Hydrol. Earth Syst. Sci. Discuss., 6, S267–S270, 2009
www.hydrol-earth-syst-sci-discuss.net/6/S267/2009/

Hydrology and Earth System Sciences Discussions

Interactive comment on "Modeling nutrient in-stream processes at the watershed scale using Nutrient Spiralling metrics" *by* R. Marcé and J. Armengol

M. Rafa

rafamarce@ub.edu

Received and published: 10 March 2009

Although we will answer all referees' concerns in our final response, we find very interesting to discuss now the main concern expressed by 'Anonymous Referee #1': a possible spurious correlation between Sw and Q in our work.

First, we want to thank the referee for constructive comments, and also for the reference by Wollheim. The contents in this paper and references herein will add a lot to the discussion of our work.

The referee argues that calculating a correlation between Q (streamflow) and Sw (nutrient uptake length) is misleading because we can define the variables as:

Sw = u/kc

and

Q = uA

where u is water velocity, kc is a nutrient uptake rate coefficient, and A is river cross-sectional area. Since both Sw and Q share a common component in their definitions (u), the referee considers correlation between Sw and Q is spurious. Thus, the referee suggests that discussions grounded on Q vs. Sw correlations (Figure 7 in the paper) are worthless.

However, we cannot agree with this view for two reasons:

1. Sw and Q values used to construct correlations in Figure 7 are MEASURED in the field, not computed from 'u' and the rest, and the field methods to measure these variables are independent. It is true that the MATHEMATICAL formulation of these variables indicate that some relationship should exist between them (if field work is accurately done), and actually we mentioned this in the paper (page 517, line 20).

2. However, the above reasoning is not the fundamental basis to reject the idea that our correlations are spurious. Even if Sw and Q are calculated using known 'u' values, correlations could still be statistically valid and useful. The reason is that the referee is using a wrong reasoning about spurious correlation in ecology that was already detected by others (Prairie and Bird, 1989). Although we strongly recommend this paper for a full understanding of our point, we will try to explain in brief the referee's mistake.

Quoting Prairie and Bird:

'...the claim that the correlation between variables sharing a common term is spurious is a pervasive and unfortunate misconception within the ecological literature. The correlation between such composite variables is always legitimate provided: 1) they satisfy the assumptions of correlation analysis, 2) the variables are meaningful, that is,

HESSD

6, S267–S270, 2009

they represent the concepts of interest and not just a component of them, and 3) the variables do not share a large measurement error term'.

In short, correlation analysis is compromised only in the case composite variables are rates or standardized variables, but we interpret results of the correlation regarding the untransformed variables. There are several beautiful examples in Prairie's paper. For example:

'For example, to consider the correlation between ratios to be representative of the relationship between numerators alone is simply faulty reasoning and this represents a case of spurious inference. An example of spurious inference in ecology is the interpretation of the relationship between bacterial abundance and organic matter content in lake sediments (e.g. Rublee 1982). Both bacterial numbers and organic matter are expressed per unit dry weight of sediment and therefore produce ratio variables sharing a common denominator. The difficulty here is that the relationship was thought to express an association between bacteria and organic matter alone without regard to other sediment characteristics. However, the hypothesis had been mathematically formulated in terms involving the amount of sediments. The correlation is not spurious, only its interpretation was. In the preceding terminology, the flaw here was that these ratios were not the variables of interest, only the numerators were'.

Obviously, this is not our case, since we are interested in Sw (and not kc) and Q (not A). Moreover, although we can expect some relationship between Sw and Q due their definitions, it is impossible to derive the results of the correlation before calculating it. There is nothing wrong in calculating a correlation between these variables, and we have to take in mind that we are COMPARING different correlations. That is, we are not claiming that the existence of this correlation is a great discover (first, it was described earlier elsewhere, and second, we already note in the paper that such correlation is an a priori expected result). We do claim that there exist a DIFFERENT relationship between Sw and Q in pristine and impaired streams.

HESSD

6, S267–S270, 2009

Interactive Comment

Finally, since field procedures to calculate Q and Sw are independent, we do not see any reason to think that these variables could share a high common measurement error term. For example, if we calculate Sw and Q from water velocities measured with great error, correlation could be artificially enhanced. However, most frequently Sw is calculated from nutrient concentration decline along a reach.

All in all, we think that correlations between Q and Sw are totally acceptable and very informative (in fact, these graphs have been used in many influential contributions).

We hope 'Anonymous Referee #1' will find our explanation convincing.

Rafa Marcé

Pairie I. and Bird D. 1989. Some misconceptions about the spurious correlation problem in the ecological literature . Oecologia 81: 285-288

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 6, 501, 2009.