Data, models, and decision support

Course instructors
Dr. Amy Hurford, Biology Dept, and Dept of Mathematics & Statistics, Memorial University
Dr. Julien Arino, Dept of Mathematics, University of Manitoba (AARMS summer school)
Dr. Amy Greer, Ontario Veterinary College, University of Guelph (AARMS summer school)
Dr. Jane Heffernan, Dept of Mathematics and Statistics, York University (AARMS summer school)
Dr. James Watmough, Dept of Mathematics and Statistics, University of New Brunswick (AARMS summer school)

Course description Mathematical models provide decision support through forecasting, describing unrealized past outcomes (counterfactuals), and in communication. This course will teach the technical skills to parameterize dynamical systems models in ecology and epidemiology, and quantify prediction uncertainty. The course will examine case studies to understand the use of models in decision support.

Pre-requisites Students should have completed undergraduate courses in mathematical modelling and statistics, and be able to code in *R* (or have sufficient coding experience that this would not be a challenge).

Textbook Ecological forecasting (EF) by Mark Dietze.

Grading Assignments (40%), Final project (60%)

Rationale During the COVID-19 pandemic mathematical modellers responded to urgent requests for decision support. In a March 2023 presentation, Dr. Nicholas Ogden (Public Health Agency of Canada modelling team leader), described three types of contributions: forecasting, counterfactuals, and communication. To teach the skills needed for forecasting, we will consider *Ecological forecasting* by Mark Dietze as this book defines a framework applicable to classical population dynamic models (i.e., the logistic growth equation and beyond), and describes a careful mechanistic treatment of how to include uncertainty in models. Approximately, 90% of the course content is focused on learning the skills to formulate, parameterize, fit, and assess forecast models. The remaining course content considers decision support with counterfactuals, and investigates how modelling was used to support communications during the COVID-19 pandemic.

Course outline

[1] Understanding the past, present, and future of modelling to support decision making. Readings: Ch. 1. of EF, Ch. 1 of *A Biologists Guide to Mathematical Modelling* by Sarah P. Otto and Troy Day (available as a .pdf from the MUN library). *All models are wrong?* by Enderling and Wolkenhauer; Five ways to ensure models that serve society: a manifesto by Saltelli et al. Simulating the pandemic: what COVID forecasters can learn from climate models by Adam.

[2] From models to forecasts. Describes how classical models in population biology can be extended to forecast models. Ch 2 in EF.

[3] Parameter estimation from the literature. During the COVID-19 pandemic many modelling approaches estimated model parameters from the literature. We will study the parameterization of *Projected effects of nonpharmaceutical public health interventions to prevent resurgence of SARS-CoV-2 transmission in Canada* by Ng et al.

[4] Introduction to Bayes. Bayesian analysis provides a framework to combine models and data using probabilities. Includes: The Likelihood; Numerical methods for Bayes; Evaluating MCMC output. Ch 5 in EF.

[5] Characterizing uncertainty. Bayesian methods are used to characterize and partition sources of uncertainty moving beyond the classical assumptions of independent, homoscedastic (constant variance), and Normal residual error. Includes: Observation error; Missing data and inverse modeling; and Hierarchical models and process error. Ch 6 in EF.

[6] Latent variables and State-Space Models. A dynamic modelling framework applicable when the variable of interest is unobserved, or observed with error. Includes: Latent variables; State-space models; Hidden Markov Time-Series Models. Ch 9 in EF. See also Towards an improved understanding of causation in the ecological sciences by Addicott et al.

[7] Propagating uncertainty. Describes techniques to propagate uncertainties into forecasts and diagnose what is driving the uncertainty. Ch 11 in EF.

[8] Data assimilation. Updating forecasts given new information. The forecast cycle. Kalman filter and Particle filter. Ch 13. In EF.

[9] Assessing model performance. Considers visualization and data mining residuals. Ch 16 in EF.

[10] Models for decision support. Considers the use of models to support decision-making during the COVID-19 pandemic. Examples: flatten the curve, R_t/doubling time, herd immunity, total mortality for highly transmissible variants and the use of counterfactuals and forecasting. Ch 17 in EF. See also Special Report: The models deriving the world's response to COVID-19 by Adam.

Possible assignments

[1] A critical assessment of modelling to support decision-making. Students should define important components and discuss challenges and future directions with cited literature and/or example models and policy questions.

[2*] Perform Monte Carlo simulations for an ordinary differential equation model with uncertainty in the parameter estimates and uncertainty in the initial conditions. Show 90% and 95% confidence intervals for the results. [based on this]

[3*] Parameter estimation using Bayesian analyses and Markov Chain Monte Carlo (MCMC). Evaluate the MCMC chain for convergence. Report parameter summary table and plot marginal distributions. Describe and explain the parameter covariances. [based on this]

[4] Students review the source publications for parameter estimates, critiquing the application to the model for the policy question, and/or quantify the distribution of parameter values from the published literature. Students might independently parameterize an already parameterized model from the literature and provide a written evaluation of the agreement.

[5] How does the parameterization approach of Ng et al. compare to the methods described in Ch. 9 Fusing data sources of EF?

[6*] Implement a state-space model for time series data. This modelling approach considers an observation error model, and a process model. This exercise builds in complexity of the process model: a) a random walk (i.e., the future state is normally distributed around the current state); b) a linear dynamical system; to c) a logistic equation (non-linear). [based on this]

[7*] Students will incrementally add sources of uncertainty to a state-space model to understand the propagation of uncertainty in forecasts. Sources of uncertainty are: initial condition, parameter, extrinsic driver, process, and random effect. [based on this]

[8*] A Kalman filter assignment. [based on this]

[9] A particle filter assignment. [based on this]

[10*] Students will assess model performance using visualization and statistical methods. Considers Taylor diagrams, patterns in Bayesian p-values, Continuous Rank Probability Score, mining the residuals. [based on this]

[11*] A decision support assignment.

*Assignments that will be completed.