

The Sediment Quality Guideline, ERL, is not a chemical concentration at the threshold of sediment toxicity

Thomas P. O'Connor
NOAA N/SCI1
National Centers for Coastal Ocean Science
1305 East West Highway, Silver Spring, MD 20910
tom.oconnor@noaa.gov

Abstract

While it is being used as such, the sediment quality guideline ERL (Effects Range Low) is not a threshold of any chemical concentration in sediment at which the probability of toxicity shows an abrupt increase. Similarly, while it has been done, there is no basis for assuming that multiple concentrations above an ERL increase the probability of toxicity.

Introduction

There are several instances (Van Dolah et al., 1999; EPA, 2001a; SFEI, 2003; and Kiddon et al., 2003) where authors have attributed undue biological significance to the sediment quality guideline called the Effects Range Low (ERL) (Long and Morgan, 1990). The point of this short note is to clarify that an ERL is not a threshold. Rather, an ERL is simply a low point on of a continuum of bulk chemical concentrations in sediment that roughly relate to sediment toxicity. For illustrative purposes I use concentrations of Cu, though any chemical with an ERL could serve as well.

Connecting sediment chemistry with toxicity

The ERL and ERM (Effects Range Median) are the 10th and 50th percentiles, respectively, on an ordered list of concentrations in sediment found in the literature that co-occur with any biological effect. For Cu, the original ERL (Long and Morgan, 1990) was 70 ppm because 51 concentrations were found in association with an effect and 7 were less than 70 ppm (to be precise that ERL was the 14th percentile but 3 of the lower 7 concentrations were almost identical at 67, 68.2 and 68.4 ppm). Long et al. (1995) reassessed ERLs and ERMs with a larger dataset that, for Cu, contained 74 concentrations associated with effects. The ERL for Cu dropped to 34 ppm, with 6 lower concentrations also co-occurring with effects. Long and Morgan (1990) originally defined an ERL as the upper end of a range of concentrations where effects rarely would be observed. Later, Long et al. (1995) refined the definition to “concentrations below the ERL value represent a minimal-effects range.....in which effects would be rarely observed. Concentrations equal to and above the ERL, but below the ERM represent a possible-effects range...” This latter range was quite large, for example, the 1995 ERL and ERM for Cu were 34 and 270 ppm, respectively. The mathematic definitions of ERLs and ERMs do not convey any predictive value. They are particular percentiles over a range of concentrations found to co-occur with toxicity but there is an overlapping concentration

range without co-occurring effects. Neither the ERL nor the ERM predicts the likelihood that a particular concentration will or will not be found with a co-occurring effect

Since 1995, the predictive powers of ERLs and ERMs have been tested against large datasets obtained in EPA and NOAA monitoring programs containing simultaneously measured chemical and toxicological parameters. O'Connor et al. (1998) compiled 1508 records with sediment chemistry and 10-day amphipod toxicity measurements. Among those 1508 records, 239 had at least one concentration above an ERM and 38% (90) of those were toxic. There were 481 samples without an ERL exceedance and 5% (26) were toxic. With additional data O'Connor and Paul (2000) expanded the dataset to 2475 records and the results were basically unchanged. Among the 453 samples with an ERM exceedance, 41% (186) were toxic and of the 730 samples without an ERL exceedance 5% (697) were toxic. A 2001 data compilation (EPA, 2001b) contains a total of 2761 records from EPA and NOAA and local programs that measured both amphipod toxicity and sediment chemistry. Eight percent of the 919 samples without an ERL exceedance were toxic. So the main conclusion has been consistent with the original definition of an ERL, i.e. "... a concentration below which effects are rarely observed," except that now "rarely" means observed with a 5-8% frequency. The numerical similarity between the 5-8% predictive ability of ERLs and their being 10th percentiles is purely coincidence, but it has become somewhat of a rule of thumb that ERLs correspond to a 10 percent probability of toxicity.

Field et al. (2002) took a major step forward when they determined a continuous relationship between concentration and toxicity such that

$$\ln(p/1-p) = B_0 + B_1(\log(x))$$

where p is the probability of toxicity and $B_0 + B_1$ are coefficients derived from the logistic regression. For Cu, $B_0 = -5.79$ and $B_1 = 2.93$ based on a dataset with 2580 samples with simultaneous measurements of Cu concentration and amphipod toxicity.

After deriving the coefficients for ten elements and 27 organic compounds or compound groups Field et al. (2002) determined that the probability of toxicity in samples with many chemicals (P(mixture)) was best estimated by the element or compound with maximum probability as:

$$P(\text{mixture}) = 0.11 + 0.33 p_{\max} + 0.4 p_{\max}^2$$

So, for example, if a sediment sample had a Cu concentration of 200 ppm(dry), the probability of toxicity based on Cu would be calculated as

$$\ln(p/1-p) = -5.79 + 2.93(\log(200)) = 0.952, \text{ so } p = 0.72$$

If among all the chemicals in that sample, the p based on Cu was the p_{\max} , then the actual P for the mixture would be

$$P(\text{mixture})=0.11 +0.33(0.72) + 0.4(0.72)^2 = 0.55$$

If the Cu concentration were 34 ppm (the ERL) and the corresponding $p (=0.21)$ were the p_{max} , the calculated $P(\text{mixture})$ would be 0.20

The assignment of higher probability of toxicity to an ERL than was found in previous examinations of large datasets may be due to the fact that Field et al.(2002) used a less stringent definition of toxicity. They declared a sediment sample to be toxic if there was less than 90% survival over ten days. All prior discussions defined toxicity as less than 80% survival, a value determined by Mearns et al. (1986) who found that cutoffs greater than 80% led to inconsistent results in an interlaboratory test of toxicity on common samples.

For the discussion at hand, the important point illustrated by Field et al.(2002) is that the connection between bulk chemistry and toxicity is a continuum. There is no concentration threshold above which toxicity is possible and below which toxicity is impossible.

Influence of particle size

It is well established that chemical concentrations in sediment vary inversely with particle size. One expects concentrations to be greater in fine- relative to coarse-grained sediment. Given that no account is taken of grain-size when determining ERLs, one would further expect ERL exceedances to increase in proportion to the fine-grained portion of sediment. That is, in fact, the case as shown on Table 1 based on the 2165 samples with grain-size determinations from the dataset used by O'Connor and Paul (1998). The frequency of ERL exceedances increased monotonically from 24 to 80% as the fine fraction increased. Toxicity, on the other hand, was unrelated to grain size except for being notably infrequent in the coarsest sediment samples.

Table 1. Frequency of ERL exceedances and toxicity among 2165 sediment samples (O'Connor and Paul, 2000) partitioned according to fine-grain fraction

percent silt&clay	number of samples	percent with an ERL exceedance	percent toxic
<20%	618	24	8
20-40%	280	51	20
40-60%	274	62	27
60-80%	379	74	26
>80%	614	80	22

Critiques of ERL uses

It is a mistake to treat ERLs as thresholds. This was done, for example, by the San Francisco Bay Estuarine Institute in their Pulse of the Estuary Report (2003), and by

Kiddon et al. (2003) in their assessment of the ecological condition of the U.S. Mid-Atlantic estuaries. In both of these any ERL exceedance was judged an indicator of degradation. SFEI (2003) uses ERLs as “contaminant guidelines [that] are generally intended to indicate if ...sediment is safe.” Kiddon et al (2003) refer to ERLs as “threshold concentrations below which adverse effects are rarely expected.” SFEI (2003) explained that “For sediment, 90% of samples collected from 1997-2001 exceeded a threshold for possible effect” and Kiddon et al (2003) concluded that “For the [Mid-Atlantic] region as a whole $30\% \pm 7\%$ of the estuarine area exceeded at least one...ERL..” In both cases, because of how ERLs were used, the situation with regard to sediment contamination was made to appear worse than it is.

If, rather than using ERLs as thresholds, the authors had indicated that an ERL corresponds roughly to a 10% likelihood of toxicity, readers would have a better understanding of the extent of the problem. Furthermore, the authors could have indicated that toxicity is hardly impossible at lower concentrations and that finer sediments tend to have more ERL exceedances than coarser ones, without any increase in toxicity. Kiddon et al. (2003) did provide perspective. They observed that sediment toxicity (less than 80% ten-day amphipod survival) was present in only 1% of the area and that it occurred with, and without, accompanying ERL exceedances. Nonetheless, their overall assessment of sediment contamination in the region was based on the percentage of sediment samples with any chemical concentration above an ERL.

Long and MacDonald (1998) improved the predictability of ERMs by introducing the idea of basing predictions on multiple exceedances. They concluded that with 6-10 ERM exceedances the probability of a sample's being toxic to amphipods would exceed 50%. However, neither they nor anyone else has ever tested the ability of multiple ERL exceedances to predict toxicity. Nonetheless, in the National Coastal Condition Report (EPA, 2001a), EPA effectively equated five ERL exceedances with 1 ERM exceedance when stipulating that sediments were degraded if either condition applied. There is simply no basis for such a claim.

Van Dolah et al. (1999) deemed a sediment sample to be degraded if its chemical concentrations included three or more ERL exceedances; there is no basis for attributing any significance to multiple ERL exceedances. In fact, in this case, the use of ERLs led to a deeper complication. Van Dolah et al. (1999) were classifying sediment to establish a benthic-index of biotic integrity (B-IBI). Their derived B-IBI is based on numbers and types of benthic organisms but it first had to be defined by separating reference from degraded sites. No account was made of sediment grain-size, but one criterion for that separation was three or more ERL exceedances. So because they naturally have higher chemical concentrations, fine-grained sediments had a tendency to fall into the degraded category, leaving coarse sediments as the reference condition for defining the B-IBIs. The B-IBIs in fine-grained sediment could be lower on a scale defined by benthic communities in coarse sediment; the difference would not be due to degradation but rather the fact that fine-grained sediments, by their very nature, harbor different benthic communities than sandy sediment.

Conclusion

ERLs are not thresholds below which sediment toxicity is impossible and above which it is likely. They are just concentrations at the low end of a continuum roughly relating bulk chemistry with toxicity. Categorizing sediments on the basis of whether their chemical concentrations include one or more ERL exceedance leads to misperceptions of the actual probability that such sediments are toxic.

References

- EPA. 2001a. National Coastal Condition Report. EPA-620/R-01/005 U S. Environmental Protection Agency. Washington, DC 204 pp.
- EPA. 2001b. National Sediment Quality Survey. Database – 1980 to1999. EPA-823-C-01-001. U S. Environmental Protection Agency. Washington, DC Compact Disc
- Field, L.J, D.D. MacDonald, S.B. Norton, C.G. Ingersoll, C.G. Severn, D. Smorong, and R. Lindscoog. 2002. Predicting amphipod toxicity from sediment chemistry using logistic regression models. *Environm. Toxicol. Chem.* 21: 1993-2005
- Kiddon, J.A., J.F. Paul, H.W. Buffum, C.S. Strobel, S.S. Hale, D. Cobb, and B.S. Brown 2003. Ecological condition of US Mid-Atlantic estuaries, 1997–1998 *Mar. Poll Bull.* 46:1224-1244
- Long, E.R. and L.G. Morgan. 1990. The Potential for Biological Effects of Sediment-Sorbed Contaminants Tested in the National Status and Trends program. NOAA Technical Memorandum NOS OMA 52, Seattle, WA 175 pp & appendices
- Long, E.R., D.D. MacDonald, S.L. Smith, and F.D. Calder. 1995. Incidence of adverse biological effects within ranges of chemical concentrations in marine and estuarine sediments. *Environm. Mgmt.* 19: 81-97
- Long, E.D. and D.D. MacDonald. 1998. Recommended uses of empirically derived, sediment quality guidelines for marine and estuarine ecosystems. *Human and Ecol. Risk Assess.* 4: 1019-1093
- Mearns, A. J., R. C. Swartz, J. M. Cummins, P. A. Dinnel, P. Plesha, P., and P. Chapman. 1986. Inter-laboratory comparison of a sediment toxicity test using the marine amphipod, *Rhepoxynius abronius*. *Mar .Environ .Res .*19: 13-37
- O'Connor, T. P., K. D. Daskalakis, J. L. Hyland, J. F Paul, and J. K. Summers. 1998. Comparisons of measured sediment toxicity with predictions based on chemical guidelines. *Environ. Toxicol. Chem.* 17: 468-471
- O'Connor, T. P. and J. F. Paul. 2000. Misfit between sediment toxicity and chemistry. *Mar. Poll. Bull.* 40: 59-64

SFEI. 2003. *The Pulse of the Estuary: Monitoring and Managing Contamination in the San Francisco Estuary*. SFEI Contribution 74. San Francisco Estuary Institute, Oakland, CA. 40 pp.

Van Dolah, R.F., J.L. Hyland, A.F. Holland, J.S. Rosen, and T.R. Snoots. 1999. A benthic index of biological integrity for assessing habitat quality in estuaries of the southeastern USA. *Mar. Environm. Res.* 48: 269-283