# 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**: Kayleigh Greenwood

**Overall Miniproject Mark**: 60%

## Overall Project Organization

All your directories are present and minimally cluttered, though we would have expected the Plots subdir to be inside the results subdir.

You have included a comprehensive `readme` file, which breifly describes the project, lists the programming languages used (with version numbers), lists dependencies/packages, and lists and details the key scripts.

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 relatively clean and well organised project with good documentation. Nicely done.

## The Code

Your chosen coding tools are appropriate. It is fine to have a preferred language like R but be open to using Python or C for more computationally complex tasks in future. You are appropriately sparing with your use of packages which is good – over reliance on packages can limit your development as a programmer and pose problems for reproducibility in future.

Your code is clearly and sensibly commented, giving a nice at-a-glance sense of what is happening in each section of your code. For future, more complex projects you might consider writing separate scripts to hold all your functions and the main body of your code.

We encountered one error when running your workflow, in that we had to create the Plots subdirectory ourselves in order for figures to successfully be saved to it. You successfully fit 2 models (cubic and logistic) and compare them using AIC and RSS. However, we note that you chose to log the population sizes for both the cubic and logistic models. This is an unusual choice, and technically means you have chosen to investigate whether (for the cubic model for example) the log of the population follows a polynomial or logistic relationship w.r.t time, rather than the population itself. One would typically fit these models to non-logged data, incontrast to models explicitly designed for log inputs (e.g. Baryani, Gompertz).

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). Also, put in checks so that the computational workflow aborts if any step in the analysis gives an error. Reporting that error to the user is a good idea too.

Your workflow did not incorporate progress updates displayed to the terminal, except for a flood of messages during the model fitting script. Although it is good to keep track of errors and warnings during model fitting, when many warnings or other outputs appear in quick succession it becomes essentially impossible to determine whether or not they are meaningful. It would be good practise to redirect repeated warnings and other non-critical outputs to log files, and simply flag on the terminal that warnings are being logged if they begin to appear. This would also free up terminal space during your run for you to print meaningful indicators of progress or even time estimates.

Your project ran extremely quickly ( 5s). This is partly because you elected not to implement any computationally expensive initial parameter optimisation, but also partly an indicator of efficient coding, so well done.

Overall a solid, well-organised and well-documented project that meets the project brief and executes reasonably efficient and nearly error-free code. Nicely done.

**Marks for the project and computational workflow**: 62%

## The Report

You have met the brief and written a rather commendable abstract/intro. However your report in general is quite short, and the discussion in particular is somewhat lacking, to the detriment of your overall report grade.

Title: Specific, informative, concise.

Abstract: Pretty good. Background and motivation present and touches immediately on mechanistic-phenomenological and model comparison. Methods and study objectives clear and results summarised. Take-home message (cautiously) proposed. (72%)

Intro: Pretty good. Background competently expanded upon with thorough reference to the literature. Study objectives and research question thereby well motivated, though a little more specificity about the approach employed in the study to achieve the objectives would be welcome. (72%)

Methods: Broadly adequate but with some flaws – a cubic polynomial does not have two peaks for example. No discussion of initial parameters for the logistic fit. Computing tools present. (58%)

Results: Relatively brief but the main results (number of fits, AIC and RSS values) are all present. Figure text is uncomfortably small. (62%)

Discussion: Summarises main findings and gives a straightforward interpretation of the AIC/RSS results. No discussion of the wider implications or links to the literature. It is generically mentionedthat e.g. larger sample sizes and additional statistical analyses might improve the robustness of the results, but there is no deeper discussion of caveats or shortcomings of the study or its methods. Take-away is present but somewhat noncommital. (55%)

(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**: 58%

**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!