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: Izie Wood

Overall Miniproject Mark: 51%

Overall Project Organization

All your directories are in place and uncluttered.

You have included a `readme` file briefly describing the project aims, project structure, programming language used (with version number) and dependencies/packages. This is great practise, though you could also have included a list of key files and functions in your workflow in addition to the file structure of the repo.

You could have put the writeup LaTeX source files and pdf in a separate directory – this is what you should aim to do for your final dissertation.

Overall a clean and logical project organisation, reasonably documented.

The Code

Your choice of coding tools is appropriate, and having a preferred language is fine, though you should remain open to using Python or C in future, more computationally complex projects. You make judicious use of packages which is also good – over-reliance on packages stunts your development as a programmer and can lead to reproducibility issues with your workflow.

Your code is thoroughly commented and gives a quick at-a-glance sense of what is being done in each section of the code. Do be careful to eventually remove bits of code that you comment out entirely and end up not using at all (e.g. lines 76-78 in FindStartingValues.R) so that these fragments do not accumulate over time. In general your code is well partitioned into scripts for different subtasks, but your script names are sometimes a little confusing. For example ModelData.R contains only the code to fit the linear models, whereas the Gompertz model fitting happens in FindStartingValues.R. It is also considered best practise to define functions at the start of your scripts, just after importing any necessary packages, rather than adding definitions through the body of the file.

We encountered the following errors in running your master script. Firstly, we had to add "require(dplyr)" to FindStartingValues.R and ModelData.R, and had to change "pdflatex FirstExample.tex" to "pdflatex Miniproject.tex" to get your report to compile. The report compilation also exhibited a (silent) issue: To compile your report the enter key had to be held down, but there was no indication that this needed to happen, silent errors are particularly dangerous! You successfully fit 3 models (quadratic, cubic and Gompertz) to your data, and compare them

using AIC. However, we note that you chose to log all the population sizes even for the polynomial models. This is an unusual choice, and technically means you have chosen to investigate whether the log of the population follows a polynomial relationship w.r.t time, rather than the population itself. A better option might have been to fit the polynomial models to non-logged data, and the Gompertz models to logged data, and to manually calculate non-logged residuals for these so that you can still perform model comparison using AIC/BIC.

Recall that you should write into your workflow commands that will delete all existing output files every time the workflow is run (they should be re-generated afresh).

Your workflow could have incorporated progress updates printed to the terminal to signify progress through your code to users while they are running it.

Your project ran relatively quickly (40s) with most time spent in FindStartingValues.R where the initial parameters for Gompertz were estimated.

Overall your project meets the brief by fitting and comparing 3 models. The project is reasonably organised and well documented. However running the workflow required us to fix some fairly fundamental errors which ought to have been caught and fixed prior to submission. Once these errors were fixed the code ran relatively smoothly (Latex silent error aside) and in reasonable time.

Marks for the project and computational workflow: 54%

The Report

You clearly knows and understand the relevant background and available computing tools quite well, but did not allocate nearly enough time to complete the writing of the report. As a result the quality varies wildly from genuinely impressive to basically empty. Nonetheless you have met the project brief by fitting 3 models and comparing them.

Title: Clear, concise and informative.

Abstract: No abstract. (0%)

Intro: Impressive presentation of the philosophical background to the mechanistic- phenomenological distinction in scientific modelling. You might have added a paragraph or so describing the salience/importance of the specific kind of data you are working with here (growth rates), and a little additional text could have been dedicated to explaining that you would compare models by fitting them to the same data and using specific model comparison tools. (76

Methods: Quite good but perhaps a little rushed? Somewhat sparse description of the process of model fitting – not quite clear to me from the text what approach was taken for parameter estimation. Computing tools section present. (63%)

Results: Clearly ran out of time. Only a summary table and a couple of lines describing it. The text refers to additional tables but these are not present in the pdf or the tex file. Just about meets the brief since the table does compare the models using AIC. (50%)

Discussion: Merely reiterates the result presented immediately prior. (10%)

(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: 48%

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 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!