Commentary

Assessing the relationship between biodiversity and stability of ecosystem function—is the coefficient of variation always the best metric?

Tim Carnus, John A. Finn, Laura Kirwan, and John Connolly

Tim Carnus, (tim.carnus@gmail.com), Teagasc, Environment Research Centre, Johnstown Castle, Co. Wexford, Ireland and UCD School of Mathematical Sciences, University College Dublin, Dublin 4, Ireland

John A. Finn (john.finn@teagasc.ie), Teagasc, Environment Research Centre, Johnstown Castle, Co. Wexford, Ireland

Laura Kirwan (laura.kirwan@outlook.com), Teagasc, Environment Research Centre, Johnstown Castle, Co. Wexford, Ireland and Waterford Institute of Technology, Cork Road, Waterford, Ireland

John Connolly (john.connolly@ucd.ie), UCD School of Mathematical Sciences, University College Dublin, Dublin 4, Ireland

Abstract

The role of biodiversity in regulating the stability of ecosystem functioning has importance for the reliable delivery of ecosystem services. To date, ecological studies that aim to measure stability in ecosystem function across a range in diversity have almost universally used the coefficient of variation (CV, the ratio of standard deviation (SD) of functional response to its mean), or its inverse $1/\text{CV}$, in reaching conclusions. We argue that the use of CV for this purpose can lead to misleading conclusions on stability. We consider that defining the stability of ecosystem functioning solely in terms of the CV confines the term stability to a usage and context that is not natural or intuitive to many who wish to address questions about the reliability with which ecosystems deliver services. We use illustrative scenarios to show that an assessment of stability based on the CV is not as effective in many cases as one based on joint consideration of mean and standard deviation, and may be completely misleading, especially where low values of functional response are a desirable outcome. Faced with similar questions, agronomic studies that aim to assess the stability of ecosystem function (comparison of yield of different varieties within and across different sites) take both the average response and variability within- and between-sites into

consideration. We argue that the way stability is measured should be appropriate for the questions about the delivery of ecosystem services that are being addressed. We suggest approaches based on the joint modelling of the mean and standard deviation as a basis for addressing questions of biodiversity and stability of ecosystem function. Assessment of the importance of diversity in providing ecosystem services for society is more likely to be made on socio-economic evaluation of trade-offs between mean and variability of the function rather than its stability as measured by the coefficient of variation.

Keywords: biodiversity, ecosystem function, stability, temporal stability, mixed models, threshold, penalty, coefficient of variation.

Introduction

Stability in provision of ecosystem function is an important property of ecosystems, and the ability of ecosystems to maintain good levels of desired functions with relatively low levels of variation is inherently

appealing, allowing for trade-offs with other desirable functions. There is strong evidence that biodiversity positively influences ecosystem functioning in many ecosystems (Sala et al. 2000, Balvanera et al. 2006, Cardinale et al. 2006, Kirwan et al. 2007). However, despite receiving considerable attention and multiple methodological developments (de Mazancourt et al. 2013) its role in regulating the stability of ecosystem functioning (functional or temporal stability) is still unclear (Hooper et al. 2005, Griffin et al. 2009). This lack of clarity is due in part to issues with defining stability: the term ‘stability’ has been used with very different meanings in many different ecological contexts (Grimm and Wissel 1997, Tilman 1999, Ives and Carpenter 2007). In many ecological studies, a key metric for assessing stability has been based on the coefficient of variation (CV) of the functional response, defined as the ratio of the standard deviation to the mean response. The metric $1/CV$ has also been proposed (Tilman 1999, Lehman and Tilman 2000). Here we examine this metric and analyse its appropriateness to address many questions on the ability of ecosystems to deliver reliable functioning and services, questions that might appear to users to fall within the ambit of the term stability.

The way in which stability is measured should be appropriate to address questions posed in respect of the reliable delivery of ecosystem services, and how delivery may be affected by changing levels of biodiversity, environments, time and space. The influence of diversity on stability has been assessed by regression of the CV of the functional response on some measure of either species richness (McNaughton 1977, Tilman et al. 2006, Hector et al. 2010, Campbell et al. 2011) or functional richness (Weigelt et al. 2008). Systems with lower CV values are said to have greater stability, and declining CV with increasing diversity is interpreted as an increase in stability with increasing diversity (Tilman et al. 2006, van Ruijven and Berendse 2007, Weigelt et al. 2008, Isbell et al. 2009, Hector et al. 2010, Roscher et al. 2011). Components of the CV that might be driving the relationship with biodiversity are often also assessed: mean-standard deviation scaling (van Ruijven and Berendse 2007), species covariances (Isbell et al. 2009), standard deviation and mean (Hector et al. 2010). However, the fundamental metric for stability remains the CV. Most recently, Gross et al. (2014), while using stability as defined by the CV, explicitly couch their approach to stability in the context of ‘deep implications for ecosystem management where sustainability and reduced risk are often primary goals’. The goals of sustainability and reduced risk do encompass stability as defined by the CV but extend far beyond that range and require many other approaches to assess the achievement of these goals. One of our concerns is that the narrow definition of stability implicit in the CV as the primary metric may confuse other biologists, ecosystem

managers and policymakers not so close to the detail of the stability debate.

There is a parallel and substantial literature in agronomy that addresses similar questions about stability. In agronomy, the concept of stability is applied through a focus on the selection and recommendation of cultivars or systems of production for use in different regions (Smith et al. 2005). For example, does one cropping system or cultivar show higher stability than another across geographical sites and do these effects persist over time? Answering this type of question is based on quantifying both the mean and variability of the desired function(s) (e.g. yield, weed invasion, pests or diseases) across multiple sites and times (e.g. Piepho et al. 2012, Mühleisen et al. 2014). Stability is defined in terms of these means and variability estimates and cultivar selection usually involves a trade-off between them. Both mean and variability may include separate components for space, time and environment in addition to the variation due to replication (for reviews, see Piepho 1998, Smith et al. 2001, Smith et al. 2005). Mean and variability are jointly modelled using standard statistical mixed-models methodology (e.g. Pinheiro and Bates 2000), which can incorporate structural sources of variation (time, space and replication) and estimate standard deviations across these.

The CV, as the ratio of standard deviation to mean functional response, conflates the separate patterns of response of mean and standard deviation in the face of temporal and environmental change in a single metric; a specific value of CV can arise as the result of many different influences on the two components of the ratio. It is not always clear how to link stability defined using this metric with the importance of diversity in providing reliable ecosystem services for society. Here, we contend that describing stability solely in terms of the Coefficient of Variation (CV) is often inadequate and can be misleading. We argue that an approach based on the CV is not appropriate for many of the questions that can arise in Biodiversity-Ecosystem-Function research, questions that could equally be seen as exploring the stability of ecosystem function. We discuss the inappropriateness of the approach based on the CV when low values of the ecosystem are preferred. We contrast the CV-based approach with an alternative based on the joint use of the mean response and its standard deviation (SD). The proposed alternative approach can be generalised to incorporate socio-economic and biological complexities (e.g. limiting thresholds) when appropriate.

Comparison of ecosystem function based on different methods: illustrative scenarios

Common goals of ecological and agronomic studies, as well as their common methodological challenges, are

reflected in the illustrations in Figure 1. Using hypothetical communities with varying mean and SD, we discuss possible outcomes, and interpretations about the stability of the ecosystem, when using either the CV or a mean and SD approach. We consider three scenarios. In all scenarios we compare monoculture communities A and B with AB which is an equal mixture of A and B. In scenario (ii) we introduce another hypothetical equal mixture AB' with different characteristics to AB. Figure 1 shows mean and SD assumptions for each community. We assume that the functional response is normally distributed for all communities.

- (i) We assume (Figure 1 (i)) that there is some overyielding, i.e. $\text{mean}(AB) > (\text{mean}(A) + \text{mean}(B)) / 2$. Then, based on comparison of the CV values, AB is more stable (*sensu* CV) than either of the monocultures (A, B). This conclusion agrees with the hypothesis that diversity increases stability (*sensu* CV).
- (ii) In a second scenario (Figure 1(ii)), there is considerable transgressive overyielding i.e. $\text{mean}(AB) > \max(\text{mean}(A), \text{mean}(B))$. Here, diversity does not increase stability (*sensu* CV). However, even with its higher SD and CV, the mean response of AB is more than 2SD greater than the mean performance of A or B and so most occurrences of AB would exceed the responses of A or B. Despite its greater CV, mixture AB might be regarded as better (a more desirable state or a more stable outcome by the users of the ecosystem function) than A or B because it consistently has a higher ecosystem function than either of them, even allowing for its greater instability. So, in this context, what is the insight from stability (*sensu* CV)?

AB' is an alternative, less clear scenario with an equal mixture of A and B that also displays transgressive overyielding but again with greater CV than A or B and so less stable (*sensu* CV). B is the better monoculture having lower CV than A and is doubly preferred to A as it also has a higher mean functional response. So, is B preferred to AB' due to its greater stability (lower CV) in this case? Using the mean-SD approach, 95% prediction intervals for single realisations of B and AB' are (16, 44) and (11, 59) respectively. Comparing the prediction intervals for B and AB' shows that many occurrences of AB' would be less than those from B. Here the comparison of AB' with B and judgement on which is better is not so clear. Any assessment of the relative performance of B and AB' should involve a trade-off between higher average ecosystem function and greater variability. Assessment based solely on the CV ignores the

obvious dilemmas in selecting one or other community in the belief that it is more stable.

In general, when comparing more diverse with less diverse mixtures and monocultures the following question will be relevant: does the increased mean functional response in more diverse mixture communities (usually desirable) compensate for their greater variability (usually undesirable)? Here, higher functioning comes at the cost of higher variability, and selection of one community over another depends on how risk-adverse the criteria for selection are. It is difficult to envisage a scenario in Fig. 1 (ii) in which A or B would be preferred to AB, but B might be preferred to AB' if it was vital to avoid even the occasional occurrence of a very low value of functional response. If the primary goal of ecosystem management is to avert the risk of low levels of ecosystem function then, despite its lower mean functional response, community B may be preferred to AB'. In contrast, community AB' would be chosen if the basis of selection was to choose communities with higher average levels of ecosystem function. It is not clear how stability (*sensu* CV) informs this discussion.

Questions in scenario (ii) are of central interest to ecologists in assessing whether biodiversity is an important element in the reliable provisioning of ecosystem services across space and/or time. Our main point here is that such decisions in scenario (ii) are not assisted by information on stability as measured by the CV of the functional response. These two scenarios suggest that the use of CV to define stability is associated with strong caveats that need to be considered if it is the only metric used. Indeed, we find it difficult in terms of scenario (ii) to identify an ecological question that is addressed conclusively by the CV.

- (iii) We introduce a third scenario (Figure 1 (iii)), in which low values of ecosystem function are preferable to high ones (e.g. weed yield, greenhouse gas emissions). Here, AB (the equal mixture) has a lower mean (desirable) and standard deviation (desirable) than either of A or B (the monocultures) but a higher CV than both of them. Here greater mean and greater variability are both undesirable and any sense of one compensating for the other is not present (as it could be in scenario (ii)). Selection of A or B over the mixture AB as being more stable (*sensu* CV) would be a totally incorrect choice in terms of any reasonable definition of stability. This example illustrates how decision-making based on CV alone could obscure rather than illuminate the issues.

Figure 1. Illustration of three scenarios discussed in text. All scenarios compare levels of ecosystem function for three hypothetical communities A, B and AB, with A and B monocultures and AB an equal mix of A and B. The top of each panel shows the assumed mean, SD and CV for each community. Scenario (ii) includes community AB', an equal mixture of A and B with alternative mean – SD properties to AB. Functional response in all communities is assumed to be normally distributed and each community is represented in the appropriate panel by its mean and 95% prediction interval. (i) Based on the value of CV, community (AB) is the most stable. (ii) AB is less stable than A or B (*sensu* CV). However, most functional responses from AB would exceed those from A or B and on that basis it might be preferred as more stable. Community AB' is less stable than A or B (*sensu* CV). Its mean response is greater than that of B but its prediction interval overlaps that of B at low levels; a decision between communities may depend on the importance of avoiding very low values of the functional response. (iii) Where low values of the functional response are desirable, the CV can be very misleading. The mixture AB is less stable (*sensu* CV) but comparison of the prediction intervals for the three communities shows that almost all examples of community AB would have a lower, preferred, level of functional response than either of the monocultures.

Comparison of ecosystem function based on threshold levels: illustrative scenarios

The scenarios discussed above point to a need for improved definition of the characteristics of ecosystem function that are used to define and measure 'stability of ecosystem function'. The stability of ecosystem function is also of practical significance to land use managers who wish to optimise the provision of ecosystem services. The choice of a stability metric should be appropriate to the context in which it is being used and may go beyond either the CV or a simple assessment of mean and variability. Here we outline one possible alternative, the comparison of ecosystems on the basis of the probability of the functional response reaching some threshold, perhaps combined with a penalty for crossing the threshold.

The idea of risk aversion naturally leads to the idea of comparing ecosystems on the basis of the probability of functional response exceeding or falling below threshold values. An ecosystem function may fluctuate within a range of levels that are considered to supply an acceptable level of ecosystem function. The acceptable limits can be defined by the presence or absence of tipping points or thresholds (Scheffer et al. 2001, Groffman et al. 2006); threshold levels may be upper or lower limits (or both) of a single (or multiple) function (Gamfeldt et al. 2008). Essentially, the setting of thresholds establishes a framework for the calculation of the risk of an ecosystem function reaching an unacceptable level such as, for example, the minimum level of required economic return on a cropping system (Mead et al. 1986), maximum levels of nutrient concentration in lake ecosystems (Zhang et al. 2003), or water quality levels of nitrate-loading in streams (Cardinale 2011). The limits and thresholds are defined in terms of a desired range of functioning. For an ecosystem manager, the metric of stability can be the probability of ecosystem function lying within the desired range. This probability will depend on both mean and standard deviation of the ecosystem function. Thresholds have been used in a recent study on multi-functionality as a basis for combining several functions to assess the effects of diversity (Byrnes et al. 2014), and this framework could be extended to assessing stability, simply through the incorporation of an estimation of the variation around the mean and hence of the probability for a given function to fall under (or over) a threshold.

We use three simple scenarios (Figure 2) to illustrate how the probability of exceeding a defined threshold level of ecosystem function will depend on both the mean and the SD of the function. We consider three communities (A, B and C) that have different mean and SD values for their provision of a normally distributed ecosystem function. We consider a threshold level of

Figure 2. Distribution of a measure of ecosystem function for three hypothetical communities for which the mean and SD differ: (A) mean = 10, SD = 2; (B) mean = 25, SD = 7; (C) mean = 30, SD = 10 with a threshold level of ecosystem function set at 20, below which values of ecosystem function are undesirable. Community A is 100% below the threshold, while B and C overlap the threshold to different degrees.

ecosystem function (= 20) that it is desirable to exceed. The shaded area under the curves in Figure 2 corresponds to the probability of a function falling below the defined threshold level and the unshaded area is the probability that a function is above the threshold and can be considered as the measure of stability of the system. Using the CV only, it is not possible to assign values to this probability; both the mean and the SD are necessary (Figure 2). In a comparison of either the CV or SD alone among communities A, B and C, community A has the lowest values and would be considered the most stable. However, this community is also 100% beneath the threshold, and thus cannot be considered adequate. When comparing communities B and C with the CV or the SD alone, B is more stable. However, its mean function is lower than that of C, and it is not possible to make further inference about the stability of these two communities (*sensu* probability of exceeding the threshold) without jointly considering the mean and the SD to calculate the probabilities that the functioning of each community exceeds the threshold level.

The threshold approach can readily become both more complex and more realistic if there is a penalty for not achieving the threshold. If this penalty changes strongly with the extent of failure to meet the threshold

then a definition of stability that incorporates both the probability of exceeding the threshold and the penalty for not achieving it might be more appropriate. Indeed a combination of the distribution of the functional response and a penalty function may well provide a measure of stability that is more directly attuned to the question being addressed, rather than a measure based on the CV. Thus, deciding which community is most stable may be influenced by judgements in setting threshold levels and/or the socio-economic consequences of not achieving certain levels of functional response. Closely related to this will be the need to assess the extent to which diversity, both among monocultures and across levels of richness (Schmid et al. 2008) can affect the variability of functional responses. This is readily achievable using standard statistical methods (Pinheiro and Bates 2000) and experimental designs that include monocultures and polycultures, with appropriate replication (Hector et al. 2009).

Discussion

Judgements on how stability is best measured require clarity about the specific questions to be resolved. A definition of ecosystem stability in any context should aim to resolve a specific question about ecosystem performance. A clear logical pathway from the definition of stability to the ability to discriminate between several outcomes is necessary. Our various scenarios and threshold examples illustrate the inadequacies of the definition based on the CV in making a sensible contribution to addressing several questions around the delivery of ecosystem services. We sketched out several scenarios, where alternative definitions of stability based on the mean ecosystem function, its variability and penalties for extreme responses could provide a more satisfactory framework for assessing ecosystem stability. We consider that defining the stability of ecosystem functioning solely in terms of the CV confines the term stability to a usage and context that is not natural or intuitive to many who wish to address questions about the reliability with which ecosystems deliver services (Campbell et al. 2011). The use of the CV to define stability where low values of the function are preferable is not appropriate.

In many ecological studies, the CV remains the metric of choice in assessing the stability of ecosystem function, but the rationale for its use is not always clear; indeed, some of the studies that rely on CV also recognise that its use requires qualification (Steudel et al. 2011). The limitations of the use of CV as a measure of stability need to be carefully considered, and the associated caveats more fully appreciated. Interpretation based on the CV is likely to be confounded by the separate effects of factors that simultaneously influence the numerator (standard deviation of functional res-

ponse) and the denominator (mean functional response) of the ratio. Thus, an observed difference in CV between two ecosystems is explainable in many different ways, and may be different because the mean is greater or the variation less or a combination of both.

Some of the arguments about the size of the CV are grounded in work on the dynamics of mixed-species communities (Tilman 1996, de Mazancourt et al. 2013). It is well accepted for many functional responses that positive interactions among species in mixtures can both increase the mean response over what might be expected by averaging across monocultures and may also influence interspecific dynamics so as to reduce the relative variation of functional response from mixtures as measured by the CV (Tilman et al. 1998). While the use of the CV arises naturally in models of systems dynamics, the assignment of the term 'stability' to its interpretation seems to rely more on historic usage rather than on any intrinsic argument (de Mazancourt et al. 2013).

In diversity experiments, jointly quantifying the effects of multiple factors such as space and time on the functional response and its variability for different levels of species diversity provides a starting point for an assessment of stability that would simply not be available with methods based on CV alone. The variability of ecosystem function may be affected by time, location, environment and diversity; this can be captured using mixed model approaches (Piepho 1997, Pinheiro and Bates 2000, Kirwan et al. 2007, Keith et al. 2008). This can lead to a range of means and SDs that vary across these factors. The flexibility of an approach based on these means and SDs that incorporates information across multiple scales and linked with socio-economic evaluation of risk avoidance (see for example proposition 1 in Baumgärtner 2007) offers the prospect of a more general framework within which to address questions on stability than has currently been utilised.

Acknowledgements

T. Carnus was supported by the Teagasc Walsh Fellowship scheme. A contribution to the research leading to these results was conducted as part of the Animal Change project (JAF and JC, grant agreement 266018), which received funding from the European Community's Seventh Framework Programme (FP7/2007–2013).

Referees

Jonathan S. Lefcheck – jslefche@vims.edu
Virginia Institute of Marine Science

Claire de Mazancourt –
Claire.demazancourt@ecoex-moulis.cnrs.fr

References

- Balvanera, P., Pfisterer, A., Buchmann, N., He, J.-S., Nakashizuka, T., Raffaelli, D. et al. 2006. Quantifying the evidence for biodiversity effects on ecosystem functioning and services. *Ecology Letters* 9: 1146–1156. [CrossRef](#)
- Baumgärtner, S. 2007. The insurance value of biodiversity in the provision of ecosystem services. *Natural Resource Modeling* 20:87–127. [CrossRef](#)
- Byrnes, J.E.K., Gamfeldt, L., Isbell, F., Lefcheck, J.S., Griffin, J.N., Hector, A. et al. 2014. Investigating the relationship between biodiversity and ecosystem multifunctionality: Challenges and solutions. *Methods in Ecology and Evolution* 5: 111–124. [CrossRef](#)
- Campbell, V., Murphy, G. and Romanuk, T. 2011. Experimental design and the outcome and interpretation of diversity–stability relations. *Oikos* 120: 399–408. [CrossRef](#)
- Cardinale, B. 2011. Biodiversity improves water quality through niche partitioning. *Nature* 472: 86–89. [CrossRef](#)
- Cardinale, B.J., Srivastava, D.S., Emmett Duffy, J., Wright, J.P., Downing, A.L., Sankaran, M. et al. 2006. Effects of biodiversity on the functioning of trophic groups and ecosystems. *Nature* 443: 989–992. [CrossRef](#)
- de Mazancourt, C., Isbell, F., Larocque, A., Berendse, F., De Luca, E., Grace, J.B. et al. 2013. Predicting ecosystem stability from community composition and biodiversity. *Ecology Letters* 16: 617–625. [CrossRef](#)
- Gamfeldt, L., Hillebrand, H. and Jonsson, P. 2008. Multiple functions increase the importance of biodiversity for overall ecosystem functioning. *Ecology* 89: 1223–1231. [CrossRef](#)
- Griffin, J.N., O’Gorman, E.J., Emmerson, M.C., Jenkins, S.R., Klein, A.M., Loreau, M. et al. 2009. Biodiversity and the stability of ecosystem functioning. In: Naeem, S., Bunker, D.E., Hector, A., Loreau, M. and Perrings, C. (eds.), *Biodiversity, ecosystem functioning, and human wellbeing: An ecological and economic perspective*. Oxford University Press, pp. 78–93. [CrossRef](#)
- Grimm, V. and Wissel, C. 1997. Babel, or the ecological stability discussions: An inventory and analysis of terminology and a guide for avoiding confusion. *Oecologia* 109: 323–334. [CrossRef](#)
- Groffman, P., Baron, J., Blett, T., Gold, A., Goodman, I., Gunderson, L. et al. 2006. Ecological thresholds: The key to successful environmental management or an important concept with no practical application? *Ecosystems* 9: 1–13. [CrossRef](#)
- Gross, K., Cardinale, B.J., Fox, J.W., Gonzalez, A., Loreau, M., Polley, H.W., et al. 2014. Species richness and the temporal stability of biomass production: A new analysis of recent biodiversity experiments. *The American Naturalist* 183:1–12. [CrossRef](#)
- Hector, A., Hautier, Y., Saner, P., Wacker, L., Bagchi, R., Joshi, J. et al. 2010. General stabilizing effects of plant diversity on grassland productivity through population asynchrony and overyielding. *Ecology* 91: 2213–2220. [CrossRef](#)
- Hooper, D.U., Chapin, F.S., Ewel, J.J., Hector, A., Inchausti, P., Lavorel, S. et al. 2005. Effects of biodiversity on ecosystem functioning: A consensus of current knowledge. *Ecology Monographs* 75: 3–35. [CrossRef](#)
- Isbell, F., Polley, W. and Wilsey, B. 2009. Biodiversity, productivity and the temporal stability of productivity: Patterns and processes. *Ecology Letters* 12: 443–451. [CrossRef](#)
- Ives, A. and Carpenter, S. 2007. Stability and diversity of ecosystems. *Science* 317: 58–62. [CrossRef](#)
- Keith, A.M., van der Wal, R., Brooker, R.W., Osler, G.H.R., Chapman, S.J., Burslem, D.F.R.P. et al. 2008. Increasing litter species richness reduces variability in a terrestrial decomposer system. *Ecology* 89: 2657–2664. [CrossRef](#)
- Kirwan, L., Luscher, A., Sebastia, M. T., Finn, J.A., Collins, R.P., Porqueddu, C. et al. 2007. Evenness drives consistent diversity effects in intensive grassland systems across 28 european sites. *Journal of Ecology* 95: 530–539. [CrossRef](#)
- Lehman, C. and Tilman, D. 2000. Biodiversity, stability, and productivity in competitive communities. *The American Naturalist* 156: 534–552. [CrossRef](#)
- McNaughton, S.J. 1977. Diversity and stability of ecological communities: A comment on the role of empiricism in ecology. *The American Naturalist* 111: 515–525. [CrossRef](#)
- Mead, R., Riley, J., Dear, K. and Singh, S.P. 1986. Stability comparison of intercropping and monocropping systems. *Biometrics* 42: 253–266. [CrossRef](#)
- Mühleisen, J., Piepho, H.P., Maurer, H.P., Longin, C.F. H., and J.C. Reif. 2014. Yield stability of hybrids versus lines in wheat, barley, and triticale. *Theoretical and Applied Genetics* 127:309–316. [CrossRef](#)
- Piepho, H.P. 1997. Analyzing genotype-environment data by mixed models with multiplicative terms. *Biometrics* 53: 761–766. [CrossRef](#)
- Piepho, H.P. 1998. Methods for comparing the yield stability of cropping systems. *Journal of Agronomy and Crop Science* 180: 193–213. [CrossRef](#)
- Piepho, H.P., Möhring, J., Schulz-Streck, T., and J.O. Ogutu. 2012. A stage-wise approach for the analysis

- of multi-environment trials. *Biometrical Journal* 54:844–860. [CrossRef](#)
- Pinheiro, J. and Bates, D. 2000. *Mixed-effects models in s and s-plus (statistics and computing)*. Springer.
- Roscher, C., Weigelt, A., Proulx, R., Marquard, E., Schumacher, J., Weisser, W. et al. 2011. Identifying population- and community-level mechanisms of diversity–stability relationships in experimental grasslands. *Journal of Ecology* 99: 1460–1469. [CrossRef](#)
- Sala, O., Stuart, C., Armesto, J., Berlow, E., Bloomfield, J., Dirzo, R. et al. 2000. Global biodiversity scenarios for the year 2100. *Science* 287: 1770–1774. [CrossRef](#)
- Scheffer, M., Carpenter, S., Foley, J.A., Folke, C. and Walker, B. 2001. Catastrophic shifts in ecosystems. *Nature* 413: 591–596. [CrossRef](#)
- Smith, A., Cullis, B., and R. Thompson. 2001. Analyzing variety by environment data using multiplicative mixed models and adjustments for spatial field trend. *Biometrics* 57:1138–1147. [CrossRef](#)
- Smith, A.B., Cullis, B.R., and R. Thompson. 2005. The analysis of crop cultivar breeding and evaluation trials: an overview of current mixed model approaches. *The Journal of Agricultural Science* 143:449–462. [CrossRef](#)
- Steudel, B., Hautier, Y., Hector, A. and Kessler, M. 2011. Diverse marsh plant communities are more consistently productive across a range of different environmental conditions through functional complementarity. *Journal of Applied Ecology* 48: 1117–1124. [CrossRef](#)
- Tilman, D. 1996. Biodiversity: Population versus ecosystem stability. *Ecology* 77: 350–363. [CrossRef](#)
- Tilman, D. 1999. The ecological consequences of changes in biodiversity: A search for general principles. *Ecology* 80: 1455–1474.
- Tilman, D., Lehman, C. and Bristow, C. 1998. Diversity-stability relationships: Statistical inevitability or ecological consequence? *American Naturalist* 151: 277–282. [CrossRef](#)
- Tilman, D., Reich, P. and Knops, J. 2006. Biodiversity and ecosystem stability in a decade-long grassland experiment. *Nature* 441: 629–632. [CrossRef](#)
- van Ruijven, J. and Berendse, F. 2007. Contrasting effects of diversity on the temporal stability of plant populations. *Oikos* 116: 1323–1330. [CrossRef](#)
- Weigelt, A., Schumacher, J., Roscher, C. and Schmid, B. 2008. Does biodiversity increase spatial stability in plant community biomass? *Ecology Letters* 11: 338–347. [CrossRef](#)
- Zhang, J., Jørgensen, S., Beklioglu, M. and Ince, O. 2003. Hysteresis in vegetation shift—lake morgan prognoses. *Ecological Modelling* 164: 227–238. [CrossRef](#)