Practice with with `morse` and `MOSAIC`

Sandrine CHARLES & Aude RATIER

`Contact:sandrine.charles@univ-lyon1.fr`

Master EPET the 31^{th} of March, 2021

Introduction

The purpose of this document is to guide you through the practical session with the R-package `morse`. This document complements the slides of `02-lecture-morse.pdf` which provides an introduction to the R-package `morse`.

This document will help you to run the first example as described in the lecture, as well as other applications. The solution (R code, outputs and discussion points) is in `support-morse.pdf`.

First, you will work with the survival and reproduction data set *chlordan* (section 1). Then you will analyse another data set with toxicity data on *D. magna* exposed to cadmium (section 2), as well as a data set with toxicity data on snails also exposed to cadmium (section 3). In section 4, you will be introducted to dose-response analyses of quantitative continuous data with the MOSAIC$_{growth}$ web application. In section 5, the MOSAIC$_{bioacc}$ will allow you to get bioaccumulation factors, while the last section is dedicated to GUTS models.

1 *D. magna* exposed to chlordan

The idea of this first practical part is to redo the analyses and plots as in the lecture.

Open the R script `lecture-morse.R`

Step 1.1 Run instructions line by line in order to understand what each line is providing. Explore each newly created object to understand its content. Try the different options in the functions to explore all their features.

Step 1.2 Try another target time.

Step 1.3 Get LC_x and EC_x for $x = 5, 15, 25, 45, 75$. What do you notice about 95% credible intervals? Try to make an appropriate plot to visualize LC_x or EC_x versus x.

Step 1.4 Change colors of the fitting plot: the median curve in `darkslategray4` the 95% credible band in `darkslategray1`.

If needed, look at the help about function `plot.survFitTT()`.

2 *D. magna* exposed to cadmium

Step 2.1 Load data set `cadmium1` from package `morse`. Look at the raw data. Which differences do you notice with the previous data set?

Step 2.2 Plot the raw survival data as well as the concentration-response curve at the final time of experiment (default target time).

Step 2.3 Fit a survival model at target time and get some LC_x estimates of interest.

Step 2.4 Plot your fit results in different ways. Give your comments and conclusions.

Step 2.5 Perform a complete reproduction analysis. Look at the results and make some conclusions.

Step 2.6 Compare the sensitivity of *D. magna* to chlordan to the sensitivity of *D. magna* to cadmium, regarding both survival and reproduction.

3 *L. stagnalis* exposed to cadmium

Load the data set `cadmium2` from package `morse` and perform both a survival and a reproduction analyses.

Step 3.1 Make comments and give conclusions on this data set.

Step 3.2 Compare sensibilities of *D. magna* and *L. stagnalis* to cadmium for both survival and reproduction.

4 Quantitative continuous data analysis

Step 4.1 Go back to `http://mosaic.univ-lyon1.fr/` and choose menu `Growth`. This is the way to use MOSAIC at your office. During this training course, please use the pedagogical server which ensures that all participants can perform exercises at the time:

`http://umr5558-shiny.univ-lyon1.fr/mosaic-growth/`.

Step 4.2 Data visualisation

- Try first with example `plant07.txt`. Notice that either a plot and a table are provided to check data have been correctly uploaded. Note that the user may personalize X- and Y-axis labels and choose the appropriate units for exposure, measurements and time.
- An innovative feature in this module is the possibility to upload several files **at the same time**. Try to upload these additional examples `chlordan-daphnia.txt`, `cadmium-daphnia.txt`, `subst01-lymnaea.txt`, `plant03.txt` and `plant10.txt`. What do you notice? Which data set is displayed? Try to change file.

Step 4.3 Dose-response analysis

- Keeping several files uploaded, go to sub-menu `Dose-response analysis`. Run the analysis for example `plant07.txt`. What is the estimate of the EC_{50} (denoted ER_{50} within the module). Provides the median and the 95% credible interval.
- From goodness-of-fit results, get the correlation plot. What do you notice? Don't hesitate to read the help text below the figure.
- Ask for the calculation of ER_{25}, ER_{50} and ER_{75}. What do you notice in terms of precision of the estimates?
- Run additional analyses for examples `plant03.txt` and `plant10.txt`. Look at the ER_{50} estimate from file `plant10.txt`. What do you notice? How could you explain such a result?
- Is the PPC plot suitable enough to trust provided results for further ERA?
- Ask the calculation of ER_{50} for `plant03.txt` and `plant10.txt`. Ask for the table displaying all ER_{50} for files `plant03.txt`, `plant07.txt` and `plant10.txt`. Which plant species is the most sensitive?

Step 4.4 Downloads

- Go to sub-menu `Downloads`.
- For example `plant07.txt`, download the single report (choose the Word format).
- Download the text file of the joint posterior distribution.

Step 4.5 Prediction tool

- Go to sub-menu `Prediction tool`.
- Enter a concentration range, for which parameter value of the three-parameters log-logistic model will be available. For example: `40; 80; 160; 320; 640`.
- **As a first try**, choose non distributed parameters (default option).
- **Option (1)**: enter a single value for each parameter, based on your experience, based on expert knowledge or coming form the literature for example. Keep in mind that these values where obtained for a given exposure duration, chosen by the experimenter, usually equal to the duration of the experience (e.g., the standard duration of 21 days for most of the toxicity tests performed under an OECD guideline). Try option (1) with parameter values entered by default as they are; please note that parameter b and e are expected to be given in log_{10}. What do you notice?
- Change the previous value of parameter e from $log_{10}(e) = 2.2$ to $log_{10}(e) = 3$. What do you notice?
- **Option (2)**: get median parameter estimates from a DR analysis that has already been performed on a data set (either with $\text{MOSAIC}_{\text{growth}}$ or with any other software), under experimental conditions that are close enough to those for which you want to make the prediction; in particular pay attention to simulate for a similar exposure duration. Try option (2) with parameter estimates obtained from `plant07.txt` data set. What do you notice?
- Try again option (2), but with parameter estimates obtained from `plant10.txt` data set. Keep again the same concentration range. What do you notice?
- Based on the previous simulation, find a concentration range that would be more appropriate.

- **As a second try**, choose **distributed parameters**. This necessary means that you previously downloaded a joint posterior distribution of the DR model parameters, what is today only possible with the $\text{MOSAIC}_{\text{growth}}$ module, or that you previously performed a DR analysis with $\text{MOSAIC}_{\text{growth}}$.
- **Option (1)**: coming back to the initially proposed concentration range (namely `40; 80; 160; 320; 640`), upload file:

  `outputPosterior_plant07.txt`

  You may have already downloaded earlier. If not, go to option(2). Which difference(s) do you observe with previous predictions?
- **Option (2)**: choose option `from a previous DR analysis`, that is supposed to have been performed on data set `plant07.txt` earlier. What do you notice?
- Change the concentration range and see what happens.

For further details on $\text{MOSAIC}_{\text{growth}}$, you can download and read the Tutorial available on line from here: `http://umr5558-shiny.univ-lyon1.fr/mosaic-growth/Tutorial.pdf`
A vignette also accompanies the Tutorial, with details on the modelling framework used in $\text{MOSAIC}_{\text{growth}}$: `http://umr5558-shiny.univ-lyon1.fr/mosaic-growth/vignette.pdf`.
Its reading is optional for the use of the application.

5 Bioaccumulation factors calculation

Step 5.1 Go back to `http://mosaic.univ-lyon1.fr/` and choose menu `bioacc`. This is the way to use MOSAIC at your office. During this training course, please use the pedagogical server which ensures that all participants can perform exercises at the time:

`http://umr5558-shiny.univ-lyon1.fr/mosaic-bioacc/`.

For further details on $\text{MOSAIC}_{\text{bioacc}}$, you can download and read the user-guide available on line by clicking on the dedicated button.

Step 5.2 Data visualisation

- Try first with example `Male_Gammarus_Single.txt`. Notice that either a plot and a table are provided to check data have been correctly uploaded. Note that the user has nothing to do, everything is automatically filled in.

- Download example file `Male_Gammarus_Single.txt` and browse the file. Notice that the user has to enter the correct separator, the time unit and the duration of the accumulation phase. Based on the description of the example, fill in the correct inputs.
- Try with example file `Oncorhynchus_two.csv` and look at the data. Notice that the user has now the choice between two exposure concentrations.
- Try with example file `Male_Gammarus_seanine.csv`. Notice that in plot visualisation, parent concentrations, metabolite concentrations and growth measurements are available.

Step 5.3 Model and parameters

- Upload again example file `Male_Gammarus_Single.txt` and go to sub-menu `Model and parameters`. The parameters of the model are automatically selected according to the uploaded data. In addition, the equations of the model (for the deterministic part only) are displayed with the corresponding variables. For this data set, how many parameters are there? And How many equations? Can you deduce the exposure routes (water, sediment, pore water and/or food?) and elimination processes (excretion, growth dilution and/or biotransformation?) that were considered in the experiment?
- Idem with example file `Chironomus_benzo-a-pyrene.csv`.
- Idem with example file `Male_Gammarus_seanine.csv`.
- Upload again example file `Male_Gammarus_Single.txt` and run the analysis with a click on the "Calculate and display" button. Calculations can take several minutes depending on the data set.

Step 5.4 Results and goodness-of-fit criteria

- From the analysis of example file `Male_Gammarus_Single.txt`, what is the estimate of the BCF_k? Provide the median and the 95% credible interval.
- Is it reasonable to ask for the BCF_{ss}? If yes, provide its median and its 95% credible interval.
- What are the estimates of parameters k_{uw} and k_{ee}? Provide their median and 95% credible intervals.

- From `Goodness-of-fit` sub-menu, get the Posterior Predictive Check (PPC). What is the percentage of data within the prediction intervals? Do not hesitate to read the help-text next to the figure.
- Get the `Priors and Posteriors` plot. Are posterior distributions narrower than prior ones?
- Go to the `Correlation plot`. What do you notice?
- Go to the `Potential Scale Reduction Factors` (PSRF). Are the values acceptable for convergence for all parameters?
- Do the MCMC chains overlap well for each parameter?
- Idem with example files `Chironomus_benzo-a-pyrene.csv` and `Male_Gammarus_seanine.csv`. Note that calculations can take several minutes, especially for the second file. It would be quicker to directly use the R code in RStudio, but that would reuiqre to deeply go to the code itself (in case you would like to do such a way, please ask us).
- What do you notice for the bioaccumulation factor with data set`Chironomus_benzo-a-pyrene.csv`?

Step 5.5 Downloads

- Go to sub-menu `Downloads`.
- For example `Male_Gammarus_Single.txt`, download the report (choose the Word format).
- Download the text file of the joint posterior distribution and the R code.

6 Using GUTS models (survival data over time)

Step 6.1 Go to `http://mosaic.univ-lyon1.fr/` and choose Menu `surv`, then sub-menu `GUTS-fit`.

Step 6.2 Regarding the provided examples, which difference(s) do you notice with the other MOSAIC modules?

Step 6.3 Try with example `chlordan`. Choose the GUTS-RED-SD model and click `run`. While getting the fit results (it can take few minutes), is the 'Data description' section different from the one in `surv-Standard` module?

Step 6.4 Once fit results have been obtained (don't forget to **export posteriors**), what do you notice?

Step 6.5 What about parameter estimates?

Step 6.6 What about the LC_x estimates?

Step 6.7 What about the PPC plot?

Step 6.8 Try with example `cadmium1`. What do you notice? What about the LC_{50}? Do you have the same result with the GUTS-RED-IT model?

Step 6.9 Try with example `Ring-test Dataset A IT`. Fit a GUTS-RED-IT model. What do you notice?

Step 6.10 Try with example `Ring-test Dataset B-cst`. Fit a GUTS-RED-SD model. Check you have the same result as in Fig. 16 of the Scientific Opinion from EFSA PPR Panel (2018, p57).

Step 6.11 Check that the PPC plot is also the same as Fig. 17 of the Scientific Opinion from EFSA PPR Panel (2018, p58).

Step 6.12 What a about the dose-response curve? What are the differences with Fig. 18 of the Scientific Opinion from EFSA PPR Panel (2018, p59)? How to bridge the gap?

Make predictions from the calibration results

In order to perform predictions by using GUTS models, for example from time-variable exposure concentration profiles, you need to go to `http://mosaic.univ-lyon1.fr/` and choose menu `surv`, then sub-menu `GUTS-predict`.