Oecologia (1997) 110:557–559

Peter M. Vitousek

On regression and residuals: response to Knops et al.

Received: 16 December 1996 / Accepted: 16 December 1996

Knops, Koenig and Nash (1997; hereafter KKN) conclude that (1) my earlier work demonstrating that forests in low-nutrient sites produce litter with wide dry mass/nutrient ratios (low nutrient concentrations) (Vitousek 1982) is flawed, and (2) a reanalysis using appropriate statistical procedures yields no relationship between soil fertility and dry mass/nutrient ratios in litterfall. Their analysis is based on a misinterpretation of simple linear regression; once this is corrected, it is as clear in their analysis as it was in mine that forests in low-nutrient sites produce litter with low nutrient concentrations.

I believe it is fair to summarize a main line of argument in KKN as follows.

(1) In 1982, I suggested that the inverse of element concentrations in litterfall could be used as an index of nutrient use efficiency in forest ecosystems, and that the quantity of a biologically important nutrient circulated through litterfall could be used as an index of availability of that nutrient (Vitousek 1982). I then put together a database from the literature on nutrient cycling in forests, and showed that the ratio of dry mass to nutrients in litterfall is systematically greater in sites circulating small quantities of nutrients through litterfall. Finally, I speculated that higher mass-to-element ratios in litterfall in low-nutrient sites could lead to slower rates of decomposition and nutrient mineralization, thereby driving a positive feedback from low nutrient availability to efficient nutrient use by plants to slow decomposition to low nutrient availability.

(2) KKN suggest that my analysis was flawed because plots showing the quantity of nutrient circulated on the *x*-axis and dry mass divided by the quantity of nutrient circulated on the *y*-axis are autocorrelated; they automatically yield an apparent pattern of increased dry mass/element ratios in low-nutrient sites, even for randomly generated data.

(3) If the analysis of how dry mass per unit nutrient varies with nutrient supply is done correctly, there is "no indication that nutrient use efficiency is greater in nutrient-poor ecosystems" (Knops et al. 1997).

The remainder of KKN includes some speculation about mechanisms underlying the lack of a relationship between soil fertility and nutrient use efficiency – speculation that is useful only if the two are indeed unrelated.

The description in point 1 above is exactly what I did. Were I to do it over 15 years later, I would use the conceptually richer views of nutrient use efficiency that have been developed by Shaver and Melillo (1984), Berendse and Aerts (1987) and others. More importantly, I would draw upon the large number of recent studies in which nutrient availability (or some soil-based index of it) has been measured directly and then compared with nutrient concentrations in plants or litterfall.

Point 2 is right concerning the pitfalls of autocorrelation – but KKN are wrong that "this problem is more severe than previous investigators have assumed." In 1982, I said "correlation of the axes could cause a relationship to emerge because the *y*-axis... reduces to litterfall mass divided by the *x*-axis. Consequently, an inverse relationship...would be expected if there were no functional relationship between nutrient circulation and litterfall mass." For that reason, all of my statistical tests were based on analyses of plots of litterfall dry mass against quantities of nutrients – just as those of KKN are.

Point 3 is the key one. KKN use subsets of my dataset, and a larger one from Vogt et al. (1986), to analyze the relationship between litterfall dry mass and nutrients in three ways – using simple linear regressions (their preferred approach), linear regressions forced through the origin and polynomial regressions (which are not evaluated statistically and which I will not discuss here). For the simple linear regressions, a single example might

---

P.M. Vitousek
Department of Biological Sciences, Stanford University,
Stanford, CA 94305, USA
fax: 415-725-1856; e-mail: vitousek@leland.stanford.edu

involve litterfall dry mass (dependent variable) regressed on the quantity of N in litterfall (independent variable). After calculating the regression, KKN test to see if the residuals from that regression are correlated with N, the original independent variable. They invariably find no significant correlation, and therefore reject the hypothesis that N use efficiency is greater in low sites. However, residuals are defined as the variation *not* explained by the original independent variable; if the regression is done correctly, they cannot be correlated with that variable. It is useful to plot the distribution of residuals against the independent variable (to look for non-linearity, to examine assumptions about variance). However, to interpret lack of correlation between the residuals and the original independent variable as if that represented hypothesis testing is to betray a fundamental misunderstanding of regression and residual analysis.

If you want to test whether there is significantly more dry mass per unit of N in low-N sites, you should evaluate the *y*-intercept of litterfall dry mass regressed against litterfall N, rather than the residuals of this linear regression. KKN suggest that their simple linear regressions theoretically should pass through the origin – as indeed they should, *if* there is no systematic variation in the quantity of dry mass per unit of nutrient. However, they find a (generally very highly) significant positive *y*-intercept in every case involving comparisons across sites. Assuming that linear regression is the appropriate model, these positive *y*-intercepts establish that there is systematically more dry mass per unit of nutrient in litterfall in the low-nutrient sites. That is the implication of a positive *y*-intercept in a simple linear regression, and KKN's analysis shows that is the way it is in the world.

KKN then calculate linear regressions forced through the origin, as I did in 1982. These regressions provide a clear statement of the null hypothesis that there is no systematic variation in the quantity of litterfall dry mass per unit of nutrient, as a function of quantities of nutrients in litterfall. If the distribution of data points deviates systematically and significantly from this forced regression line, the null hypothesis must be rejected. Both KKN and I show that points with low nutrients (at least low N, and often P) are systematically and significantly above the line, while those with high nutrients are below it; their residual-based approach is more straightforward than my polynomial-based approach. KKN dismiss this very clear result by stating that the positive *y*-intercepts observed in simple linear regressions result by default in an artifactual negative relationship between the residuals and the variable of interest in no-intercept regressions. There is, however, nothing artifactual about either the positive *y*-intercepts themselves, or the consequent negative relationships between nutrients and residuals of the no-intercept linear regressions. Both occur *because* forest ecosystems systematically produce more litterfall dry mass per unit of nutrient in sites with less aboveground nutrient circulation.

It is reasonable to ask if this well-established pattern represents something we really want to know. I think my use of indices of nutrient availability and nutrient use efficiency was appropriate in 1982; it allowed me to employ large quantities of data (collected for other purposes) to test patterns in forest nutrient cycling and use. However, the more fundamental question is how tissue and litter chemistry, and components of nutrient use efficiency, vary as a function of soil fertility. There are now many more studies than in 1982 in which soil fertility or nutrient availability have been measured directly across a range of sites, with simultaneous assessments of plant and/or litter chemistry (cf. Pastor et al. 1984; Silver 1994; Bridgham et al. 1995; Crews et al. 1995). There are also many studies in which nutrient availability has been manipulated by fertilization and the consequences for plant and/or litter chemistry assessed (cf. Miller et al. 1976; Birk and Vitousek 1986; Berg and Tamm 1991; Aber et al. 1995; Vitousek et al. 1995). Analyses of these measurements and experiments are more straightforward than are attempts to infer nutrient availability from plant or litter nutrients, and they do not face the perils of autocorrelation. Analyses of both fertility gradients and fertilizer experiments generally show that forests with low levels of nutrient supply systematically produce plant tissue litter with lower concentrations of nutrients (higher mass-to-element ratios). The relationship between tissue chemistry and nutrient use efficiency is more complex than I recognized in 1982 (Berendse and Aerts 1987; Bridgham et al. 1995), the mechanisms involved are interesting and important (Aerts and de Caluwe 1994), and good work remains to be done on the implications of this pattern for plant-soil-microbial feedbacks. However, the underlying pattern of low nutrient concentrations in low-nutrient sites appears solidly based experimentally as well as comparatively.

Acknowledgements I thank D.U. Hooper for comments on earlier drafts of this response.

References

Aber JD, Magill A, McNulty SG, Boone RD, Nadlehoffer KJ, Downs M, Hallett R (1995) Forest biogeochemistry and primary production altered by nitrogen saturation. Water Air Soil Pollut 85: 1665–1670

Aerts R, Caluwe H de (1994) Nitrogen use efficiency of *Carex* species from heathlands. Oecologia 84: 391–397

Berendse F, Aerts R (1987) Nitrogen-use-efficiency: a biological meaningful definition? Funct Ecol 1: 293–296

Berg B, Tamm CO (1991) Decomposition and nutrient dynamics of litter in longterm optimum nutrition experiments. I. Organic matter decomposition in *Picea abies* litter. Scand J For Res 6: 305–321

Birk EM, Vitousek PM (1986) Nitrogen availability and nitrogen use efficiency in loblolly pine strands. Ecology 67: 69–79

Bridgham SD, Pastor J, McClaugherty CA, Richardson CJ (1995) Nutrient-use efficiency: A litterfall index, a model, and a test along a nutrient-availability gradient in North Carolina peatlands. Am Nat 145: 1–21

Crews TE, Kitayama K, Fownes JH, Riley RH, Herbert DA, Muller-Dombois D, Vitousek PM (1995) Changes in soil phosphorus fractions and ecosystem dynamics across a long chronosequence in Hawaii. Ecology 76: 1407–1424

Knops JMH, Koenig WD, Nash TH III (1997) On the relationship between nutrient use efficiency and fertility in forest ecosystems. Oecologia 110: 550–556

Miller HG, Cooper JM, Miller JD (1976) Effects of nitrogen supply on nutrients in litterfall and crown leaching in a stand of Corsican pine. J Appl Ecol 13: 233–248

Pastor J, Aber JD, McClaugherty CA, Melillo JM (1984) Aboveground production and N and P cycling along a nitrogen mineralization gradient on Blackhawk Island, Wisconsin. Ecology 65: 256–268

Shaver GR, Melillo JM (1984) Nutrient budgets of marsh plants: efficiency concepts and relation to availability. Ecology 65: 1491–1510

Silver WL (1994) Is nutrient availability related to plant nutrient use in humid tropical forest? Oecologia 98: 336–343

Vitousek PM (1982) Nutrient cycling and nutrient use efficiency. Am Nat 119: 553–572

Vitousek PM, Gerrish G, Turner DR, Walker LR, Mueller-Dombois D (1995) Litterfall and nutrient cycling in four Hawaiian montane rainforests. J Trop Ecol 11: 189–203

Vogt KA, Grier CC, Vogt DJ (1986) Production, turnover, and nutrient dynamics of above- and below-ground detritus of world forests. Adv Ecol Res 15: 303–377