The case of the Ontario Rabies Model

Antoinette Ludwig[1,3]; Philippe Berthiaume[1,3]; Michel Bigras-Poulin[2,3], Rowland Tinline[4]

1 Public Health Agency of Canada – 3200 Sicotte C.P. 5000- Saint Hyacinthe QC J2S 7C6
2 Faculty of Veterinary Medicine – Universiy of Montreal - 3200 Sicotte C.P. 5000- Saint Hyacinthe QC J2S 7C6
3 Groupe de recherche en épidémiologie des zoonoses et santé publique - 3200 Sicotte C.P. 5000- Saint Hyacinthe QC J2S 7C6
4 Department of Geography, Queen's University, Kingston, ON, K7P 3C7

Abstract: Dynamic mathematical modeling and stochastic simulation of disease-host systems for the purpose of epidemiological analysis offer great opportunities for testing hypotheses, especially when field experiments are impractical or when comparing multiple experimental scenarios. This, combined with the ever increasing computer power available to researchers, has contributed to the development of many mathematical models of epidemics. In the particular case of spatially distributed dynamic stochastic models, such as the Ontario Rabies Model (ORM), it is important to determine the impact of the intrinsic variability produced by the stochastic mechanism used. Properly assessing the behaviour of such models requires an adequate testing protocol and sufficient simulation runs. Here we present a procedure for testing for directional bias in the ORM, a model of raccoon rabies epidemic spread, in absence of clustered heterogeneity. We did this by comparing spatial descriptive statistics of 100,000 epizootics resulting from running 100 iterations for every one of 100 unique baseline populations all using an identical target carrying capacity over the entire study region. This procedure permitted the discovery of an anomaly in the behaviour of the model for which the origin could later be found in the source code. Another aspect of this work consisted in defining a framework to determine the number of simulations required to capture enough model variation depending on the objective of the chosen research question by potential end-users. We observed that with only 100 simulations (10 populations and 10 trials), we were able to observe output that captured much of the variation in the 100,000 epizootics but, as the range of output variation is very large , doing more simulations would be appropriate if we were concerned about identifying extreme results. Both aspects constitute important factors likely to influence the level of use this model or other models will receive from their targeted end-users. We conclude that although validation is a resource hungry task, it should pay off by providing a better credibility to the models and inform the intended users on the limits inherent to the model.

Background

Common challenges to using a spatio-temporal stochastic model

- Modelling imperfectly represents biological processes
- Difficulties in interpretation of **variability** resulting from profusion of results

All of these may cause a lack of confidence or ownership by targeted end-users…

Solutions

- Thorough **assessment of model behaviour**
- Choosing appropriate experimental designs to find an appropriate compromise between "lower effort" scenarios (minimizing the number of trials/output, minimizing the simulation time) and sufficient representation of complexity

Objectives

- Providing a simple intuitive methodology for **validating model behaviour** against directional spatial bias in stochastic spatially distributed individual-base models
- Presenting a **framework for optimizing time and resources** required

Methods

Our experimental vehicle: the Ontario Rabies Model (ORM)

- Stochastic model
- Models the spatio-temporal dispersion of rabies in a raccoon population
- Tracks individual raccoons moving through hexagonal grid and infecting animals in each grid cell

Simulation strategy

- Size and shape: theoretical hexagonal study area (radius = 116 km)
- Random generator used: linear congruential
- 100 baseline populations started from an Adam & Ève couple
- 100 epizootics (trials) for each starting population
- Outbreak initiated in the central cell
- Duration of simulation: 3.5 years
- Homogeneous carrying capacity over the entire study area (≈10 raccoons /km^2)
- Capture data on rabies cases and population at the individual cell level

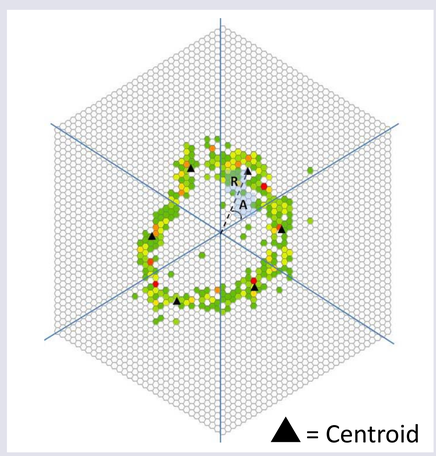

Fig.1 Spatial analysis framework

Output variables

- The polar coordinates ((R)adius and (A)ngle) for the weighted mean location (centroid) of the number of cases by sector and time for 100 starting populations and 100 trials for each starting population (10,000 epizootics)

$$R = \frac{\sum w_i R_i}{\sum w_i} \qquad A = \frac{\sum w_i A_i}{\sum w_i}$$

With w_i = number of cases located in the same ORM cell
R_i = Radius for the weighted mean location
A_i = Angle for the weighted mean location

Analysis

- Visual analysis of plots showing centroids at 6 month intervals, i.e. 0.5, 1,... 3.5 years.
- Comparison of the mean value of the radius for the cases by sectors at six months intervals
- Comparison of the range of variation of the radius for the cases by number of baseline populations and number of trials used for the simulation

Results

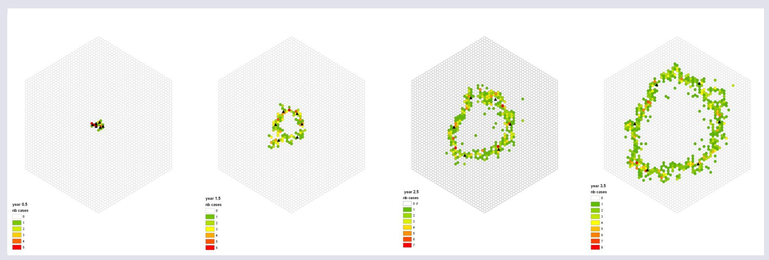

Fig. 2 Evolution of the density of rabies-infected raccoons

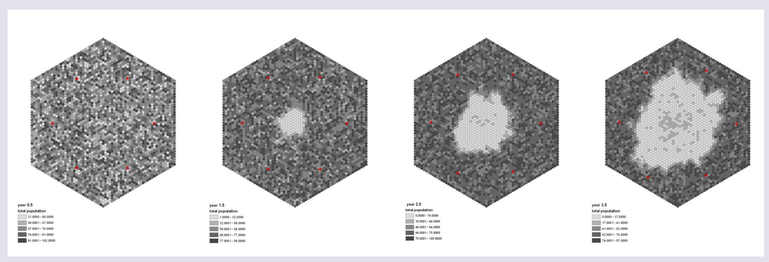

Fig3. Evolution of the density of total raccoon population

Number of simulations and variability in the outputs

- Increasing the number of populations and/or trials increases the range of variation of the observed values of the centroid radius
- The **maximum variation was obtained with 30 populations and 40 trials** – an important result that is repeated from year 1 to year 3.5
- After 3.5 years of simulation, the **variation of extension of the epidemic wave can be as high as 70 km**

Model behaviour assessment

- The centroid points are **representative of a typical direct contact disease spread** (Figures 1-3)
- The epizootic behaviour is supported by field observations
  - ➢"Donut" shape epidemic centered on the first infected cell (Grenfell, 2002)
  - ➢ Depletion of the population behind the front (Rosatte, 2007)
- The epidemic seems to **run slower in sector 2 and 3** than in other sectors – The difference in radius between sectors can be as high as 13 km after 3.5 years of simulation (Fig.4)

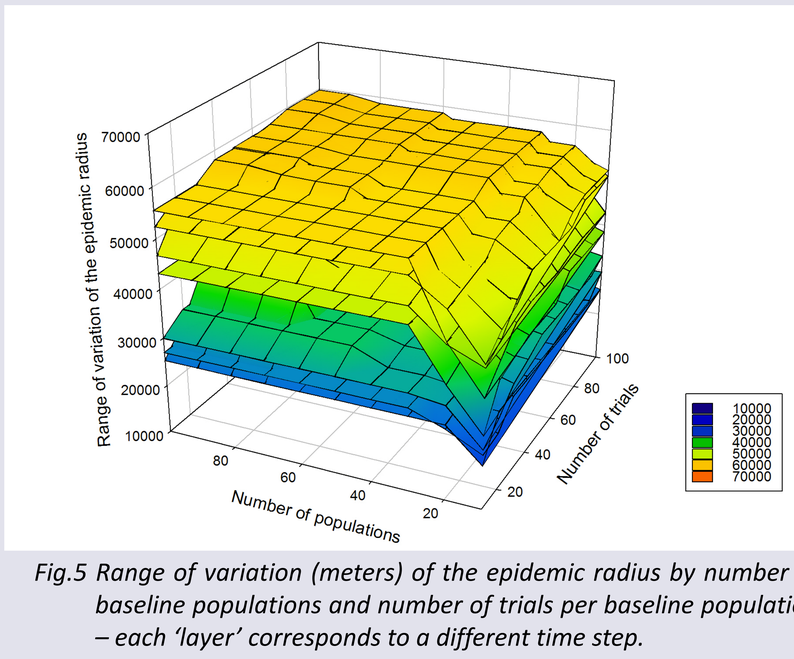

Fig.5 Range of variation (meters) of the epidemic radius by number of baseline populations and number of trials per baseline population – each 'layer' corresponds to a different time step.

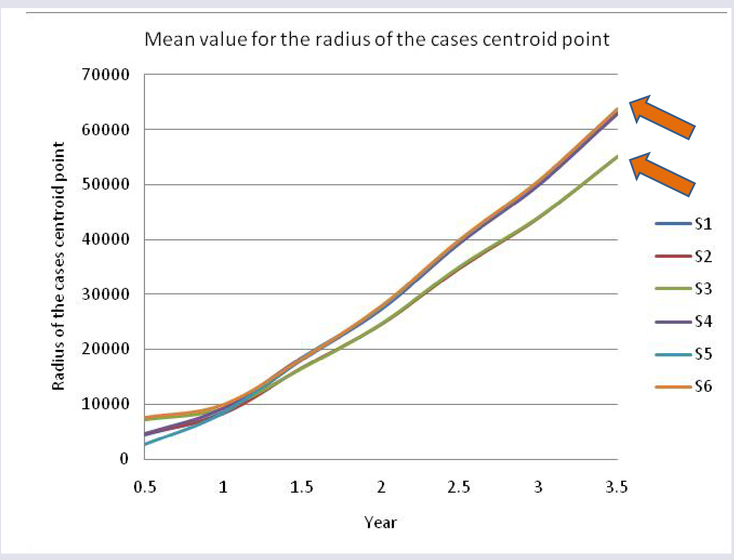

Fig.4 Mean evolution of the disease spread through time by sector

Number of simulations and computer resources

- Simulations in a Windows XP environment are computer intensive (2 months using 5 medium performance computers)
- Adaptation of the ORM code for parallel processing and use in Unix clusters is presently being performed in hope of greatly **reducing simulation time**.

Discussion

- The **geometric method used proved to be very effective** in identifying directional bias
- The observed directional bias has lead to a revision of the computing code
- The directional bias could not have been identified easily without a large number of simulation and geometric analysis of the output
- For most practical purposes experiments **with only 10 starting populations and 10 trials generate around 95% of the variation observed** with larger numbers of populations and trials
- Having 2 versions of the ORM code provide more flexibility for the user depending on his research objective:
  - The Linux version provides an optimal environment for large scale simulation - Large scale simulations are important if the objective is to capture all the variability in the outputs
  - The Windows version is still appropriate for small scale simulations run to capture the mean behaviour of the experience

Conclusions and perspectives

- Modelling biological systems often results in sophisticated models for which the diagnosis of error may be particularly difficult
- Model behaviour validation should be implemented as a standard step to increasing confidence and use by end-users
- It also helps in rationalizing resources when answering scientific questions.

Acknowledgements: Many thanks to the Ontario Ministry of Natural Resources Rabies Research Unit under whose auspices the Ontario Rabies Model was developed in collaboration with Queen's University. Many thanks to Jacques Richer, from the RQCHP, and David Ball, the creator of the ORM code for their careful and expert work in recompiling the ORM in Linux. Thanks to Dre Denise Bélanger, UdM, for her support to the project

Litterature:
Grenfell, B., 2002, Rivers dam waves of rabies, PNAS (99) 3365-3367;
Rosatte, R. and Al, 2007 The elimination of raccoon rabies from Wolfe Island, Ontario: Animal density and movements, Journal of Wildlife Diseases (43) 242-250.
Réseau Québécois de Calcul Haute Performance (RQCHP) (https://rqchp.ca/)