

CMEE Masters: Miniproject Assessment

February 14, 2022

Assignment Objectives: To address on a model-fitting problem using computational methods, and produce a written report, all in a coherent, reproducible, modular workflow under version control.

Student's Name: Amelia Wake

Overall Miniproject Mark: 42%

Overall Project Organization

Your directories are generally in place and uncluttered, with the exception of the Results directory which needed to be added (see below).

Your naming convention is consistent and the filenames are clear and informative. Nonetheless it is good practise to include a **readme** file documenting the versions of R/Python/any other tools you used, describing your different files, their functions and their dependencies, and giving clear instructions on how to run the project. We would strongly recommend including a readme for your final dissertation.

You could have put the report .tex source files and pdf in a separate directory – also a recommendation for your dissertation.

Overall the organisation of the project is clean and logical, but in future you should take care to ensure that all the directories needed for your workflow to run to completion are present.

The Code

Your choice of coding tools is generally sensible, and sticking to a preferred language like R is perfectly fine, particularly for less computationally intensive tasks such as you are doing here. Do remain open to using Python or C for more intensive work in future though. Your choice of packages is also generally sensible, and mostly makes use of tidyverse wrangling tools or extremely common plotting tools like ggplot2. This is good practise - excessive reliance on packages would restrict your growth as a programmer, and obscure packages are generally bad for reproducibility.

Your code commenting is good, and gives an at-a-glance sense of what is being done. You have also done a reasonable job of partitioning your code into functions/modules, although future and more complex projects might benefit from separate scripts for defining functions and for the execution of your code.

Your project master script encountered two errors when run on our machines. Firstly, as mentioned above, the Results subdirectory was not present in your submitted repo, and needed to be manually created in order for the ModelFitting.R script to successfully produce plots. Secondly, each of your R scripts tried to load data using absolute rather than relative file paths, and so needed to be changed to run on our machines. The ModelFitting.R script generated warnings

regarding the number of iterations in nlsLM. Ideally these should be addressed so that no warnings are given at all, but this is not always possible given time constraints, and is OK so long as the warnings aren't a sign of something fundamentally flawed. The Latex compilation section of the code threw up multiple non-fatal errors for undefined references, but did still successfully generate the report pdf.

You successfully fit 3 models (cubic, Gompertz and logistic) to your data. The cubic and Gompertz plots appear to demonstrate convincing fits, but the logistic plots are generally either constantly near-zero relative to the scale of the other models or in some cases negative, suggesting that you may have forgotten to "un-log" the logistic model estimates before plotting them on the same scale as the cubic and the Gompertz estimates (which you did un-log). Besides the plots for each subset, your workflow does not perform any model comparison using AIC, BIC or other statistical tests of model fitting, nor does it produce a summary of which model fits best overall or on any collection of the subsets.

Ideally your workflow would have included commands to delete all output files each time the workflow is run anew (i.e. so that they are always re-generated with each run). Although you do this for some of the auxiliary Latex files, you do not for the report pdf nor your R-generated results.

Your workflow did print information to the console that could be used to determine the progress of your scripts, but did not include explicit progress updates or time estimates, which you might consider in future to help the user understand where exactly they are in the workflow at any given moment while executing your code.

Your project ran in reasonable time, taking 45s in total, with the largest amount of time spent in ModelFitting.R, though adding the missing model comparison analyses would likely have increased this somewhat.

Overall, the components of your project that are present are relatively good, despite producing minor errors. Your organisation is clean and sensible and your code is reasonably commented. Ultimately however you are significantly held-back by not performing any direct model comparison after fitting your models, which was one of the main objectives of the project, and by the errors that needed to be manually fixed in order to run your project.

Marks for the project and computational workflow: 49%

The Report

Overall: It is clear that you did not allocate enough time for report writing. The majority of the report is simply not present, while the parts that are vary from perfunctory to just about adequate.

Title: Minimally descriptive but not inaccurate.

Abstract: Single-sentence description of the aims of the paper, mentioning mechanistic and phenomenological model fitting to functional response data. No background, results or take-home messages. (30%)

Introduction: Somewhat better. Reasonable background info motivating the use of model comparison tools like AIC to compare models. However no mention of the actual functional

trait (growth rates) measured by the input data, nor why it is important. Besides the background/motivation, no mention of the aims of the study or a specific research question. (42%)

Methods: Starts out reasonably well, describing the data wrangling approach, the choice of models, and the choice of model comparison algorithm. Clearly though the student ran out of time as the section ends abruptly with the statement of the logistic equation. (51%)

All remaining sections missing: (0%)

(Some specific feedback is in the attached pdf, and we can also discuss more aspects of your write-up in our 1:1 feedback meeting)

Marks for the Report: 35%

Signed: Samraat Pawar & Alexander Kier Christensen

February 14, 2022

Notes on Assessment :

- This written feedback will be discussed in a 1:1 session scheduled after this assessment has been given to you.
- The coursework marking criteria (included in this feedback at bottom) were used for both the computing and report components of the Miniproject Assessment. *In contrast*, Your final dissertation project marks are going to be based pretty much exclusively on the written report and viva (not code). Expect your final dissertation report to be marked more stringently, using the dissertation marking criteria (also included in this report).
- In the written feedback, the markers may have contrasted what you have done with what you should do in your actual dissertation. *This does not mean that you were penalized* — one of the main goals of the miniproject is to provide feedback useful for your main dissertation. However, there may be cases where what you have done is just really bad practise (for example missing line numbers or abstract), irrespective of whether it is a mini- or main- project report – you will be penalized in that case.
- The markers for this assessment are playing the role of somebody trying to understand and use your project organization and workflow from scratch. So it will seem like the feedback is particularly pedantic in places — please take it in the right spirit!