

Response to Comment on “Nature of Phosphorus Limitation in the Ultraoligotrophic Eastern Mediterranean”

T. F. Thingstad,^{1*} C. S. Law,^{2,8} M. D. Krom,³ R. F. C. Mantoura,^{2,9} P. Pitta,⁴ S. Psarra,⁴ F. Rassoulzadegan,⁵ T. Tanaka,^{1,5} P. Wassmann,⁶ C. Wexels Riser,⁶ T. Zohary⁷

With no requirement for synoptic treated (IN) and control (OUT) stations, analysis of covariance is an interesting statistical technique for testing IN-OUT differences in Lagrangian experiments, but it has inherent limitations due to its assumption of linear responses. With this limitation properly considered, we find that analysis of covariance strengthens, not weakens, experimental support for the food-web transfer mechanisms we proposed.

Hypothesis testing in mesoscale Lagrangian experiments can be separated into two decision problems: (i) to identify whether the IN water mass is different from the OUT water mass(es) (a problem we refer to as DP1) and (ii) to identify whether this difference is caused by the treatment (DP2). We regard our experiment (1) as primarily hypothesis generating rather than hypothesis testing but acknowledge that use of analysis of variance (ANOVA) for DP1 does not properly take into account the evolution over time. We therefore found the suggestion of Hale and Rivkin (2) to use analysis of covariance (ANCOVA) to separate out the time effect to be most interesting. Carrying out the analysis, however, we reach conclusions different from those in (2). To understand why, consider the three hypothetical responses illustrated in Fig. 1. An ANCOVA with a linear regression for each of the two categories (IN and OUT) against time as the continuous predictor permits testing for significant differences in the slope and mean levels (and intercepts) for the two categories. For responses increasing with time (type I), the slopes of the two regression lines would be different, allowing a test for significant effects of time. Also the means would differ in type I cases. For bell-shaped responses (type B), one would still expect IN-OUT differences in means. The slopes may, however, be similar. Finally, with oscillatory response patterns (type O), neither

slopes nor means (nor intercepts) need to differ. ANCOVA is thus not a suitable test for IN-OUT differences in type O responses.

Hale and Rivkin (2) focus only on significant differences between slopes. This may be adequate in some of the iron experiments with long-lasting responses, many of which possibly are classified as type I. In the Cycling of Phosphorus in the Eastern Mediterranean (CYCLOPS) experiment, however, our interpretation was that the IN system returned to the original P-deficient state after about 7 days because of dilution, with sampling continuing to day 9 for most parameters (1). Responses were therefore mostly of type B. It should thus come as no surprise that ANCOVA (3) does not give significant differences in slopes for several of our variables

(Table 1). The exception is bacterial production, for which we got a significant slope. Bacterial production was in our case technically closer to a type I response because there was only a single measurement point (day 8) in the period after the return to P deficiency. However, for three of the four measured parameters in Table 1, we found significant differences in means ($P < 0.05$), as would be expected for type I and type B responses (for V^{\max} , $P = 0.055$). Contrary to Hale and Rivkin, we therefore conclude that use of ANCOVA based on differences in means supports, rather than weakens, our claim for IN-OUT differences in chlorophyll, particulate P, and bacterial production (1). A special case is the V^{\max} parameter, which appeared to have an overshoot in the presumed transition phase between N and P deficiency (1, 4), giving it more of a type O response, not well suited for ANCOVA. For this particular parameter, we actually have a replicate experiment. In a pilot experiment in the same region 1 year earlier (May 2001), we observed a drop in V^{\max} inside the patch, similar to that reported for the 2002 experiment (1). V^{\max} is thus a parameter where reproducibility supports our belief that the difference observed indeed is a treatment effect (DP2).

For copepod egg abundance, for which there was only one OUT sampling point, we cannot use statistical arguments and instead present as the simplest explanation that a one-order-of-magnitude increase shortly after phosphate addition is a treatment effect. For DP2, we otherwise primarily base our arguments on other supportive lines of reasoning. One is that

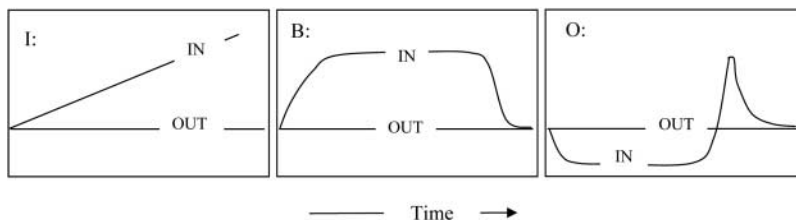


Fig. 1. Three hypothetical response patterns, increasing (I), bell shaped (B), and oscillatory (O) in treated (IN) water masses, combined with an assumed steady state in control (OUT) water masses of a Lagrangian experiment. The suitability of ANCOVA to test for significant IN-OUT differences in each of the three cases is discussed in the text.

Table 1. ANCOVA (2) for data used by Thingstad *et al.* (1) as a basis for generating hypotheses about P transfer in the pelagic food web. IN and OUT used as categorical separator and sampling day as continuous separator. Classification in type I, B, and O responses is based on the observation reported in (1) and the modeling results reported in (3). *, statistically significant differences in slopes or means ($P < 0.05$).

Variable	Slopes		Means		N_{OUT}	N_{IN}	Response type
	F	P	F	P			
Chlorophyll	0.144	0.71	5.864	0.029*	7	11	B (inverted)
Bacterial production	7.71	0.020*	5.237	0.045*	8	5	I
Particulate P	0.005	0.94	29.193	0.000121*	11	5	B
V^{\max}	0.716	0.41	4.528	0.055	10	5	O

¹Department of Biology, University of Bergen, Bergen, Norway. ²Plymouth Marine Laboratory, Plymouth, UK. ³School of Earth and Environmental Sciences and Earth and Biosphere Institute, Leeds University, Leeds, UK. ⁴Hellenic Centre for Marine Research, Heraklion, Crete, Greece. ⁵Station Zoologique, Villefranche-sur-mer, France. ⁶Norwegian College of Fishery Science, University of Tromsø, Tromsø, Norway. ⁷Israel Oceanographic and Limnological Research, Kinneret Limnological Laboratory, Tiberias, Israel. ⁸National Institute of Water and Atmospheric Research, Wellington, New Zealand. ⁹Marine Environment Laboratory, International Atomic Energy Agency, Monaco.

*To whom correspondence should be addressed. E-mail: frede.thingstad@bio.uib.no

all the basic biological mechanisms required for the proposed bypass and tunneling pathways are consistent with the literature (1), another that assembling these well-documented biological mechanisms in an idealized numerical food-web model allowed us to simulate the (unexpected) microbial responses observed (4). We could thus demonstrate the consistency of the observed system-level behavior with accepted cell-level mechanisms and form a coherent logical basis for the phosphate transfer mechanisms proposed (1).

Replication and independence are less straightforward concepts in mesoscale experiments than in laboratory bottles. Replicate patches would likely experience different physical environ-

ments leading to different dilution histories (5), thus creating a variance component from physics easily blurring the biological signal in exploratory, hypothesis-generating experiments such as ours. One could also argue that repeated samplings within one fertilized patch are samplings of different water parcels with limited interconnectivity and thus represent truly independent replicates, at least when restricting the question to that of IN-OUT differences.

With proper consideration of the linearity assumption in an ANCOVA analysis, we found it a useful statistical tool and conclude that it supports, not weakens, our experimental evidence for IN-OUT differences and thus further justifies our use of the experimental results as a

platform for generating hypotheses concerning trophic pathways for P transfer (1).

References and Notes

1. T. F. Thingstad *et al.*, *Science* **309**, 1068 (2005).
2. M. S. Hale, R. B. Rivkin, *Science* **312**, 1748 (2006); www.sciencemag.org/cgi/content/full/312/5781/1748c.
3. Our ANCOVA was done using the STATISTICA software (StartSoft, Inc., Tulsa, OK, USA) with log transformation of all data. For each station, data are averaged for samples taken at depths <20 m.
4. T. F. Thingstad, *Deep-Sea Res. II* **52**, 3074 (2005).
5. E. R. Abraham *et al.*, *Nature* **407**, 727 (2000).
6. We thank V. Vandvik, J. Birks, T. Miki, and R. Primicerio for comments and discussions on statistical aspects.

10 March 2006; accepted 25 May 2006
10.1126/science.1126408