2012). A tabulation of the details for 60 cases shows the variety of guises in which the error appears. These include multiway contingency tables as well as multiway ANOVAs. A major cause of pseudofactorialism is the widespread failure of statistical texts, the primary literature, and documentation for statistics software to distinguish the 3 major components of experimental design—treatment structure, design structure, response structure (see Hurlbert 2013)—and clearly and correctly define key terms such as experimental unit, evaluation unit, split unit, factorial, and repeated measures.

REFERENCES

Hurlbert SH. 1984. Pseudoreplication and the design of ecological field experiments. *Ecol Monogr* 54:187–211.

Hurlbert SH. 2009. The ancient black art and transdisciplinary extent of pseudoreplication. *J Comp Psychol* 123:434–443.

Hurlbert SH. 2012. Pseudofactorialism, response structures, and collective responsibility. *Austral Ecol* 38:646–663.

Hurlbert SH. 2013. Affirmation of the classical terminology for experimental design via a critique of Casella's Statistical Design. *Agronomy J* 105:412–418.

Hurlbert SH. 2015. [Pseudoreplication update for ESA "Notable Papers" project]. Centennial Special: Notable Papers: Ecological Monographs. Available from: http://www.esajournals.org/page/centennial-ecologicalmonographs or contact author for pdf.

Hurlbert SH, Lombardi CM. 2009. Final collapse of the Neyman-Pearson decision theoretic framework and rise of the neoFisherian. *Ann Zool Fenn* 46:311–349.

Hurlbert SH, Lombardi CM. 2012. Lopsided reasoning on lopsided tests and multiple comparisons. *Austral New Zealand J Stats* 54:23–42.

Hurlbert SH, White MD. 1993. Experiments with freshwater invertebrate zooplanktivores: Quality of statistical analyses. *Bull Mar Sci* 53:128–153.

Lombardi CM, Hurlbert SH. 2009. Misprescription and misuse of one-tailed tests. *Austral Ecol* 34:447–468.

Siegel S. 1956. Nonparametric statistics for the behavioral sciences. New York (NY): McGraw-Hill. 399 p.

SSD MODELING—IT'S ALL ABOUT F!

David R Fox*†‡
†*Environmetrics Australia, Beaumaris, Victoria*
‡*University of Melbourne, Parkville Victoria, Australia*
*david.fox@environmetrics.net.au

DOI: 10.1002/ieam.1724

The ubiquity and pervasiveness of the species sensitivity distribution (SSD) in ecotoxicology has been well documented (Posthuma et al. 2002 and references therein). Articles have been published on many facets of the SSD modeling approach including, but not limited to

- Assumed randomness of the sample of species used to generate the data
- Limitations and difficulties due to extremely small sample sizes
- Mathematical and statistical considerations to do with functional form of the SSD, estimation strategies, and the inferential framework for a derived HCx

Although the assumption of randomness has been universally acknowledged (Forbes and Forbes 1993; Van der Hoeven 2004; ECHA 2008) as a necessary and key requirement demanded by statistical theory to ensure the validity of the approach, little or nothing has been done to address the invariable violations of this assumption in practice. Indeed, the advice given in most guideline documents actually guarantees nonrandomness. For example, the revised Australian and New Zealand Water Quality Guidelines recommend using toxicity data from at least 8 species from at least 4 taxonomic groups (Batley et al. 2014). Such purposive sampling is the antithesis of randomness. To further complicate matters, there has been a complete absence of any studies to quantify and describe the impact nonrandom species selection has on the fitted SSD and quantities (such as the HC_x) subsequently derived from it.

Motivated by this gaping hole in current ecotoxicological practice and my involvement in the preparation of Australia's revised guidelines, which recommended further research be undertaken on this unresolved issue, I took up the challenge. In a recent article I detail the results of those investigations (Fox 2015). At least initially, I have only investigated what happens to an SSD assumed to follow a log-logistic distribution when the selection of species used to generate the sample data has been biased. Using the very flexible beta distribution to characterize the selection function, it has been shown that the actual distribution arising from nonrandom sampling is *F*. This important result is conveniently and compactly summarized as follows.

The actual distribution when an assumed log-logistic SSD having parameters (α,β) is used to describe toxicity data (X) that have been selected according to a beta distribution with parameters (a,b) is a modified *F* distribution given by the probability density function (PDF) $g_x(x;a,b,\alpha,\beta) = \frac{b\beta}{a\alpha}\left(\frac{x}{\alpha}\right)^{\beta-1} dF\left[\frac{b}{a}\left(\frac{x}{\alpha}\right)^{\beta};2a,2b\right]$ where the notation $dF(\cdot;\nu_1,\nu_2)$ denotes a standard *F* distribution having ν_1 and ν_2 degrees of freedom. A requirement satisfied by this PDF is that when species selection is truly random (corresponding to the special case of the beta distribution with $a=b=1$) the distribution is the assumed log-logistic.

A useful outcome of this result is that it is possible to construct a bias correction factor (BCF), which adjusts the HC_x derived from the fitted SSD to compensate for the nonrandom selection of species data. This BCF is readily computed using intrinsic functions found in Microsoft Excel. Analysis of the BCF for various selection function shapes suggests that, if the toxicity data are biased toward the more sensitive species as has been suggested (Versteeg et al. 1999), then the common practice of using the lower limit of a confidence interval for the estimated HC_x may be compensating in the wrong direction. As I note (Fox 2015), this is an issue that requires further research and evaluation as the implications for what has hitherto been understood to be a "protective concentration" may be profound. For example, the SSD methodology is routinely used in Australia to determine a "safe" dilution for effluents from wastewater treatment plants and desalination plants. This, in turn, dictates the depth of an ocean outfall to achieve that dilution. Given that nonrandom species selection for the SSD can result in HC_x errors of a factor of 20 or more, it is apparent that either the environmental or monetary cost is potentially significant depending on the direction of this error.

REFERENCES

Batley GE, Braga O, Van Dam R, Warne MStJ, Chapman JC, Fox DR, Hickey C, Stauber JL. 2014. Technical rationale for changes to the method for deriving Australian and New Zealand water quality guideline values for toxicants. Sydney (AU): CSIRO Land and Water Report. 46 p. Available from: http://bit.ly/1SBtlot

[ECHA] European Chemicals Agency. 2008. Chapter R.10: Characterisation of dose [concentration]-response for environment. In: Guidance on Information Requirements and Chemical Safety Assessment. Helsinki (FI): European Chemicals Agency.

Forbes TL, Forbes VE. 1993. A critique of the use of distribution-based extrapolation models in ecotoxicology. *Funct Ecol* 7:249–254.

Fox DR. 2015. Selection bias correction for species sensitivity distribution modelling and hazardous concentration estimation. *Environ Toxicol Chem* 34:2555–2563.

Posthuma L, Suter GW II, Traas TP. 2002. Species sensitivity distributions in ecotoxicology. Boca Raton (FL): CRC Press. 616 p.

Van der Hoeven N. 2004. Current issues in statistics and models for ecotoxicological risk assessment. *Acta Biotheoretica* 52:201–217.

Versteeg DJ, Belanger SE, Carr GJ. 1999. Understanding single species and model ecosystem sensitivity: Data-based comparison. *Environ Toxicol Chem* 18:1329–1346.

ISSUES WITH USING ONLY REGRESSION MODELS FOR ECOTOXICITY STUDIES

John W Green*†
†*DuPont Applied Statistics Group, Newark, Delaware, USA*
*John.w.green@dupont.com
DOI: 10.1002/ieam.1723

INTRODUCTION

There is considerable momentum to move away from using NOECs to evaluate ecotoxicity studies to regression models. However, regression models have limitations that are sometimes poorly addressed in regulatory guidelines and by scientists trying to meet those guidelines. Tools for evaluating regression models are available and there are numerous examples of problematic data and also some types of data for which no models are currently available.

The distribution of effect concentration x (ECx) estimates from carefully designed simulation studies is a critical evaluation tool and supplements more traditional model selection tools such as Akaike's Information Criteria with finite sample correction (AICc) (Motulsky and Christopoulos 2004), lack-of-fit tests comparing error about regression to pure error, diagnostic plots, confidence interval width, sensitivity analysis, and common sense, an underused concept. The NOEC often provides good information when no sound regression model exists.

DISCUSSION

Ecotoxicity data sometimes need to be transformed before analysis to satisfy model requirements. Traditional models often assume normality and variance homogeneity of the response; failure to satisfy those requirements can bias the estimated ECx. When data are transformed, the meaning of an x% effect changes. For example, a 20% change in the logarithm of length is not equivalent to a 20% change in length. In analyzing the proportion of seeds that emerge or eggs that hatched or phenotypically male fish exposed as embryos to an endocrine disrupting chemical, the need for a normalizing, variance stabilizing transform before the use of statistical hypothesis testing or traditional modeling is well known. This is not a concern for hypothesis testing, but it is for regression modeling. Alternative generalized linear mixed models (GLiMM) can be used for proportion data (Cameron and Trivedi 2013). Similar rethinking is needed for other types of responses.

Another key requirement of almost all models is that observations should be independent. For nontarget plant studies, the International Organization for Standardization (ISO 2005) advises normalizing responses to the control, expressing for example, the number of emerged seedlings for each pot as a percentage of the mean emerged in control pots. The normalized response is

$$P_{nij} = \frac{P_{ij} - P_0}{P_0},$$

where P_0 is the control mean and P_{ij} is the mean of the j^{th} replicate of the i^{th} treatment. Because these ratios all have the control mean, a random variable, in the denominator, they

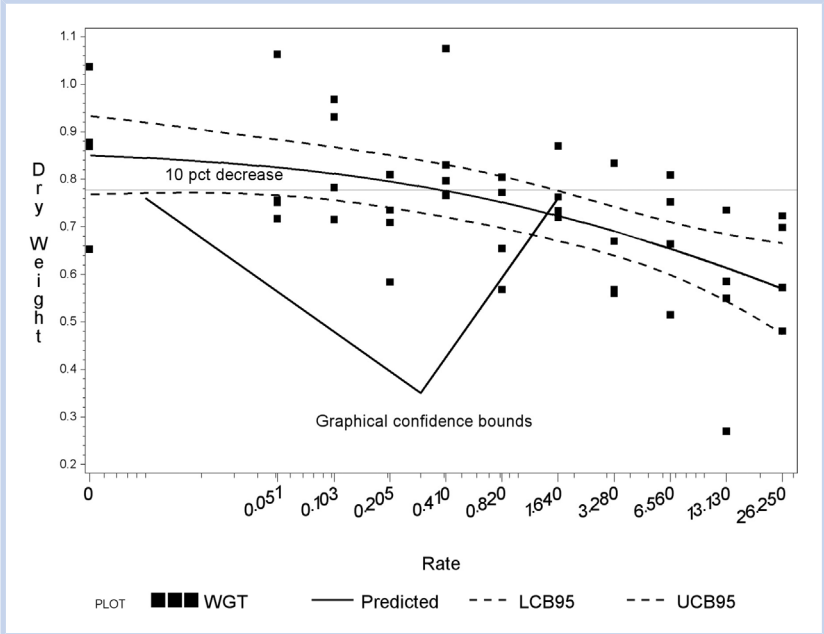

Figure 1. Dry weight of oat plants in a vegetative vigor study. WGT = oat dry weight in grams; Predicted = the fitted model; LCB95 and UCB95 = pointwise 95% confidence bounds for the predicted values (respectively lower and upper bounds), application rates in grams per hectare.