‘Bigger data’ on scale-dependent effects of invasive species on biodiversity cannot overcome confounded analyses: a comment on Stohlgren & Rejmánek (2014)

Jonathan M. Chase1, Kristin I. Powell2 and Tiffany M. Knight1,3

1German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig 04103, Germany
2Center for Tropical Forest Science—Forest Global Earth Observatory, Smithsonian Tropical Research Institute, Washington DC 20560, USA
3Department of Biology, Washington University in Saint Louis, Saint Louis, MO 63130, USA

A recent study by Stohlgren & Rejmánek [1] (SR) recently attempted to evaluate a general pattern suggested by Powell et al. [2] (PCK) that invasive plant species have large effects on species richness at small spatial scales and smaller effects at larger spatial scales, leading to predictable shifts in the parameters of the species–area relationship (SAR). To do so, SR collated data from hundreds of SARs and found little correlation between SAR parameters and invasive plant cover, which they took to indicate that the results observed in PCK were not general. We discuss several fundamental flaws in SR’s approach.

First, contrary to SR’s claims, the generalities suggested by PCK were not based only on data from three sites, but emerged from a synthesis of well-supported results across scales. Hundreds of studies and meta-analyses have shown that at small spatial scales, dense populations of invasive species strongly reduce native species richness (e.g. [3]). At regional scales, invasive plants rarely drive native species extinct (e.g. [4]). Because invaders reduce local, but not regional diversity, it necessarily follows that, in general, invasive plants reduce the intercept, but increase the slope of the SAR. The lack of relationship in SR’s results is at odds with this voluminous literature.

Second, more than 70% of the sites in each dataset used by SR had less than 5% cover of invaders and less than 5% of sites had more than 50% cover of invaders. Such extreme bias in the dataset would obscure any significant relationships, should they exist, especially because invasive species effects are known to be nonlinear (e.g. the effects of invaders are usually negligible until they constitute at least 50% cover; [5]).

Third, the correlative analyses by SR do not allow them to disentangle the causal influence of invasive species owing to indiscriminant compilations of study sites with confounding conditions. On the contrary, PCK compared paired invaded and uninvaded sites to disentangle confounding factors...
Three of SR’s most incongruous assumptions are that: (1) invasive cover varies independently from environmental variation; (2) environmental variation has minimal influence on SAR parameters; and (3) the causal relationship between invader cover and species richness is unidirectional. Instead, environmental factors (e.g. productivity and disturbance) are a primary influence of invasive plant cover (e.g. [6]) (contra no. 1), as well as SAR parameters (e.g. [7]) (contra no. 2), and there is a strong bidirectionality between invader cover and native plant richness (e.g. [8]) (contra no. 3).

As an example, we find that an estimate of site productivity is positively correlated with the intercept and negatively correlated with the slope of the SAR (figure 2), rendering it impossible to correctly detect an effect of invaders without controlling for the influence of site productivity. Indeed, a subset of the same data was used to show that productivity was positively correlated with both native species richness and invader cover [9], emphasizing SR’s inappropriate inference of causality.

\[ Y = -0.44x + 0.311, \ p < 0.001 \]
\[ Y = -0.386x + 0.160, \ p < 0.001 \]
\[ Y = -0.118x + 0.426, \ p < 0.001 \]
\[ Y = -0.476x + 0.021, \ p < 0.001 \]

**Figure 1.** Overview of the factors that influence SAR parameters, \( c \) (intercept) and \( z \) (slope). Black arrows indicate the causal effects of invaders on the SAR, while grey arrows are just a few of the many confounding variables that were controlled using the paired design by PCK, but uncontrolled by SR.

**Figure 2.** Relationship between a proxy for productivity (log\(_{10}\) per cent plant cover) and the slope and intercept of the SAR from the Modified Whittaker vegetation plots (a,b) and the USDA Forest Health Monitoring plots (c,d) from the data provided by SR. (Online version in colour.)
While we acknowledge that the paired design by PCK cannot disentangle all potential confounds, it is the most rigorous approach available in lieu of long-term experiments.

In conclusion, we agree with SR that much will be gained by broadening the scope of research on invasive species effects on SAR parameters by including data from more sites with variability in invader abundances and diversity. To properly analyse such data, however, would require a more sophisticated analysis that explicitly examines multiple causality (e.g. path analysis and structural equation modelling) (e.g. [10]). At their best, synthesized datasets like those used by SR can provide important insights into generality and nuance, but at their worst, impede advancement in the field.

References

2. Powell KI, Chase JM, Knight TM. 2013 Invasive plants have scale-dependent effects on diversity by altering species–area relationships. Science 339, 316 – 318. (doi:10.1126/science.1226817)