

Calibrating Models to Data: A comparison of Methods

by

Zenabu Suboi

*Thesis presented in partial fulfilment of the requirements
for the degree of Master of Science in Mathematics in the
Faculty of Science at Stellenbosch University*



Department of Mathematical Sciences,
University of Stellenbosch,
Private Bag X1, Matieland 7602, South Africa.

Supervisor: 1. Dr. Marijn C. Hazelbag

2. Prof. Wim Delva

December 2020

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Signature:

Zenabu Suboi

Date: March 26, 2020

Copyright © 2020 Stellenbosch University
All rights reserved.

Abstract

Calibrating Models to Data: A comparison of Methods

Zenabu Suboi

*Department of Mathematical Sciences,
University of Stellenbosch,
Private Bag X1, Matieland 7602, South Africa.*

Thesis: MSc. (Mathematics)

December 2020

dddddddddddddddddddddddddddddddddddddd.

Keywords:

.....

Acknowledgements

I would like to express my sincere gratitude to the following people and organisations ...

Dedications

T.....

Contents

Declaration	i
Abstract	ii
List of Figures	vi
List of Tables	vii
1 Introduction	1
2 Methods	3
3 Simulation Procedure	4
4 Results	5
5 Discussion and Conclusion	6
5.1 Discussion	6
5.2 Conclusion	6
Appendix	7
List of references	7

List of Figures

List of Tables

Chapter 1

Introduction

Most sciences today use mathematical and computer simulation models to approximate the real-world processes under study [2, 4, 9]. For example, models play a significant role in health policymaking by estimating the impact of interventions in situations where empirical studies may be time-consuming, costly and impractical [7]. Developing a model calls for a trade-off between computational cost and accuracy; simple models require little computation time but can be a poor description of the real-world process, whereas complex models allow for a more accurate description of the process at the cost of increased computational cost. After model development, it is imperative to know how well the model represents reality. Model calibration, or fitting the model to data, increases the confidence that the model provides a realistic approximation to the real-world process [7, 9] but this process involves running the model many times. This becomes a problem with complex models for which model run-times are very long (Hladish).

Calibration is the process of comparing model outputs with empirical data to identify the model parameter values that achieve a good fit to data [5, 9]. Researchers commonly use calibration methods to find parameter values in case parameter estimates are not available in the literature [8]. The main components of calibration are; summary statistics (targets), the parameter-search strategy, the goodness-of-fit (GOF) measure, acceptance criteria and stopping rules. Improving the computational cost of the calibration involves specification of a more efficient parameter search strategy. In this study, we focus on sampling algorithms as the parameter search strategy, since sampling methods obtain valid estimates

of parameter uncertainty and correlations between parameters. Several more efficient sampling algorithms have been proposed and the number of studies that apply these algorithms is proliferating [9].

Existing literature compares the performance of alternative algorithms for calibrating the same model but does not allow us to draw general conclusions [1, 6]. [3] highlights the need for simulation-based studies that inform the choice of the parameter search strategy in terms of correct estimation of the posterior in different scenarios in terms of contextual variables (i.e. the number of target statistics, the number of calibrated parameters). the performance, strengths and limitations of different model calibration methods. algorithm implementation vs “How the algorithm works on paper”

Because there are many model calibration methods with little or no consensus on their performance, we perform a simulation study to compare the performance of model calibration methods using a simple stochastic Susceptible-Infected-Recovered (SIR) model. The methods to be compared are Rejection Approximate Bayesian Computation (Rejection ABC), Sequential Approximate Bayesian Computation (Sequential ABC) and Bayesian Maximum Likelihood estimation (BMLE).

Outline to be completed when thesis is fully written.

Chapter 2

Methods

Chapter 3

Simulation Procedure

Chapter 4

Results

Chapter 5

Discussion and Conclusion

5.1 Discussion

5.2 Conclusion

List of references

- [1] Issa J Dahabreh, Jeffrey A Chan, Amy Earley, Denish Moorthy, Esther E Avendano, Thomas A Trikalinos, Ethan M Balk, and John B Wong. A review of validation and calibration methods for health care modeling and simulation. In *Modeling and Simulation in the Context of Health Technology Assessment: Review of Existing Guidance, Future Research Needs, and Validity Assessment [Internet]*. Agency for Healthcare Research and Quality (US), 2017.
- [2] Anthony T Fojo, Emily A Kendall, Parastu Kasaie, Sourya Shrestha, Thomas A Louis, and David W Dowdy. Mathematical modeling of “chronic” infectious diseases: unpacking the black box. In *Open forum infectious diseases*, volume 4, page ofx172. Oxford University Press US, 2017.
- [3] C Marijn Hazelbag, Jonathan Dushoff, Emanuel M Dominic, Zinhle E Mthomboti, and Wim Delva. Calibration of individual-based models to epidemiological data: a systematic review. *medRxiv*, page 19006056, 2019.
- [4] Marc C Kennedy and Anthony O’Hagan. Bayesian calibration of computer models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(3):425–464, 2001.
- [5] Nicolas A Menzies, Djøra I Soeteman, Ankur Pandya, and Jane J Kim. Bayesian methods for calibrating health policy models: a tutorial. *Pharmacoeconomics*, 35(6):613–624, 2017.
- [6] Amanda Minter and Renata Retkute. Approximate bayesian computation for infectious disease modelling. *Epidemics*, 29:100368, 2019.
- [7] Natasha K Stout, Amy B Knudsen, Chung Yin Kong, Pamela M McMahon, and G Scott Gazelle. Calibration methods used in cancer simulation models and suggested reporting guidelines. *Pharmacoeconomics*, 27(7):533–545, 2009.

-
- [8] Elske van der Vaart, Mark A Beaumont, Alice SA Johnston, and Richard M Sibly. Calibration and evaluation of individual-based models using approximate bayesian computation. *Ecological Modelling*, 312:182–190, 2015.
 - [9] Tazio Vanni, Jonathan Karnon, Jason Madan, Richard G White, W John Edmunds, Anna M Foss, and Rosa Legood. Calibrating models in economic evaluation. *Pharmacoeconomics*, 29(1):35–49, 2011.