

Determination of Lethal Concentrations Using an R Software Function Integrating the Abbott Correction

Merveille Koissi Savi¹, Essomanda Tchanda Mangamana¹, Jean Marcel Deguenon², Castro Gbèmemali Hounmenou¹ and Romain Glèlè Kakaï¹

1. *Laboratoire de Biomathématiques et d'Estimations Forestières, Faculty of Agronomic Sciences, University of Abomey-Calavi, 04 BP 1525, Cotonou, Benin*

2. *Department of Entomology and Plant Pathology, North Carolina State University, Campus Box 7616, Raleigh, USA*

Abstract: Several thousands of chemical substances are registered every year for different purposes, and sometimes many of them are claimed to play the same role. To establish and compare their toxicities, the determination of the lethal concentrations is usually necessary and should account for natural mortality. However, many of the statistical software packages used for that purpose do not readily integrate control mortality or adjust the best link function to the data during the process. This manuscript proposes an “lc” function in the R open source that aims at the effective determination of lethal concentrations. Furthermore, it performs the procedure with the appropriate link function. The “lc” application on the example provided revealed that the complementary log link function is adequate.

Key words: Toxicity, lethal concentration, Abbott correction, binomial family links, generalized linear models, pesticides.

1. Introduction

Every year, several thousands of organic or inorganic substances are produced or improved, and further registered by the chemical abstract service. According to the chemical abstract service report, more than 122,740,000 substances are registered to date [1]. Among those substances, pesticides designed to solve pest problems are of particular interest, as they affect several aspects of life (animals, humans, plants and environment) [2]. To establish and compare their toxicity, the determination of the median lethal concentration (LC₅₀), which is the inhaled concentration of a substance that will lead to the deaths of 50% of the dosed population [3], is the most commonly used method [4, 5]. According to literature, lethal concentrations are mostly determined using probit or logit link function [5, 6]. The use of such links assumes that the dose tolerance curve has an inverse normal cumulative or logistic distribution [7].

When the dose tolerance is assumed to follow the Gumbel distribution, the complementary log link is used [8, 9]. The misspecification of the binomial link (probit, logit or complementary log) results in asymptotic bias and inefficiency of maximum likelihoods estimators (MLEs) [10]. As MLEs are used to fit LC₅₀, the misspecification will generate incorrect LC₅₀ values [11].

Because most testing procedures involve control treatments, the Abbott correction was developed to account for natural factors causing mortality in the control checks and to provide the actual “percent control” by the pesticide or corrected efficacy [12].

Although the use of Abbott correction has become a routine in bioassays of pesticides efficacy, it is not readily integrated in most of models regardless of the statistical software.

Pacheco and Rebelo [13] proposed a procedure in R software [14] for the determination of lethal concentrations, using a dataset from Calazans et al. [15] on NaCl toxicity to mollusks. The authors did not

Corresponding author: Merveille Koissi Savi, M.Sc., research field: biostatistics.

include the Abbott correction due to absence of mortality in the control treatment. Furthermore, the R package “ecotoxicology” [16] was proposed for the secured determination of lethal concentrations. However, its implementation lacks the selection of the best link in the binomial family. The package usage hence requires knowledge on the dose tolerance distribution.

The majority of computer based programs available for the LC₅₀ determination require the user to purchase a license [11, 17]. This manuscript proposes a computer based function in the free R software for the estimation of lethal concentrations (LC₅₀, LC₉₀ and LC₉₅). The proposed function integrates the Abbott correction and adjusts the best link function.

2. Materials and Methods

A function denoted “lc” was written for the determination of lethal concentrations in the open source R [14]. First, the “lc” function provides the corrected efficacy or actual percent control for the different concentrations by applying the Abbott correction. Second, to choose the best link function, it computes the value of the residual deviance for three links: probit, logit and complementary log. As the deviance is a measure of goodness of fit of the model, the best link will have the lowest value. Third, the concentrations tested are arbitrary log transformed as Raymond [18] recommended. The lethal concentrations (LC₅₀, LC₉₀ and LC₉₅) are then determined using the R library Modern Applied Statistics with S (MASS) [19].

At the end of the procedure, a goodness of fit test is proposed to the users to ascertain that the best

binomial family link was selected. Thus, a Chi square test is used to compare the residual deviance of the link function selected to the residual degrees of freedom, and the Naglekerke R square is computed [20]. In order to show how the proposed function works, a running example with a dataset (Table 1) from the USEPA office of water is provided [21]. The study measured fathead minnow larval survival when exposed to a control and five effluent concentrations in freshwater.

3. Results and Discussion

The code is designed to integrate the Abbott correction (Abb) first. This correction is given by the Eq. (1):

$$\text{Abb} = \frac{X - Y}{X} \quad (1)$$

where, X denotes the number of living organisms in the control and Y is the living organisms in the treatment at the end of the exposure time [12].

The Abbott correction gives values ranging between -1 and 1. For a corrected mortality of value below zero, the program returns the value zero, because the mortality rate could not be below zero [21]. The three link functions in the binomial family further tested by the “lc” program are, respectively, “probit”, “logit” and “complementary log”. The following Eqs. (2)-(4) show the transformations that are performed after the Abbott correction (Abb).

With the “probit” link, Bliss [22] assumed that the response variation (here Abbott corrected mortalities (Abb)) followed a lognormal distribution. Therefore the function should be written as Eq. (2):

Table 1 Mortality data from a fathead minnow larval survival test exposed to five effluent concentrations and control USEPA [21].

Effluent concentration (%)	Mortality	Total
0.00	2	40
6.25	2	40
12.50	0	40
25.00	0	40
50.00	26	40
100.00	40	40

$$g(\text{Abb}) = \Phi^{-1}(\text{Abb}) \quad (2)$$

According to Berkson [23], the link function “logit” assumes that the response variation (Abb) follow a logistic distribution as displayed in Eq. (3):

$$g(\text{Abb}) = \log\left(\frac{\text{Abb}}{1 - \text{Abb}}\right) \quad (3)$$

The complementary log “cloglog” link function is asymmetric and produces different results from the “probit” and “logit” link functions [8, 9], as Eq. (4):

$$g(\text{Abb}) = \log[-\log(1 - \text{Abb})] \quad (4)$$

After the selection of the best link function, the “lc” function computes and provides the lethal concentrations as well as the 95% confidence intervals.

In the provided case study (Table 1), the best link function selected is the complementary log. The lethal concentration LC_{50} using sigmoidal regression was 49.44% (effluent concentrations were provided as percentages of active ingredient) (Fig. 1). The test of the goodness of fit of the link function revealed that there is no significant difference between the residual

deviance and the residual degrees of freedom. Such result demonstrates that the adequate link function was selected and the lethal concentrations were correctly determined. The Naglekerke R square is equal to 1, which means that the link function used explained 100% of mortality variation.

The “lc” function developed provides the possibility to use corrected mortality rates to determine lethal concentrations. The Abbott correction was developed since 1925 and is widely used in establishing pesticide efficacy or toxicity level. Incorporating its automatic use through a lethal concentration determination function in a statistics-oriented open source, such as R, is important and will reduce computing time to users as well.

Moreover, as the “lc” function compares the residual deviance of the link function and the degrees of freedom for the selection of the best binomial family link, it also reduces biases during the calculations [10].

The results from the “lc” function applied to the example showed that the complementary log link was the best. In fact, several studies have reported the “probit” as the best link when compared to the “logit”

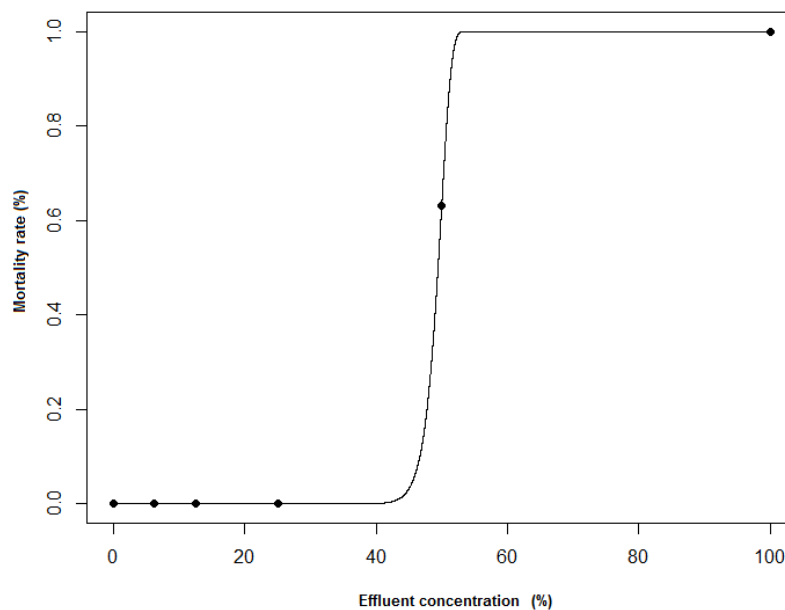


Fig. 1 Mortality rate of fathead minnow in larval state exposed to control and five effluent concentrations.

link for computing lethal concentrations [24, 25]. However, based on the results in this study, more attention should be paid to the selection of the adequate link function (testing more than probit and logit), which normally guaranties effective results.

4. Conclusions

The “lc” program was written and the open source R software code is provided for the determination of lethal concentrations. The “lc” integrated the Abbott correction to account for control mortality and helps select the adequate link function. The application of “lc” on the running example suggested that additional studies are needed for the binomial link functions comparison in the context of lethal concentration determination.

References

- [1] Chemical Abstracts Service (CAS). 2016. “CAS Registry—The Gold Standard for Chemical Substance Information.” Accessed November 03, 2016. <https://www.cas.org/content/chemical-substances>.
- [2] Geiger, F., Bengtsson, J., Berendse, F., Weisser, W. W., Emmerson, M., Morales, M. B., Ceryngier, P., Liira, J., Tschamtker, T., Winqvist, C., Eggers, S., Bommarco, R., Pärt, T., Bretagnolle, V., Plantegenest, M., Clement, L. W., Dennis, C., Palmer, C., Oñate, J. J., Guerrero, I., Hawro, V., Aavik, T., Thies, C., Flohre, A., Hänke, S., Fischer, C., Goedhart, P. W., and Inchausti, P. 2010. “Persistent Negative Effects of Pesticides on Biodiversity and Biological Control Potential on European Farmland.” *Basic and Applied Ecology* 11 (2): 97-105.
- [3] Gad, S. C. 2014. “LD₅₀/LC₅₀ (Lethal Dosage 50/Lethal Concentration 50).” In *Reference Module in Biomedical Sciences: Encyclopedia of Toxicology*, 3rd ed.. Cary, NC, USA: Elsevier Inc., 58-60.
- [4] Huertas, M., Gisbert, E., Rodriguez, A., Cardona, L., Williot, P., and Castello-Orvay, F. 2002. “Acute Exposure of Siberian Sturgeon (*Acipenser baeri*, Brandt) Yearlings to Nitrite: Median-Lethal Concentration (LC₅₀) Determination, Haematological Changes and Nitrite Accumulation in Selected Tissues.” *Aquatic Toxicology* 57 (4): 257-66.
- [5] Farah, M. A., Ateeq, B., Ali, M. N., Sabir, R., and Ahmad, W. 2004. “Studies on Lethal Concentrations and Toxicity Stress of Some Xenobiotics on Aquatic Organisms.” *Chemosphere* 55 (2): 257-65.
- [6] Sanathanan, L. P., Gade, E. T., and Shipkowitz, N. L. 1987. “Tiimed Logit Method for Estimating the ED₅₀ in Quantal Bioassay.” *Biometrics* 43 (4): 825-32.
- [7] Kalish, L. A. 1990. “Efficient Design for Estimation of Median Lethal Dose and Quantal Dose-Response Curves.” *Biometrics* 46 (3): 737-48.
- [8] West, B. T., Welch, K. B., and Galecki, A. T. 2007. *Linear Mixed Models: A Practical Guide Using Statistical Software*. New York: Chapman & Hall/CRC.
- [9] Verbeke, G., and Molenberghs, G. 2000. *Linear Mixed Models for Longitudinal Data*. New York: Springer-Verlag.
- [10] Czado, C. 1992. “The Effect of Link Misspecification on Binary Regression Inference.” *Journal of Statistical Planning and Inference* 33 (2): 213-31.
- [11] O’Brien, T. E. 2016. “Efficient Experimental Design Strategies in Toxicology and Bioassay.” *Statistics, Optimization and Information Computing* 4 (2): 99-106.
- [12] Abbott, W. S. 1925. “A Method of Computing the Effectiveness of an Insecticide.” *Journal of Economic Entomology* 18: 265-7.
- [13] Pacheco, A. G., and Rebelo, M. F. 2013. “A Simple R-Based Function to Estimate Lethal Concentrations.” *Marine Environmental Research* 91: 41-4.
- [14] R Core Team. 2016. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
- [15] Calazans, S. H. C., Americo, J. A., Fernandes, F. C., Aldridge, D., and Rebelo, M. F. 2013. “Assessment of Toxicity of Dissolved and Microencapsulated Biocides for Control of the Golden Mussel *Limnoperna fortunei*.” *Marine Environmental Research* 91: 104-8.
- [16] Gama, J. 2015. “Ecotoxicology: Methods for Ecotoxicology.” R Package Version 1.0.1. Accessed October 14, 2015. <https://CRAN.R-project.org/package=ecotoxicology>.
- [17] Russell, R. M., Robertson, J. L., and Savin, N. E. 1977. “POLO: A New Computer Program for Profit Analysis.” *Bulletin of the Entomological Society of America* 23 (3): 209-13.
- [18] Raymond, M. 1985. “Log-Probit Analysis Program Presentation for a Microcomputer.” *Entomologie Médicale et Parasitologie* 23 (2): 117-21. (in French)
- [19] Venables, W. N., and Ripley, B. D. 2002. *Modern Applied Statistics with S*, 4th ed.. New York: Springer.
- [20] Nagelkerke, N. J. D. 1991. “A Note on a General Definition of the Coefficient of Determination.” *Biometrika* 78 (3): 691-2.
- [21] United States Environmental Protection Agency (USEPA). 2002. *Short-Term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Freshwater Organisms*, 4th ed.. Washington, DC, USA:

- National Service Center for Environmental Publications, 350.
- [22] Bliss, C. I. 1934. "The Method of Probits." *Science* 79 (2037): 38-9.
- [23] Berkson, J. 1944. "Application of the Logistic Function to Bioassay." *Journal of the American Statistical Association* 39 (227): 357-65.
- [24] Hoekstra, J. A. 1991. "Estimation of the LC_{50} : A Review." *Environmetrics* 2 (2): 139-52.
- [25] Chambers, E., and Cox, D. R. 1967. "Discrimination between Alternative Binary Response Models." *Biometrika* 54 (3-4): 573-8.

Appendix

```

The function lc() in the open source of R
rm(list=ls()) # Clears the memory of R
library(MASS) # Load the package useful for the determination of lethal concentrations
lc<-function(d,x,t){ # This function has three main entries that are the concentrations of effluents (d), the number of dead (x) after
treatment and the total number of species (t) that received effluent concentrations (lc)
  options(warn=-1) # This is to avoid warning alerts
  # The following function will compute the corrected mortality rates
  Abbott<-function(d,x,t){
  ab<-(1-((t-x)/(t[d==0]-x[d==0])))
  ifelse(ab < 0,0,ab[])
  Abb<-Abbott(d,x,t)
  mynewdat<-data.frame(d,x,t,Abb); (mynewdat = mynewdat[!d==0,]) # This part removed the control which is no more useful for
the determination of lc
  mod<-list()
  #The following help to select the binomial family link
  family<-function (i){
  if (i==1){family=binomial(link = "probit")}
  else{
  if (i==2){family=binomial(link = "logit")}
  else{
  if (i==3){family=binomial(link = "cloglog")}
  }}
  #return(family)
  }
  for (i in 1:3){
  mod[[i]]<-glm(Abb ~ log(d), family=family(i), data=mynewdat)
  }
  b<-which.min(c(deviance(mod[[1]]), deviance(mod[[2]]), deviance(mod[[3]]))) # Compare and select the best link
  cat("The best model is the model", b,"\n") # Gives the output of the previous comparison
  out=list()
  out$Resum<-summary(mod[[b]])
  ld<-dose.p(mod[[b]],p=c(0.50,0.90,0.95)) # Computation of lethal concentrations 50, 90 and 95 but may also 10, 20, 80 etc...
  ld.ci <- ld+ attr(ld, "SE") %*% matrix(qnorm(1 - 0.05/2)*c(-1,1), nrow=1) # Defines the confidence intervals
  out$ld.est <- round(exp((cbind(ld, attr(ld, "SE"), ld.ci[,1], ld.ci[,2]))),3)
  dimnames(out$ld.est)[[2]] <- c("LD", "SE", "LCL", "UCL") # Return the lethal concentrations and their confidence intervals
  #The following part plots the curve
  x1=seq(min(d),max(d), 0.001)
  lpredmod <- predict(mod[[b]], data.frame(d=x1), type="response")
  plot(d,Abb, pch=16, ylim=c(0,1),xlab="Concentration ", ylab = "Mortality rate")
  lines(x1,lpredmod)
  #Goodness of fit of the model
  out$comparison<-pchisq(deviance(mod[[b]]), mod[[b]]$df.residual, lower=FALSE) # Comparison of deviance
  out$R2_Naglekerke<-round((1-exp(-(mod[[b]]$dev-mod[[b]]$null)/sum(t)))/(1-exp(-mod[[b]]$null/sum(t))),3) #Determination of
Naglekerke R square
  return(out)
  }
  The example
  d<-c(0.00,6.25,12.50,25.00,50.00,100.00)
  x<-c(2,2,0,0,26,40)
  t<-c(40,40,40,40,40,40)
  lc(d,x,t)

```