Environmental Modelling & Software 77 (2016) 1–12

Contents lists available at ScienceDirect

Environmental Modelling & Software

journal homepage: www.elsevier.com/locate/envsoft

Improving the computational efficiency of an agent-based spatiotemporal model of livestock disease spread and control

R.A. Bradhurst [a], [*], S.E. Roche [b], I.J. East [b], P. Kwan [a], M.G. Garner [b]

[a] *School of Science and Technology, University of New England, Armidale, NSW, Australia*
[b] *Epidemiology and One Health Program, Animal Health Policy Branch, Department of Agriculture and Water Resources, Canberra, ACT, Australia*

ARTICLE INFO

Article history:
Received 27 February 2015
Received in revised form
10 November 2015
Accepted 23 November 2015
Available online 11 December 2015

Keywords:
AADIS
ABM
FMD
Spatial indexing
Spatiotemporal model

ABSTRACT

Agent-based models (ABMs) are well suited to representing the spatiotemporal spread and control of disease in a population. The explicit modelling of individuals in a large population, however, can be computationally intensive, especially when models are stochastic and/or spatially-explicit. Large-scale ABMs often require a highly parallel platform such as a high-performance computing cluster, which tends to confine their utility to university, defence and scientific research environments. This poses a challenge for those interested in modelling the spread of disease on a large scale with access only to modest hardware platforms.

The Australian Animal DISease (AADIS) model is a spatiotemporal ABM of livestock disease spread and control. The AADIS ABM is able to complete complex national-scale simulations of disease spread and control on a personal computer. Computational efficiency is achieved through a hybrid model architecture that embeds equation-based models inside herd agents, an asynchronous software architecture, and a grid-based spatial indexing scheme.

1. Introduction

Mathematical models of disease spread have been in use since at least the 18th century (Dietz and Heesterbeek, 2002). An example is a Susceptible, Exposed, Infectious and Recovered (SEIR) compartmental equation-based model (EBM) that uses a system of ordinary differential equations (ODEs) to predict infection state proportions in a population over time (Keeling and Rohani, 2008). The population is dynamically disaggregated into the SEIR compartments; however, individuals within any particular compartment are indistinguishable. Models such as this are termed population-based, in that top-down population-level relationships provide insight into individual-level states. Population-based models can be concise and computationally efficient, but generally assume homogeneous contact rates and susceptibility, i.e., individuals mix uniformly and randomly, and have an equal likelihood of contracting a disease. The homogeneous 'well-mixed' assumption of an aggregated population-based model is a limitation if the population is heterogeneous and mixes heterogeneously (Hethcote, 2000; Keeling et al., 2001; Kitching et al., 2006; Bansal et al., 2007; Grassly and Fraser, 2008).

Agent-based models of disease spread are formulated from the bottom-up, whereby population-level relationships emerge organically from the aggregation of individual-level behaviours. Agent-based models are well suited to capturing population heterogeneity, stochasticity, spatial relationships, adaptivity, social systems and policy elements (Parunak et al., 1998; Davidsson, 2001; Hare and Deadman, 2004; Crooks and Heppenstall, 2012). The explicit modelling of individuals in a population, however, can be computationally intensive, especially for large populations. Computational efficiency is important for models that employ Markov chain Monte-Carlo (MCMC) methods (Hamra et al., 2013). A stochastic model may be called upon to re-run a scenario with the same initial conditions thousands of times to allow trends to emerge from the underlying probabilistic mechanisms (Driels and Shin, 2004). Spatiotemporal models have additional computational challenges of efficiently managing spatial objects, spatial relationships and spatial queries (Kennedy et al., 2009). Large-scale agent-based models can require custom software implementations (Parker and Epstein, 2011), and highly parallel platforms such as high-performance computing (HPC) clusters (Carley et al., 2006; Germann et al., 2006) or general purpose computing on graphics

---

* Corresponding author. School of Science and Technology, University of New England, Armidale NSW 2351, Australia.
E-mail address: rbradhur@myune.edu.au (R.A. Bradhurst).

http://dx.doi.org/10.1016/j.envsoft.2015.11.015

processing units (Lysenko and D'Souza, 2008; D'Souza et al., 2009; Welch et al., 2014). Access to HPC clusters is largely in the hands of university, defence and scientific research organisations. This poses a challenge for those interested in modelling the spread of disease efficiently on a large scale with access to only standard hardware platforms.

Epidemiological models are used in Australia to support planning and policy development for exotic animal diseases such as foot-and-mouth disease (FMD) (Garner et al., 2007, 2014; Roche et al., 2014). FMD is a highly contagious disease of cloven-hoofed animals that significantly impacts livestock production and trade in livestock and livestock products (Buetre et al., 2013). Modelling the spread and control of FMD is complex as the virus has multiple serotypes, and spreads via multiple pathways (direct contacts, indirect contacts and aerosol plumes), to multiple host species (including cattle, sheep, goats and pigs). The environment of a potential outbreak is also complex as there is considerable heterogeneity in Australian livestock production systems, market systems, geography and climate (Animal Health Australia, 2014a). Further, disease control in Australia is managed by the individual states and territories (Animal Health Australia, 2014b), so for national planning purposes, a model needs to handle jurisdictional differences in the implementation of control programs. AusSpread is a spatially-explicit, farm-based state-transition microsimulation developed by the Australian Department of Agriculture and Water Resources to study FMD (Garner and Beckett, 2005). It is based on the MapBasic/MapInfo geographic information system (GIS) platform (Pitney Bowes, 2015). Runtime constraints limit AusSpread's practical use to studying disease at a regional level.

The Australian Animal DISease (AADIS) model is a national-scale epidemiological model developed by the Australian Department of Agriculture and Water Resources (Bradhurst et al., 2013, 2015). The AADIS model is written in Java (Oracle, 2015), and employs open-source products such as OpenMap (BBN, 2015), PostgreSQL (PostgreSQL, 2015) and SQL Power Architect (SQL Power Group, 2015). A national-scale epidemiological model must be computationally efficient while addressing the needs of disease managers in terms of capturing the disease epidemiology, regional variability in transmission (for example, due to different livestock movement patterns, production systems and climates), and jurisdictional approaches to control. In this paper we present the key design features that allow the AADIS model to run complex national-scale simulations efficiently on a personal computer: a hybrid agent-based model (ABM) architecture that embeds EBMs inside herd agents; an asynchronous software architecture that features lightweight agents in an active concurrent environment; and grid-based spatial indexing. We describe the performance gains achieved through the use of uniform grid-based spatial indexing instead of R-Tree-over-GiST spatial indexing provided in the PostGIS (OSGeo, 2015) extension to PostgreSQL. We also compare the runtime performance of the AADIS ABM with the AusSpread microsimulation and discuss some advantages and disadvantages of concurrent processing over the sequential approach taken by many microsimulations such as AusSpread.

2. Material and methods

2.1. Hybrid model architecture

The main Australian industries vulnerable to an FMD outbreak are beef, dairy, wool, sheep meat, and pigs. This implies an FMD-susceptible national population in excess of 100 million animals (Australian Bureau of Statistics, 2014). It is possible to derive realistic individual-based contact networks from identification and tracing systems such as the Australian National Livestock Identification System (Meat and Livestock Australia, 2014). Australian cattle, for example, undergo mandatory tagging with a radio frequency identification ear tag or rumen bolus that facilitates per-animal tracking from property of birth to place of slaughter/export. It is, however, computationally infeasible for a personal computer to accommodate an ABM with 100 million agents. The AADIS ABM, instead, employs the herd as the modelling unit of interest, yielding a national population of approximately 236,000 FMD-susceptible herds. This is a reasonable simplification as livestock are typically managed as single-species herds that effectively share a single contact network whilst on a farm (Kostova-Vassilevska, 2004). A herd is assumed to be homogeneous with respect to species and farming practices, and well-mixed from a disease transmission point of view. This implies that any one member of a herd has the same likelihood of contracting a disease as any other member. The number of animals in a herd is simplified to be constant over time, that is, births and incoming consignments are assumed equivalent to natural deaths and outgoing consignments. Modelling the spread of disease on a per-herd basis captures heterogeneity within multi-species farms. For example, farms that manage both sheep and cattle can be modelled as comprising multiple independent herds with distinct disease dynamics and animal management practices. An example national dataset of FMD-susceptible herd and farm types is presented in Table 1.

Each herd agent has an embedded ODE-based SEIR EBM that deterministically predicts the infected, infectious and clinical prevalence of the herd over time. The parameterisation of the EBM ODE system is dependent on the strain of FMD, the relative infectiousness and susceptibility of the species, and the production system (which influences the degree of contact between animals). Modelling the spread of disease within a herd deterministically is reasonable for such a highly contagious disease as FMD, that once introduced into a susceptible herd will typically progress

Table 1
Farm and herd types used in the AADIS ABM.

Farm type	Number of farms	Number of animals mean (min – max)	Herd type	Number of herds
Extensive beef	1331	1909 (1200–46,575)	Extensive beef	3993
Intensive beef	51,383	280 (30–7436)	Intensive beef	51,383
Feedlot	508	1825 (100–39,963)	Feedlot	508
Mixed beef/sheep	21,556	242 (30–5700)	Mixed beef	21,556
			Mixed sheep	21,556
Dairy	8675	298 (40–2742)	Dairy	8675
Small pigs	1873	244 (40–4850)	Small pigs	1873
Large pigs	333	4922 (1000–17,896)	Large pigs	333
Sheep	22,150	1649 (20–44,000)	Sheep	22,150
Small holder	103,641	5 (1–14)	Small holder	103,641
Total	**202,775**			**235,668**