Increasing model efficiency by dynamically changing model representations

Randall Gray
CSIRO Division of Marine and Atmospheric ResearchWotherspoon
University of Tasmania

February 17, 2010

1 Introduction

Models of the effects of human interactions with the environment and with animal and plant populations, are often are faced with dilemma of choosing an appropriate scale at which to model the interactions. Population level effects are likely to be modelled with an analytic approach, but if our system exhibits behaviour which is incompatible with mean-field assumptions an individual-based or agent-based approach to modelling the system may be more appropriate.

There is a body of literature stretching back three decades which discusses individual-based modelling as an effective approach to modelling systems where individual variability is perceived to be an important driver in the system's dynamics, notably JABOWA and its derivatives (Botkin, *et al.* 1972a, Botkin, *et al.* 1972b). Since the early 1980s, a rapidly increasing number of significant papers and books have appeared which address the use of individual-based models across a broader base and with discussion of the relative strengths and weaknesses of the approach (such as Huston *et al.* (1988); DeAngelis and Gross (1992) and Grimm and Railsback (2005)). The use of classical (equation based) models to explore populations and ecological systems goes back much further, arguably to 1798 with Malthus's ***An Essay on the Principle of Population***.

In some sense these approaches represent the extrema in a spectrum which exhibits, through its range, varying degrees of aggregation and resolution in time, space, and membership. Models of one extreme or the other are obvious representatives of a "choice of paradigm," but there are a large number of models which incorporate representatives from both extrema, and indeed adopt intermediate (super-individual) representations, such as described in Scheffer et al. (1995).

As ecosystem models become broader in scope, including more species and richer environments, the notion that a model can be identified with some particular region of this spectrum starts break down. Simulation models often embed the subject of study within an "environment" made up of primary data and other processes within the system. These data and processes which comprise the environment may lie in quite a number of places between purely equation-based representations and an individual-based representations. The actual implementation may be as a set of quite distinct submodels which are coupled together but retain their independence and stand, in some sense, as models in their own right; the submodels may be seamlessly integrated with the aggregated model presented, like an organism, as an indivisible entity.

There are quite a number of well established ways of integrating submodels, many of which achieve efficiency, either in run-time or accuracy, by running different components of the model at different temporal or spatial scales such as variable-speed splitting (Atlantis?, EwE?), and

model nesting (forest community models nested in annual/seasonal environment models?), and many of the numeric techniques we routinely apply in our models use adaptive steps in their solution (e.g. Runge-Kutta and root finding).

As the climate changes and the environment we live in and use drifts further from its familiar state, there is a corresponding need to manage human interaction with ecosystems more carefully. The scope of ecosystem models is steadily increasing (DeAngelis *et al.* (1998), Harvey *et al.* (2003), Fulton *et al.* (2004), Gray *et al.* (2006)). Models are including more functional groups within the ecosystem and the interactions between components are becoming more detailed. Addressing the increased demand for detail is costly in terms of computational load: individual-based models of populations may be very good at capturing the vulnerability to exceptional events, but such simulations take a long time. Worse still, much of this time may be spent with the model in a largely unchallenging or uninteresting part of its state-space.

We might consider the motion of a mass capable of acceleration as a simple example; perhaps a rocket or a taxi. The rules of Newtonian motion make a good model as long as our mass doesn't move too quickly, and, while relativistic motion is much more accurate at high speeds, it is more expensive to calculate. A simulation of the spacecraft (or taxi) that approaches the speed of light would spend a lot of time calculating relativistic motion when a simpler model would be adequate. It is easy to imagine that we might use Newton's equations when the velocities are low (most of the state space), but shift to relativistic motion when we stray from the areas where Newton is comfortably accurate.

So the notion we explore in this paper is that we might change the *representation* of a submodel based on its location in its state-space, and that by doing so we may actually be able to simulate our system more effectively. To some degree we do this in our simulation models anyway: time-steps or spatial resolutions are changed, particular code paths may be by-passed according to local conditions within the simulation, or additional calculations might be performed to reduce the error when the state is changing rapidly; but these optimisations are largely optimisations of the *encoding* of the model or submodels, rather than an actual change in representation. In contrast, moving from a population modelled by a set of differential equations to a model tracking the life-history of a large number of individuals involves a change of domain, state-space and changed assumptions concerning the homogeneity of the consituent members of the population being modelled.

This approach to the problem has been developed in the light of experience with several large scale human-ecosystem interaction models (Lyne *et al.* (1994), Gray *et al.* (2004), and now the modelling of the interactions thein a study of the region containing the Ningaloo Marine Park). In each of these studies a significant component has been simulating the interaction between organisms and contaminant plumes. The focus in Lyne *et al.* was the potential for the percolation of contaminants in industrial waste up through the food chain into commercially exploited fish stocks. Gray *et al.* developed a regional model focussed on managing the impact of human activity on the biological systems along the Northwest Shelf of Australia, one component of which was the effects of flushing the bitterns from salt production.

Monte (2009) presents a lucid discussion of *contaminant migration-population effects* models. These models incorporate the movement of populations and their internal distribution, the transport of contaminants through the system via biotic and abiotic pathways, and the changes in behaviour and population dynamics associated with contamination. Monte discusses conditions which make a method of coupling the equations which govern contaminant dispersion and population dynamics, but states that the class of systems where the "movement of animals, the death and birthrates of individuals in $\boldsymbol{x}$ at instant t depend on previously occupied positions" is not generally amenable to this approach, and suggests that repeated simulations of many individuals is an appropriate way of dealing with this situation.

The treatment of contaminants, in particular, is expensive in terms of run-time and memory use – in the model described in Gray *et al.* (2004) contaminant modelling increased the time taken by roughly an order of magnitude. In all of these studies a considerable amount of time was spent in regions where no interaction with contaminant plumes was possible.

Running a complex model and maintaining its state in a region where a simple simple model may perform better imposes a burden which is probably unnecessary. There is potential for both significant improvements in run-time and the reduction of error. There are four basic questions that arise: When should a model change representation? What data needs to persist across representations? How is the initial state for a new representation constructed? How should the error associated with the loss of state information to be managed? In general, the answer to these questions are likely to be specific to the ensemble of submodels in question. Before expending the resources and effort required to implement such a strategy in a large scale model there needs to be, *prima facie*, a demonstration that the notion is worth pursuing. The aim of this paper is to provide this demonstration rather than to develop a body of techniques supporting the approach.

We present a model of organisms moving along a simple migratory path which intersects a region with a field of contamination with varying levels. This model exhibits fundamental attributes of larger studies of pollutant/ecosystem interactions (Lyne, *et al* (1994), Gray, *et al* (2006)), and while it isn't intended to accurately represent any particular system, it might loosely correspond to some body of water influenced by contaminant loads associated with terrestrial runoff resulting from intense rainfalls. Its role is to help explore some of the issues associated with changing representations and to test the hypothesis that we might gain efficiency in simulation time, representational accuracy, or both.

Our model uses basic representations of individuals and populations as the base-line representations and the mutating model switches between them when appropriate. The population model makes the assumption that the contaminant contact with the population is distributed through the population according to the density function which describes the population and that this is adequate to characterise the uptake based on the likelihood of an individual being at any given location within the population's radius.

2 Experimentation

The comparison of the three modelling schemes is based on the following:

i. the run-times for each representation

ii. the correspondence between the total contaminant loads across the biomass through time

iii. a comparison of variability in contaminant load through time amongst the representations

iv. and the sensitivity to variability in the nature of the plume.

Our strategy is to first establish the equivalence of the results from the mutating submodel representation and the individual-based representation. This component of the study also provides us with data for comparing the relative efficiency of the representations in terms of run-time. We generate forty trials each simulating forty individuals for both the homogeneous individual-based model and the mutating model. The homogeneous population model is deterministic and one run is sufficient.

A second set of trials compare the performance of the mutating model against the homogeneous population model when the contaminant plume is elliptical. Ideally, the results with an

elliptical plume would be consistent across the model representations, though the results will show this is not the case.

2.1 The models of organisms

The test models are composed of one or more submodels which run within a simple time-sharing system. Each submodel runs for a nominated period of time and passes control to the next submodel, very much like tasks running in many modern computer operating systems.

The population-based and individual-based submodels have been kept as similar as practicable in order to minimise the sources of divergence. Both of these basic representations are used to model a group of organisms which proceed around a common circular migratory path on an annual basis. While both submodels use the same uptake-depuration model to calculate their contaminant load, the calculation of the contact with the contaminant differs since the interactions of an individual moving through a plume is, in some sense, an integration across a path, while the population's expected interactions are more analogous to the integration over the domain of the population.

The individual-based representation models the group as a set of individuals (agents, or instances of submodels) which move in a directed random walk. At each time step the trajectory of each individual incorporates a directional component toward the "target location" on the migratory path which corresponds to the notional ideal centre of the group at that time. Each agent maintains its own contaminant load and position and velocity, and these attributes are independent of the state of the other agents.

Our population-based approach models the organisms with a radially symmetric distribution whose centroid lies on the migratory path. The density of the simulated individuals around their target location arising from their movement model closely follows the probability density function of a normal distribution. From this representation of the individuals' density, we define the population's density, $\rho(\boldsymbol{p})$, to be the value of the standard 2D normal distribution with a mean of zero in each ordinate and a variance. So our population's density at $\boldsymbol{p}$is given by

$$\rho(|\boldsymbol{p}|) = S_{L} \frac{1}{2\pi\sigma^2} \exp\left(-\frac{p_x^2 + p_y^2}{2\sigma^2}\right),$$

where $\boldsymbol{p} = (p_x, p_y)$ is the position in the population disk relative to its centre, S_L is a scaling parameter (equal to 1.015) which ensures that the population density, ρ, integrated over the finite disk with a radius, τ, is one, and σ^2 is the variance for the distribution. The effective radius of the population, τ, was taken to be three times the variance in the distance of the individuals from $\boldsymbol{m}(t)$, which is about 9.5% of the migratory radius. If the distribution of individuals around the point $\boldsymbol{m}(t)$ and the distribution of the population around $\boldsymbol{m}(t)$were not consistent, then the contaminant exposure of the two submodels would be incompatible and the sort of mutation error discussed in the introduction becomes significant.

In contrast to individuals, populations do not need to maintain position or velocity since both are determined by $\boldsymbol{m}(t)$. They do, however, maintain the number of organisms they represent and, in order to make mutation simpler, populations maintain a list of contaminants. In the homogeneous population model the length of this list never exceeds one, since the population has only one mean, but in a mutating submodel, the contaminant loads of the individuals are carried along and decayed in the usual way.

2.2 The migratory circle and the plume

The migratory path is a circle which can be defined by the equation

$$\boldsymbol{m}(t) = R(\cos(2\pi t/Y), \sin(2\pi t/Y)),$$

where $\boldsymbol{m}(t)$, denotes the "location at time t" on the migratory path, R is the radius of the migratory path and Y is the period. This circle is used for all the variant submodels of the system.

The plume is modelled as a circular or elliptical cloud of some contaminant which intersects this path and has an intensity which varies sinusoidally at a frequency which is relatively prime to the annual migration cycle.

Its centroid is positioned on the migratory circle. The plume is modelled as an intensity at a location with an attenuation function. The intensity at a point, $\boldsymbol{r}$, within the plume at time t is given by

$$I(t, \boldsymbol{r}) = \frac{1}{2}(1 + \cos(2\pi t/p)) \exp(-\lambda\phi(\boldsymbol{r}, \boldsymbol{m}(t))),$$

where p is the period of the plume, λ is a decay exponent, and for a circular plume the distance function ϕ is taken to be $\phi(\boldsymbol{a}, \boldsymbol{b}) = |\boldsymbol{a} - \boldsymbol{b}|$. The intensity is made elliptical by instead taking $\phi(\boldsymbol{a}, \boldsymbol{b}) = \sqrt{(\boldsymbol{a} - \boldsymbol{b}) \cdot (\sqrt{2}, \frac{1}{\sqrt{2}})}$ since this transformation preserves the area of the plume. The effective radius of the circular plume in the model is about 3.7% of the radius of the migratory circle, R, and the period, Y is taken to be a year.

2.3 Contaminant load and contact

The model for uptake and depuration of our contaminant is the same for both representations. Initially a "total contact" is calculated for the time step. As might be expected, the total contact for a population is calculated in a different way to the total contact of an individual. In either case, this resulting contact is fed through a standard uptake-depuration model

$$dC/dt = uM - \lambda C$$

which is solved numerically with a fourth order Runge-Kutta algorithm for the value of C given a contact mass, M, and an initial contaminant value or vector of values for C.

For individuals, the mass of contaminant which is available for uptake, or contact, is taken to be the result of integrating the intensity of the plume over the path of the individual, $\boldsymbol{P}_t$ to $\boldsymbol{P}_{t+\delta}$,

$$M = \int_{\boldsymbol{P}_t}^{\boldsymbol{P}_{t+\delta}} I(\boldsymbol{p})||\boldsymbol{p}||d\boldsymbol{p}$$

where we assume that the motion between (t, ℓ_t) and $(t + \delta_c, l_{t+\delta})$ is linear. Populations are somewhat more complex. Here, we calculate the definite integral

$$M = \int_{\boldsymbol{P}_t}^{\boldsymbol{P}_{t+\delta}} 2 \int_{\boldsymbol{\Omega}} I(\boldsymbol{p}+\boldsymbol{\omega})\rho(\boldsymbol{p} + \boldsymbol{\omega})d\boldsymbol{\omega}d\boldsymbol{p}$$

where $\boldsymbol{\Omega}$ is an area over which we assess the effective area of the population and $\boldsymbol{p} + \omega$ denotes the area $\boldsymbol{\Omega}$ translated so that its centroid corresponds to $\boldsymbol{p}$. The contact equations are solved using a simple adaptive quadrature routine.

3 The mutating submodel

The mutating model is formed by interleaving the population representation and the individual-based representation. The transitions between a each submodel and its alternate representation are dictated by its proximity to the contaminant plume. Initially the model begins with the group of organisms represented as a population. The starting location is far enough away from the plume that there is no chance of any interaction between organisms represented by the population and the contaminant field. As the population moves about the circle it reaches a point at which there is a likelihood that sometime in the next few time steps an individual within its disk could conceivably encounter contaminants, and at this point the population is disaggregated into submodels representing individuals. Similarly, an individual-based submodel switches to a population submodel when there is no possibility that it will encounter the plume. These decisions obviously rely on knowledge of the system; specifically the static location of the plume, the rates of movement and the length of time steps.

The constraints posed in the presented situation aren't as unreasonable as it might seem: many of the sources of contaminants, such as stormwater outfalls and agricultural crops, are relatively stationary and in some regards their "plumes" are reasonably well bounded, and animals migrating through these areas are often moving at a relatively predictable pace.

The contaminant levels of each individual-based submodels subsumed by the population representation is maintained and updated appropriately. No relative location information is maintained.

4 Results

The analysis in this section is based on two sets of runs. The purpose of each of these runs is to provide data for a comparison of run-times, the equivalence (or lack of equivalence) amongst the models, and to examine the robustness of the representations to changes in the configuration of the plume.

The first set is comprised of forty trials of each of the homogeneous representations (individual-based and population-based) and eighty trials of the mutating model. Each of these trials tracked either forty individuals or a population which represented forty individuals or a mixed system. In this set of runs the contaminant plume was circular and its centroid was on the migratory circle. This set of runs was conducted to establish the relative cost in cpu-time of the three approaches and to experimentally verify (at least to some degree) the equivalence which ought to exist between the purely individual-based submodel and the mutating submodel.

The second set of runs consisted of eighty trials of the mutating model, and one of the homogeneous population model. The principal difference between the first and second sets is that the plume in the second set has an elliptical footprint running along a tangent to the migratory path rather than the circular plume of the first trial. Our initial analysis will only consider the trials in the first set of runs; a comparison of the reponses to the elliptical plume will be treated separately.

Both population-based and individual-based representations use a heuristic which suppresses contact evaluation when they are sufficiently far from the contaminant source. No other significant optimisations of the code have been made, and as far as possible the same routines are used in both representations. The motivation for this approach is to provide a common baseline which we can use to measure the utility of the adaptive representation.

To ensure that run-time comparisons are meaningful all of the simulations in the first set of trials were run on the same computer. Each of the configurations simulated forty organisms for twelve years.

4.1 Contaminant load correspondence between representations

The set of trajectories arising from the various representations aren't directly comparable. Our individual-based representation produces a time series of contaminant levels for each individual, while the population submodel produces a "mean load" across the whole group of entities it is representing. The mutating submodel sits between the two, sometimes producing individual time series and sometimes mean time series for varying parts of the population. We denote representations by a subscript $r \in \{i, m, p\}$, so that $C_{rkj}(t)$ is the time series associated with individual j in trial k of representation r, $C_{rk}(t)$ is the mean over all the individuals in the indicated representation and trial, and $C_r(t)$ denotes the mean of $C_{rk}(t)$ across the k trials for the indicated representation. In order to compare the dynamics of the system we generate mean time series for each of the k trials in the individual-based and mutating sets, $C_{ik}(t)$ and $C_{mk}(t)$, paying special attention to generating the correct mean in the mutating submodel from time steps which have a mixture of individual trajectories and mean trajectories from population-based representations. Each of mean time series, $C_{rk}(t)$, correspond to the mean contaminant load of the population, $C_p(t)$, produced by the population submodel; averaging them – constructing $C_r(t) = \frac{1}{k}\Sigma_{j=1}^{k} C_{rj}(t)$, where r is one of 'i' or 'm' – is equivalent to running many stochastic trials and averaging to fit the population submodel.

Using $C_{ik}(t)$, $C_{mk}(t)$ and $C_p(t)$ we find the maximum value attained for each representation, $\hat{C}_r$. We are also interested in the mean value across time of each representation,¡htab—0¿ $\bar{C}_r = \frac{1}{T}\Sigma_{t\in T} C_r(t)$ These quantities are listed in Table 1.

Time series	$\hat{C}_r$	$\bar{C}_r$
C_i	0.1787	0.0390
C_m	0.1821	0.0392
C_p	0.1387	0.0350

Table 1: Maxima and Means

4.2 Contaminant load variability

We calculated measures of variability in the time series using the aggregated time series $C_{ik}(t)$ and $C_{mk}(t)$ and their respective means across the k trials, $C_i(t)$ and $C_m(t)$. We'll take T to be the total number of time steps taken, and we take

$$\hat{\sigma}_{ab} = \max_{t\in[1,T]} \left[\frac{1}{k} \sum_{j=1}^{k} (C_{ak}(t) - C_b(t))^2 \right]^{1/2}$$

and

$$\sigma_{ab} = \left[\frac{1}{T} \sum_{t=1}^{T} \left[\frac{1}{k} \sum_{j=1}^{k} (C_{ak}(t) - C_b(t))^2 \right] \right]^{1/2},$$

to be the maximum root mean square error and the average root mean square error. Clearly we can write σ_{rr} as σ_r without introducing ambiguity, and similarly for $\hat{\sigma}_r$. The values for these measure of of variability are presented in Table 4.2.

StdDev	$r = i$	$r = m$	$r = p$	
$\widehat{\sigma}_{ir}$	0.0083	0.0084	0.0534	
σ_{ir}	0.0024	0.0024	0.0096	¡htab—0¿
$\widehat{\sigma}_{mr}$	0.0090	0.0090	0.0538	
σ_{mr}	0.0024	0.0024	0.0096	

Table 2: Deviations amongst the model runs with respect to a given mean

4.3 Sensitivity to the shape of the plume

We will use the same notation as Section 4.2 for the data derived from the circular plumes, while we will add a prime symbol to the data derived from the elliptical plumes. Thus, the mean value time series for the mutating submodel with elliptical plumes would be denoted C'_m and the mean value of that time series is $\bar{C}'$.

The data for the circular plume is presented in Table 3 and the data for the elliptical plume is presented in Table 4.

Table 3: Circular plume results

Table 4: Eliptical plume results

4.4 Run-time

Each run collected data regarding the amount of time spend in different parts of the submodel. As a basis of comparison, the overall amount of time spent running on the cpu (linux/unix "cpu seconds") for the whole run is the most important figure, though the time spent in other parts provipdes illumination into just where the effort is concentrated. Predictably, most of the effort is in calculating contact and updating contaminant loads.

The optimisation of supressing the contact calculations when a population is outside the area of potential contact seemed to make very little difference to the run-time of population submodel (of the order of 3%), and seems unlikely to make a great deal of difference to the mutating submodel. In the case of the purely individual-based submodel, this sort of optimisation is likely to play a much bigger role in that a three second penalty would be multiplied by the number of entities simulated.

The population submodel ran for 98.7 cpu seconds. This submodel is deterministic and the amount of cpu time used is very stable, so only a single run is considered for comparison. The purely individual-based submodels took just over a mean time of 4205 cpu seconds with a standard deviation of barely more than 16 seconds and the mean of the mutating submodel's run time was 1157 cpu seconds with a standard deviation of slightly over 11 cpu seconds.

5 Discussion

The data establish the relative speeds of the three approaches and support the assertion that the mutating submodel possesses the same essential dynamics as the purely individual-based

representation. The essential equivalence of the individual-based and mutating representations is well supported and there is a significant benefit in run-time with the mutating submodel when compared to the individual-based submodel. The population submodel, while fastest by several orders of magnitude, showed responses to the two distinct contaminant plumes which were significantly different to those of the other two submodels.

The divergence between the population's density function and the observed distribution of individuals in the individual-based representation is potentially a source of model error. Though the actual distribution of individuals is slightly skewed by their motion along the migratory circle, it was very close to a two dimensional normal distribution and a standard 2d normal distribution was chosen to represent the density function of population with its variance taken from the observed individuals. Such a distribution corresponds to an ideal situation – in practice there may be a number of environmental influences which make a such a simple population distribution unrealistic: in the case of a coastal population, exceptions might include mud-flats, sand bars or islands.

5.1 Mutating models

The mutating model tests whether it may be more efficient in terms of time or representational accuracy to change the representation of components of the model in response to the local states of its submodels. The specific example we are dealing with is very simple, but it possesses many of the properties which the general problem of an *ad hoc* animal/plume interaction might be expected to exhibit, namely the tests for coincidence, generating a random walks, and evaluating the uptake and decay of contaminant loads. Over a long period of infrequent contact, we would expect quite a lot of time to be spent generating random walks and evaluating the contact and uptake in places where there is no likelihood of contact between animals and the plume. This strategy of changing representation is fundamentally about managing the often competing demands of representational accuracy, mathematical tractibility, and run-time.

The mutating submodel is an ensemble of cooperating models working together like a team in a relay race. Each of the representations has its own strengths and passes the baton to another at an appropriate time. So each of the submodels must have mappings between other representations which preserve as much information as possible without imposing excessive burden. As a first step in construction we must determine what information should be maintained in each of the representations to ensure that the migration from one formulation to another doesn't introduce unreasonable error and how that information should be maintained. In our example, the *significant* information which the individual-based representation possesses which the population representation does not is the variation in contaminants amongst the individuals within the group. We assume that the role of the individual locations of these individuals in this model is not important over a large portion of the state-space since the interval between conversions from individuals to aggregates is large enough to ensure a thorough randomisation of their relative position.

The state spaces of our base-line representations differ, unlike the example of Newtonian and Relativistic motion. The population models's motion is given by the function which defines the migratory circle, and individuals follow trajectories whose mean trajectory is this circle, so the natural state space of the individual based model is clearly incorporates position information lacking in the population's state space. The transition from individuals to population and population to individuals involves the loss and reconstruction of this fine-scale position data. The assumption that we can reconstruct a plausible position for each individual from the population's distribution function is tenable because there is no behavioural change associated with contaminant load – a contaminant that made an organism sluggish would obviously skew the

distributions of both the population as a whole and the distribution of intoxicated organisms within the population.

Individuals are mapped into a population by adding their contaminant level to the contaminant vector of an appropriate population after their movement and contaminant decay have been calculated. When a population submodel is mapped back to an individual-based representation, new individual-based submodels are created with appropriate contaminant levels and random locations which are chosen to be consistent with the distribution for the population around its centroid. This mapping preserves the contaminant load present in the system and the likelihood of an encounter between a contaminated individual and new contaminants. In this way the distinct contamination levels within the system are maintained with a minimum of cost.

Some overhead is incurred in managing the swap from one representation to another. The actual conversion is done by a routine which is integrated with the scheduler which manages the transition from one time-step to the next. Each of the submodels is able to indicate to the scheduler that a change in representation may be reasonable when it determines that it has left its domain of efficiency.

5.2 Heuristics: When do we swap

The contraints we impose on this example model are quite stringent so the heuristics we can use to decide when to change representation may assume a great deal more than might be the case in a general situation: we know where the contaminants are at all times, the movement of the populations and the individuals are controlled, and the behaviour of the simulated organisms does not change through the run. In the general case, the heuristics might be much more complex and require a great deal more data. Deciding when to change the representation is central to the optimisation of the system: changing at the wrong time might sacrifice accuracy, efficiency or both.

In our case, the principal issue is whether or not an agent (population-based or individual-based) may encounter contaminants: in our case it is simple, there is a single source and the mean movement toward or away from that source is quite predictable. As a population enters the region of contaminants we can immediately determine that it will consistently encounter contaminants for some non-trivial period, and a mapping is called for.

The test which determines if an individual is mapped into a population is similarly straightforward: an individual is mapped into a population if it is far enough outside the region of contaminants that there is no chance that it might encounter them in the next few time-steps, and it is close enough to a population. If there is no suitable population, a new one is created in order to accomodate it. This spatial constraint is stringent enough, when combined with the movement constraints, that the system hasn't needed to create more than one population in any of the annual cycles in the trials.

The transition rules in this submodel, from population to individual or from individual to population, can be based solely on the distance to the contaminant source, the time-step and the speed at which the simulated group of organisms moves around its annual path. Our heuristic could be generalised to systems where there were multiple, dynamically instanciated sources, when the sources of contamination could be determined at the beginning of a time-step. This particular submodel attempts to present a greatly simplified version of the general case – if the hypothesis were to fail in this case it would be reasonable to assume that it would fail in most cases.

5.3 Means, maxima and variation

In terms of the formulation of the mutating submodel, it is difficult to see how there could be a significant difference between its contaminant loads and those of the the purely individual-based submodel. The means, maxima and the various calculations of deviation between the individual-based and mutating submodels were very close and this supports the belief that they can be treated as essentially equivalent models of the contaminant load in the group of simulated organisms. Both of these representations, however, differed noticably from the population submodel; the population's mean value, $\bar{C}_p$, was about 10% lower than that of the either $\bar{C}_i$ or $\bar{C}_m$, and the statistics on the variation between representations (σ_{ab}, and $\hat{\sigma}_{ab}$) indicate that there is very little difference between the data produced by either the base-line individual-based model and the data produce by a mutating model. The population model, in contrast, produces means and maximum values are significantly lower (11% and 22%) than those of the individual-based model.

The discrepancy between the population-based uptake and individual-based uptake is enough to suggest the following possibilities:

- the formulation of the population distribution was not consistent with the spatial dynamics of the individuals,
- the tuning of the population submodel with respect to the individuals was inadequate,
- the model of interactions between the plume and the population or the plume and the individuals was not appropriate

This is the sort of situation discussed in section 5.4: the fundamental submodels exhibit different dynamics over the domain of the simulation, and this is one of the niches which a mutating approach might be most advantageous. It is difficult to accurately describe a wild population's distribution across the whole of its domain – changes in the behaviour of individuals (arising from predation, drought or other stressors) may engender quite different population distributions, so even in the case of a purely population-based representation an argument can be made for changing the underlying representation based on the local state.

5.4 Representational sensitivity to plume distribution

The issue of whether the configuration of the plume has an effect on the results of a simulation depending on the representation of the organisms is important. It is easy to imagine calibrating a population-based representation with data from fine-scale individual simulations in order to make the most of the economies a population-based representation affords, but this may not be a good idea if there is poor correspondence between the two submodels over their domain.

A set of trials with an elliptical plume were run to generate a set of data for comparison against the the data from the trials with a circular plume. Our analysis in these trials is based on the data from the first series of trials and from eighty trials of forty individuals with a mutating representation and a corresponding population-based trial with each trial covering twelve years. The elliptical plume covers the same area as the circular plume and the integral over the area of the plumes are the same, up to the numerical error in the quadrature.

The data from the mutating model in the set of trials with elliptical plumes had means, maxima and deviations which were consistent with the set of trials which used circular contaminant fields. This matches our expectations since the total areas covered and the total contaminant load over the area are the same for the two representations, and the speed of the agents through the plume is relatively slow. In contrast, the circular and elliptical trials of the population submodel had markedly different maxima and mean values through time.

The data suggest that the individual-based representation is reasonably robust with respect to the plume shape (at least in the context of this experiment). The population submodel did not behave as expected and it may be that the model of the distribution of individuals which formed the basis of the population submodel may have systematic problems with the population wide sampling of the plume – the increase in mean and maximum levels when the plume was elliptical and oriented tangentially along the track of the population suggests that the centre of the population is oversampled. The implications of this are that we must choose our population distribution carefully, and a simple normal distribution about a centroid may not behave well with respect to the plume dynamics.

Set against this, the numeric modelling of populations is well established. It is quite likely that the individual-based submodel would fail to meet expectations with regard to the dynamics of recruitment and mortality, had those aspects of the life history of a population been included in the model. In such a situation, the mutating submodel would have performed best overall – largely maintaining the fidelity of each of the representations across the state-space.

5.5 Run-time

The population submodel is clearly much faster than either of the other two representations, and for situations where we can reasonably make mean-field assumptions and our distribution model fits well with the observed data it is likely that the economies afforded by the speed of execution and the well established mathematical understanding of such models will vastly outweigh benefits of individual-based representations. In situations where mean-field assumptions occasionally fail to hold it is similarly clear that there are benefits to using mutating submodels. Purely individual-based submodels entail significant overhead, which in our context contributes nothing to the outcome. In a richer environment with diverse responses to the local state, an individual-based submodel's 'non-productive' overhead may be competitive with the overheads associated with the management and transition of mutating submodels.

It is difficult to decide what an optimal representation of a system is – not only is the trade-off between run-time and resolution a dilemma, but we then have to consider (at least in the case of stochastic models) whether to go for minimal error, or to maximise the number of trials in the time at hand. In the case of error versus number of trials, we can make an informed decision on where that balance lies if have a good estimate of how closely the model tracks "truth" but this is often not possible, particularly when the subject of the study is poorly observed.

6 Conclusion

The case for considering modelling schemes where the representation of a system changes in response to its local state seems to have merit. There are clear advantages in run-time relative to homogeneous individual-based simulations. In contrast there would seem to be little to gain in run-time with a shift from population-based representations to individual-based representations, but in situations where there are behavioural changes – particularly changes which affect the distribution of individuals – the scope for increasing the accuracy of the models is attractive.

The principle source of extra overhead in this modelling structure is in the scheduler which manages the multiple submodels. In the homogeneous population model, there is only one submodel running, and the extra cost of the more general scheduler over a simple iterative loop is negligible. Similarly the overhead in the homogeneous individual-based model is comparable to an additional iterative loop over the agents. Neither contribute a significant amount to the total time spent running a trial; both the solution of the uptake-depuration equation and the calculation of the contact dominated the run-time in both homogeneous models.

This simple model demonstrated that changing the representation of a system from one form to another can provide a mechanism for increasing the efficiency or accuracy with only a little extra effort. A more challenging avenue for study would be a similar model which included population recruitment and mortality – particularly if there were sub-lethal effects associated with the contaminant load which changed the behaviour or fecundity of the simulated organisms.

References