between individuals for haploid, diploid, and polyploid species. *Molecular Ecology*, 14, 415-424.
- Legras, G., & Gaertner, J. C. (2018) Assessing functional evenness with the FEve index: a word of warning. *Ecological Indicators*, 90, 257–260.
- Moullot, D., Villeger, S., Scherer-Lorenzen, M., & Mason, N. W. H. (2011) Functional Structure of Biological Communities Predicts Ecosystem Multifunctionality. *PLoS One*, 6, e17476, doi:10.1371/journal.pone.0017476.
- Ricotta, C., Bacaro, G., & Moretti, M. (2014) A new measure of functional evenness and some of its properties. *PloS One*, 9, e104060.
- Scheiner, S. M. (2012) A metric of biodiversity that integrates abundance, phylogeny, and function. *Oikos*, 121, 1191-1202.
- Scheiner, S. M. (2019) A compilation of and typology for abundance-, phylogenetic- and functional-based diversity metrics. *bioRxiv*, 530782.
- Scheiner, S. M., Kosman, E., Presley, S. J., & Willig, M. R. (2017) Decomposing functional diversity. *Methods Ecol. Evol.*, 8, 809–820.
- Skolotneva, E. S., Kosman, E., Kelbin, V., Morgounov, A., Shamanin, V. P., & Salina, E. A. (2020) Virulence phenotypes of Siberian wheat stem rust population in 2017-2018. *Front. Agron.* doi: 10.3389/fagro.2020.00006.
- Taudiere, A., & Violle, C. (2016) cati: an R package using functional traits to detect and quantify multi-level community assembly processes. *Ecography*, 39, 699–708.



Tucker, C. M., Cadotte, M. W., Carvalho, S. B., Davies, T. J., Ferrier, S., Fritz, S. A., Grenyer, R., Helmus, M. R., Jin, L. S., Mooers, A. O., Pavoine, S., Purschke, O., Redding, D. W., Rosauer, D. F., Winter, M., & Mazel, F. (2017) A guide to phylogenetic metrics for conservation, community ecology and macroecology. *Biological Reviews*, 92, 698-715.

Villegger, S., Mason, N. W. H., & D. Mouillot. (2008) New multidimensional functional diversity indices for a multifaceted framework in functional ecology. *Ecology*, 89, 2290–2301.

Table 1. For a set five bat species: pairwise distances, abundances, MST attributes, and estimates of functional evenness FEve. Because no abundance data are available, we generated simulated values. The weights of the edges between each pair of MST nodes is equal to the distance between those species.

		$B_1$	$B_2$	$B_3$	$B_4$	$B_5$	abundance <sup>a</sup>	MST edges	MST edges
<i>Carollia</i>	$B_1$	0	0.333	0.333	0.333	0.167	1 20	$B_i - B_j$	we
<i>manu</i>									
<i>Chiroderma</i>	$B_2$	0.333	0	0.167	0.222	0.5	5 20	1 - 5	0.1
<i>salvini</i>									
<i>Dermanura</i>	$B_3$	0.333	0.167	0	0.167	0.5	10 20	1 - 2; 1 - 3; 1 - 4 <sup>b</sup>	0.3
<i>glauc</i>									
<i>Enchisthenes</i>	$B_4$	0.333	0.222	0.167	0	0.5	20 1	3 - 4	0.1
<i>hartii</i>									
<i>Micronycteris</i>	$B_5$	0.167	0.5	0.5	0.5	0	1 20	3 - 2	0.1
<i>megalotis</i>									

<sup>a</sup> Two sets of simulated absolute abundances: Y (bold italic ) and Z.

<sup>b</sup> Because species  $B_1$  is equally distant to three others, there are three possible MSTs. For set Y, the MSTs with edges 1 - 2, 1 - 3, and 1 - 4 resulted in FEve estimates of 0.476, 0.480 and 0.374, respectively; for set Z, the values were 0.785, 0.785 and 0.676, respectively.

Table 2. Ten microsatellite genotypes of freshwater bryozoan *Cristatella mucedo* (Cm): pairwise distances, abundances, MST attributes, and estimates of functional evenness FEve. Because no abundance data are available, we generated simulated values. The weights of the edges between each pair of nodes is equal to the distance between those genotypes.

(A)	$Cm_1$	$Cm_2$	$Cm_3$	$Cm_4$	$Cm_5$	$Cm_6$	$Cm_7$	$Cm_8$	$Cm_9$	$Cm_{10}$	abundance	MST edges
$Cm_1$	0	0.066	0.021	0.009	0.066	0.052	0.054	0.078	0.083	0.124	1	0
$Cm_2$	0.066	0	0.045	0.057	0.114	0.014	0.012	0.102	0.106	0.147	2	1
$Cm_3$	0.021	0.045	0	0.012	0.069	0.059	0.057	0.057	0.062	0.103	3	4
$Cm_4$	0.009	0.057	0.012	0	0.057	0.047	0.045	0.069	0.073	0.114	4	3
$Cm_5$	0.066	0.114	0.069	0.057	0	0.104	0.102	0.012	0.016	0.057	5	2
$Cm_6$	0.052	0.014	0.059	0.047	0.104	0	0.002	0.116	0.12	0.161	6	7
$Cm_7$	0.054	0.012	0.057	0.045	0.102	0.002	0	0.114	0.118	0.159	7	4
$Cm_8$	0.078	0.102	0.057	0.069	0.012	0.116	0.114	0	0.004	0.046	8	5
$Cm_9$	0.083	0.106	0.062	0.073	0.016	0.12	0.118	0.004	0	0.05	9	8
$Cm_{10}$	0.124	0.147	0.103	0.114	0.057	0.161	0.159	0.046	0.05	0	10	8

<sup>a</sup> Simulated absolute abundances.

<sup>b</sup> There are four different MSTs with one of these edges. These MSTs resulted in four different FEve values: 0.533, 0.553, 0.612, 0.635.

Table 3. Eleven virulence phenotypes of *P. graminis* (Pgt): pairwise distances, abundances, MST attributes, and estimates of functional evenness FEve. The weights of the edges between each pair of nodes is equal to the distance between those phenotypes. To show the effects of variation in abundances, besides the actual values, we calculated FEve for each of two pairs of altered values.

	<i>Pgt</i> <sub>1</sub>	<i>Pgt</i> <sub>2</sub>	<i>Pgt</i> <sub>3</sub>	<i>Pgt</i> <sub>4</sub>	<i>Pgt</i> <sub>5</sub>	<i>Pgt</i> <sub>6</sub>	<i>Pgt</i> <sub>7</sub>	<i>Pgt</i> <sub>8</sub>	<i>Pgt</i> <sub>9</sub>	<i>Pgt</i> <sub>10</sub>	<i>Pgt</i> <sub>11</sub>	abundance	MST edge
<i>Pgt</i> <sub>1</sub>	0	0.05	0.1	0.15	0.1	0.1	0.15	0.05	0.15	0.3	0.45	1	<i>Pgt</i> <sub>i</sub>
<i>Pgt</i> <sub>2</sub>	0.05	0	0.05	0.1	0.15	0.15	0.2	0.1	0.1	0.35	0.4	1	<i>Pgt</i> <sub>j</sub>
<i>Pgt</i> <sub>3</sub>	0.1	0.05	0	0.05	0.2	0.1	0.15	0.15	0.15	0.3	0.35	(3) <sup>a</sup>	1 - 2
												2	2 - 3
												(8) <sup>b</sup>	3 - 4
<i>Pgt</i> <sub>4</sub>	0.15	0.1	0.05	0	0.15	0.15	0.2	0.2	0.1	0.35	0.3	2	1 - 8
<i>Pgt</i> <sub>5</sub>	0.1	0.15	0.2	0.15	0	0.2	0.15	0.15	0.15	0.3	0.35	2	1 -
<i>Pgt</i> <sub>6</sub>	0.1	0.15	0.1	0.15	0.2	0	0.15	0.15	0.25	0.2	0.35	(4) <sup>a</sup>	6; 3
												1	- 6 <sup>c</sup>
<i>Pgt</i> <sub>7</sub>	0.15	0.2	0.15	0.2	0.15	0.15	0	0.2	0.3	0.25	0.3	2	1 - 5
<i>Pgt</i> <sub>8</sub>	0.05	0.1	0.15	0.2	0.15	0.15	0.2	0	0.1	0.25	0.4	9	2 - 9
													4 - 9
<i>Pgt</i> <sub>9</sub>	0.15	0.1	0.15	0.1	0.15	0.25	0.3	0.1	0	0.35	0.3	1	8 - 9
													1 - 7
													3 - 7
													5 - 7
													6 - 7
<i>Pgt</i> <sub>10</sub>	0.3	0.35	0.3	0.35	0.3	0.2	0.25	0.25	0.35	0	0.25	2	6 - 1
<i>Pgt</i> <sub>11</sub>	0.45	0.4	0.35	0.3	0.35	0.35	0.3	0.4	0.3	0.25	0	8 (2) <sup>b</sup>	10 -
													11

<sup>a</sup> Modification Y of abundances of two phenotypes.

<sup>b</sup> Modification Z of abundances of two phenotypes; the total number of individuals was not changed.

<sup>c</sup> Twenty-four different MSTs are possible with one of these edges. The multiple MSTs resulted in ten, twenty-four, and eighteen different FEve values for actual abundances, Y-modification, and Z-modification, respectively. Variability of FEve estimates is shown in Fig. 4.

Table 4. (A) Eleven microsatellite genotypes of *P. tritici* (Pt): pairwise distances, abundances and MST attributes. The weights of the edges between each pair of nodes is equal to the distance between those genotypes. To show the effects of variation in abundances, besides the actual values, we calculated FEve for two instances of altered values. (B) Estimates of functional evenness FEve.

(A)	$Pt_1$	$Pt_2$	$Pt_3$	$Pt_4$	$Pt_5$	$Pt_6$	$Pt_7$	$Pt_8$	$Pt_9$	$Pt_{10}$	$Pt_{11}$	abundances	MST edges
$Pt_1$	0	0.045	0.045	0.091	0.045	0.136	0.091	0.091	0.136	0.182	0.182	<b><i>10<sup>a</sup></i></b> <b><i>1<sup>b</sup></i></b> <b><i>20<sup>c</sup></i></b>	$Pt_i - Pt_j$
$Pt_2$	0.045	0	0.091	0.136	0.091	0.091	0.045	0.045	0.182	0.136	0.136	<b><i>12 1</i></b> <b><i>20</i></b>	1 - 2
$Pt_3$	0.045	0.091	0	0.136	0.091	0.182	0.045	0.136	0.091	0.136	0.227	<b><i>6 1</i></b> <b><i>20</i></b>	1 - 3
$Pt_4$	0.091	0.136	0.136	0	0.045	0.227	0.182	0.091	0.227	0.273	0.273	<b><i>1 1</i></b> <b><i>20</i></b>	1 - 5
$Pt_5$	0.045	0.091	0.091	0.045	0	0.182	0.136	0.045	0.182	0.227	0.227	<b><i>1 1</i></b> <b><i>20</i></b>	5 - 4
$Pt_6$	0.136	0.091	0.182	0.227	0.182	0	0.136	0.136	0.091	0.045	0.045	<b><i>2 1</i></b> <b><i>20</i></b>	2 - 7
$Pt_7$	0.091	0.045	0.045	0.182	0.136	0.136	0	0.091	0.136	0.091	0.182	<b><i>4 20</i></b> <b><i>1</i></b>	2 - 8
$Pt_8$	0.091	0.045	0.136	0.091	0.045	0.136	0.091	0	0.227	0.182	0.182	<b><i>1 1</i></b> <b><i>20</i></b>	2 - 6 3 - 9 7 - 9
$Pt_9$	0.136	0.182	0.091	0.227	0.182	0.091	0.136	0.227	0	0.045	0.136	<b><i>1 20</i></b> <b><i>1</i></b>	6 - 1
$Pt_{10}$	0.182	0.136	0.136	0.273	0.227	0.045	0.091	0.182	0.045	0	0.091	<b><i>4 1</i></b> <b><i>20</i></b>	10 -
$Pt_{11}$	0.182	0.136	0.227	0.273	0.227	0.045	0.182	0.182	0.136	0.091	0	<b><i>1 1</i></b> <b><i>20</i></b>	6 - 1

(B) MST	abundances	abundances	abundances	abundances
	actual	simulated P	simulated R	equal
1 (2-6) <sup>d</sup>	0.651	0.801	0.828	0.880
2 (3-9)	0.631	0.724	0.753	0.880
3 (7-9)	0.612	0.711	0.234	0.880

<sup>a</sup> Actual abundances (bold italic ).

<sup>b</sup> Simulation P of abundances (*italic* ).

<sup>c</sup> Simulation R of abundances (*italic underline* ).

<sup>d</sup> There are three different MSTs with one of these edges.

Figure captions

Figure 1. A community network in which the distances between all three species are identical, which results in three possible minimum spanning trees (MSTs) and multiple FEve estimates for the same community.

Figure 2. In these communities, two of the three species are equally distant in both communities ( $d_{12} = d_{23} = d_{13}$ ) with a distance that is smaller than the third distance ( $d_{13}$ ). If the abundances of the three species are  $w_1 = 1$ ,  $w_2 = 2$  and  $w_3 = 3$ , then  $FEve = 0.75$ ), even though community B seems much more functionally irregular than community A.

Figure 3. This community consists of three species in which  $d_{23}$  is larger than  $d_{13}$  and  $d_{12}$ . The abundances ( $w$ ) and distances result in values of  $EW_{12} = EW_{13} = \frac{1}{6}$ ,  $PEW_{12} = PEW_{13} = 0.5$ , and  $FEve = 1$ .

Figure 4. Variability of FEve estimates for the actual abundances of eleven virulence phenotypes of *P. graminis*, and the Y- and Z- modification of abundances (see Table 3 for details).

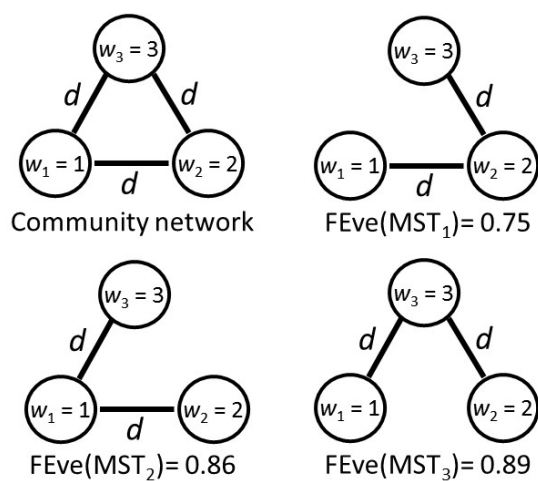


Figure 1

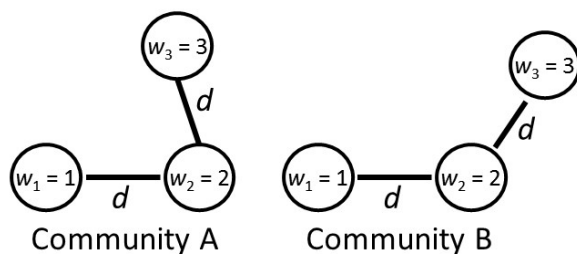


Figure 2

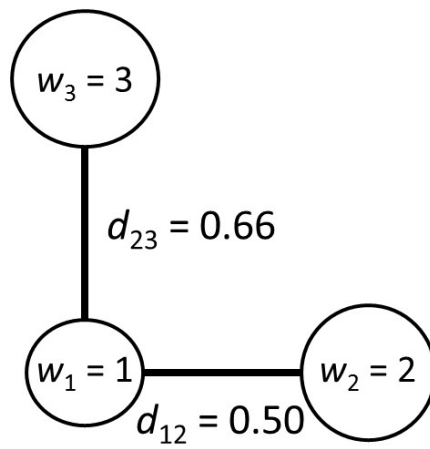


Figure 3

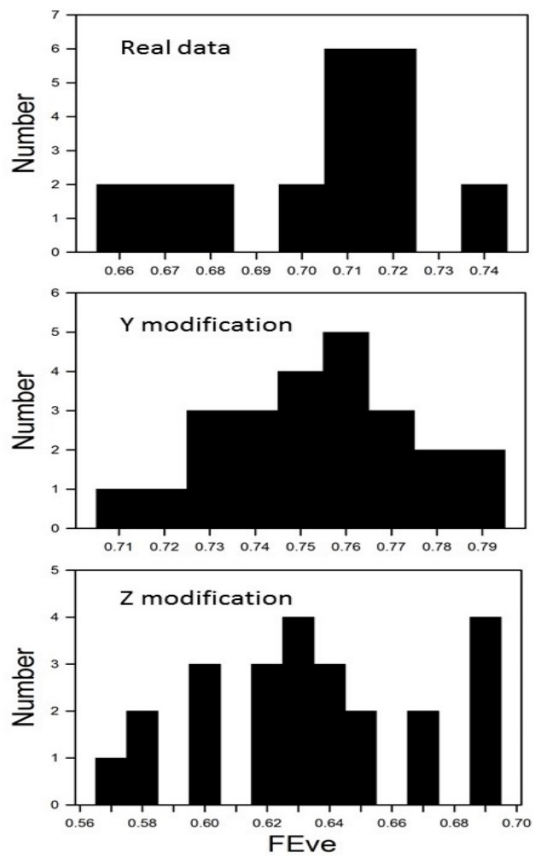


Figure 4