Process Algebra Models of Population Dynamics

Chris McCaig, Rachel Norman, and Carron Shankland

Department of Computing Science and Mathematics, University of Stirling, Stirling, FK9 4LA, UK. {cmc,ran,ces}@cs.stir.ac.uk

Abstract. It is well understood that populations cannot grow without bound and that it is competition between individuals for resources which restricts growth. Despite centuries of interest, the question of how best to model density dependent population growth still has no definitive answer. We address this question here through a number of individual based models of populations expressed using the process algebra WSCCS. The advantage of these models is that they can be explicitly based on observations of individual interactions. From our probabilistic models we derive equations expressing overall population dynamics, using a formal and rigorous rewriting based method. These equations are easily compared with the traditionally used deterministic Ordinary Differential Equation models and allow evaluation of those ODE models, challenging their assumptions about system dynamics. Further, the approach is applied to epidemiology, combining population growth with disease spread.

1 Introduction

The idea that populations cannot grow without bound has been of interest to modellers for centuries. Malthus [1], in 1798, proposed a simple exponential growth model based on compound interest but noted that this was unrealistic, since when a population becomes very large, access to resources will become restricted, restricting further growth in the population. Verhulst proposed the logistic growth model [2] to overcome this limitation and this is still widely used to describe density dependent growth. Many other models have been proposed to describe population dynamics [3–6] but it is not clear which model is most appropriate in any given situation; the logistic model is the default choice in the absence of other data. Models of population dynamics are not merely interesting in isolation. For example in our field, epidemiology, adding birth and death of individuals to a model of infectious disease spread can alter the dynamics of the epidemic. Therefore, getting a suitable model of population growth is an important step in producing realistic models of disease spread which can be analysed to provide predictive information about potential impact of epidemics, and to evaluate control strategies.

Process algebra has increasingly been used to model a wide range of biological systems [7–11]. The benefits of using process algebras to study such systems are twofold. First, process algebra allows formal, precise and unambiguous expression of a model. Second, process algebra has a formal mathematical semantics,

allowing rigorous investigation of the model via a range of techniques. For example, our work uses the discrete time process algebra *Weighted Synchronous Calculus of Communicating Systems* (WSCCS)[12]. The underlying semantics of WSCCS can be viewed as a Discrete Time Markov Chain (DTMC). Simulation can be used to explore the model. Steady state analysis can be carried out, and properties of the Markov Chain computed, e.g. probability of being in a particular state, or average number of occurrences of an action before a specific event occurs. Such investigation can be computationally expensive. Our previous work [13, 14] has been to facilitate further symbolic analyses of the model by developing a rewriting-based method to derive Mean Field Equations (MFEs) from a description of a system in WSCCS. The MFEs describe the average behaviour of the system at the population level and are analogous to traditional Ordinary Differential Equation (ODE) models of biological systems. The MFEs provide an approximation of the system dynamics of the DTMC corresponding to the WSCCS description. The derived MFEs are amenable to analysis using established algebraic techniques developed by mathematical biologists for ODEs.

The key advantage of our approach is that biological observations of individuals can be exploited in making the (individual based) WSCCS model, and the MFEs are derived automatically and efficiently. The alternative approach, used by mathematical biologists for many years, is to simply write down the MFE or ODE description assuming that behaviour at population level is well understood. While this formulation of the equations is backed up by experience, there is no rigorous relation between the actions of individuals and the outcome at a population level. Matching with disease data provides validation of ODEs, but many plausible terms can match the same data. In both approaches, facts about individual behaviour are abstracted to obtain population level equations capturing only information about the number in each class of individual. The difference is that our approach makes assumptions about behaviour explicit and that the method of abstracting is rigorous.

In this paper we consider the problem of accurately representing population growth using process algebra. Others have investigated individual based models of population dynamics and related their behaviour to population level equations. Sumpter [10] developed a simple WSCCS model of population growth and derived MFEs for the model. Brännström and Sumpter [15] presented individual based (not process algebra) models of competition which could be used to derive several existing population models but notably not Verhulst's logistic equation. The work presented here improves on previous work by applying a rigorous method across a range of different models of population growth.

Outline of the Paper. Section 2 gives a brief description of the syntax and semantics of WSCCS used in our models, and an outline of the method for deriving MFEs. In Sect. 3 WSCCS models of population dynamics are presented, which include density dependent growth in a variety of formulations (in either births or deaths, and introduced implicitly by enriching the WSCCS language or explicitly by including agents representing resources for which the population competes). The resultant changes in overall population dynamics are explored, comparing

the derived MFEs to traditional population level equations for population dynamics. In Sect. 4 we add disease spread to our model of population dynamics. Our results are summarised in Sect. 5.

2 Background

2.1 WSCCS Syntax and Semantics

In WSCCS the basic components are *actions* and the *processes* (or *agents*) that carry out those actions. The actions are chosen by the modeller to represent activities in the system. For example, *infect*, *send*, *receive*, *throw dice*, and so on. Actions form an abelian group with identity $\surd$ and the inverse of action a being $\overline{a}$. Actions occur instantaneously and have no duration. There is no notion of time in WSCCS, but there is ordering of events. WSCCS is a probabilistic process algebra, meaning that the decision to move from one state to another can be a probabilistic one. The formal syntax and semantics of WSCCS is presented in Tofts [12]. The main details are repeated here for the convenience of the reader.

The possible WSCCS expressions are given by the following BNF grammar:

$$A ::= X \mid a{:}A \mid \Sigma\{w_i.A_i | i \in I\} \mid A \times B \mid A\lceil L \mid \Theta(A) \mid A[S] \mid X \stackrel{\text{def}}{=} A \ .$$

Here $X \in Var$, a set of process variables; $a \in Act$, an action group; $w_i \in \mathscr{W}$, a set of weights; S a set of renaming functions, $S : Act \to Act$ such that $S(\surd) = \surd$ and $\overline{S(a)} = S(\overline{a})$; action subsets $A \subseteq Act$ with $\surd \in A$; and arbitrary indexing sets I. The informal interpretation of the operators is as follows:

- 0 a process which cannot proceed, representing deadlock ;
- X the process bound to the variable X ;
- $a{:}A$ a process which can perform the action a becoming the process A ;
- $\Sigma\{w_i.A_i | i \in I\}$ the weighted choice between processes A_i , the weight of A_i being w_i . Considering a large number of repeated experiments of this process, we expect to see A_i chosen with relative frequency $w_i/\Sigma_{i\in I} w_i$. Weights are generally positive natural numbers or reals, but may also incorporate the special weight ω which is greater than all natural numbers. This is used in *priority* and is written $m\omega^n$ where $m, n \geq 0$. The binary plus operator can be used in place of the indexed sum i.e. writing $\Sigma\{1_1.a{:}0, 2_2.b{:}0 | i \in \{1, 2\}\}$ as $1.a{:}0 + 2.b{:}0$;
- $A \times B$ the synchronous parallel composition of A and B . At each stage each process must perform an action with the composed process performing the composition (denoted $\#$) of the individual actions, e.g. $a : A \times b : B$ yields $a\#b{:}(A \times B)$. This is a powerful operator: models are constructed by describing simple individuals and composing a number of those in parallel. McCaig [13] introduces an extended notation $A\{n\}$ which is syntactic sugar for n instances of process A in parallel, where $n \in \mathbb{N}$;

$\dfrac{}{a{:}A \xrightarrow{a} A}$	$\dfrac{}{\sum\{w_i.A_i \mid i \in I\} \overset{w_i}{\longmapsto} A_i}$
$\dfrac{A \xrightarrow{a} A' \quad B \xrightarrow{b} B'}{A \times B \xrightarrow{a \# b} A' \times B'}$	$\dfrac{A \overset{w}{\longmapsto} A' \quad B \overset{v}{\longmapsto} B'}{A \times B \overset{wv}{\longmapsto} A' \times B'}$
$\dfrac{A \xrightarrow{a} A' \quad B \overset{w}{\longmapsto} B'}{A \times B \overset{w}{\longmapsto} A \times B'}$	$\dfrac{A \overset{w}{\longmapsto} A' \quad B \xrightarrow{a} B'}{A \times B \overset{w}{\longmapsto} A' \times B}$
$\dfrac{A \xrightarrow{a} A' \quad a \in L}{does_L(A)}$	$\dfrac{A \overset{w}{\longmapsto} A' \quad does_L(A')}{does_L(A)}$
$\dfrac{A \xrightarrow{a} A' \quad a \in L}{A \lceil L \xrightarrow{a} A' \lceil L}$	$\dfrac{A \overset{w}{\longmapsto} A' \quad does_L(A')}{A \lceil L \overset{w}{\longmapsto} A' \lceil L}$
$\dfrac{A \xrightarrow{a} A'}{A[S] \xrightarrow{S(a)} A'[S]}$	$\dfrac{A \overset{w}{\longmapsto} A'}{A[S] \overset{w}{\longmapsto} A'[S]}$
$\dfrac{A \xrightarrow{a} A' \quad X \overset{\text{def}}{=} A}{X \xrightarrow{a} A'}$	$\dfrac{A \overset{w}{\longmapsto} A' \quad X \overset{\text{def}}{=} A}{X \overset{w}{\longmapsto} A'}$
$\dfrac{A \xrightarrow{a} A'}{\Theta(A) \xrightarrow{a} \Theta(A')}$	$\dfrac{A \overset{n\omega^i}{\longmapsto} A' \nexists (j > i).A \overset{m\omega^j}{\longmapsto}}{\Theta(A) \overset{n}{\longmapsto} \Theta(A')}$

Table 1. Operational rules for WSCCS

- $A \lceil L$ a process which can only perform actions in the group L . This operator is used to enforce communication on actions $b \notin L$. Two processes in parallel may communicate when one carries out an action and the other carries out the matching co-action, e.g. *infect* and $\overline{infect}$. Communication can be used to model passing of information from one process to another, or to coordinate activity. Such communication is strictly two-way; that is, only two processes may interact on this action ;
- $\Theta(A)$ represents taking the prioritised parts of the process A only ;
- $A[S]$ represents A relabelled by the function S (we do not use relabelling in this paper, but it is included for completeness) ;
- $X \overset{\text{def}}{=} A$ represents binding the process variable X to the expression A .

The semantics of WSCCS is transition based, defining the actions that a process can perform and the weight with which a state can be reached. The operational rules of WSCCS, presented in Table 1, formalise the descriptions above. In particular note the two different arrows which feature in the table: $\xrightarrow{a}$ represents a transition associated with the action a ; and $\overset{w}{\longmapsto}$ represents a transition associated with a weight w . The auxiliary predicate $does_L(A)$, which denotes the ability of A to perform L after zero or more probabilistic actions, is well defined since only finitely branching choice expressions are allowed.

2.2 Deriving Mean Field Equations from WSCCS Models

In McCaig's thesis [13] and the related report [14] a method is described to automatically derive Mean Field Equations from WSCCS models. We give an

$$N \stackrel{\text{def}}{=} p_d.\surd : 0 + p_b.\surd : (N \times N) + (1 - p_d - p_b).\surd : N$$
$$\mathit{Population} \stackrel{\text{def}}{=} N\{n\}\lceil\{\surd\}$$

Fig. 1. Naive population model

overview of the approach here to aid understanding of the following sections. Sample derivations are given at the end of this section and in Sect. 3.2.

Consider the simple model of population growth in Fig. 1. The N agents die with probability p_d, becoming the null agent 0, give birth with probability p_b, becoming the agent consisting of two N agents in parallel, or do neither with probability $(1 - p_d - p_b)$, remaining as a single agent N. The model can be simulated, producing a single trace through the dynamics of the system. A second simulation may of course produce quite different behaviour since this is a stochastic process; therefore, of more interest is the *average* behaviour of the system as time progresses. This can be obtained by averaging the time series results of repeated simulations of the system. Clearly this becomes time-consuming, as the number of processes n and number of repetitions increases. An alternative is to generate the whole transition system for the model and to average the states produced, but as n increases the state space grows exponentially.

McCaig's method avoids generating the state space of the whole system, instead applying transformations to the WSCCS expression of the model, yielding an approximation (average) of the transition system in the form of first-order mean field equations. The approximation is shown to be "good" (i.e. lies within the standard deviation when compared with repeated simulations) in McCaig's thesis. Further, when the system becomes infinitely large, the mean of the DTMC corresponding to the transition system is proved to be equivalent to the derived MFEs. Larger populations eliminate the stochastic effects associated with low copy numbers.

The advantages of our approach are: the computational expense of generating the state space and/or simulation is avoided (the method is $O(a^2c)$ where a is the number of agents and c is the number of actions in the WSCCS description); it is a symbolic approach (avoiding questions regarding the exploration of the parameter space); and the MFEs, being a different view of the system and amenable to further analysis, offer new insight to the system.

The method applies to models of the form $A1\{n_1\}|...|Am\{n_m\}$ where the Ai communicate with each other (usually on a subset of actions). Models are limited in that steps involving probabilistic choice between actions must be separate from steps involving communication (which must have branches weighted 1).

Independently, the PEPA group [16, 17] and Cardelli [18] have proposed methods for deriving ODEs from process algebra. Their work differs in that their process algebras are continuous, based on rates rather than probabilities. Two of the methods are based on a mass action assumption, and not tied to the standard process algebra semantics. In contrast, our goal has been to pre-

serve this association, so that understanding individuals and their interactions translates automatically to population behaviour via process algebra semantics.

Transition Table: Relating Actions to Overall System Evolution. The transition system may be viewed as the evolution of the state vector $A1\{n_1\}|...|Am\{n_m\}$ For a particular Ai an action has three possible effects:

exit activity Following the action, the process evolves to some other agent Aj therefore the number of Ai agents is decreased.

entry activity In symmetry with an exit activity for Ai above must be an entry activity for Aj. The number of Aj agents increases.

none The process becomes Ai and there is no change in number of Ai agents.

WSCCS is a synchronous calculus, therefore in each time step, for every agent in the system, one of the above activities will occur. Our method is based around construction and interpretation of a transition table TT noting these exit and entry activities (Fig. 2).

```
for each agent Ai {
  for each (w_j.a_j : Ak) ∈ transitions(Ai) {
    for each Am ∈ components(Ak)
      TT[(Ai, a_j),Am] = TT[(Ai, a_j),Am] + calculateTerm(Ai, w_j, a_j)
  } }
```

Fig. 2. Constructing the transition table from a WSCCS model

The rows of TT denote the agents Ai at time t and their enabled actions aj. The columns of the transition table denote the agents Ak at the next time $t+1$. The term in cell $(Ai\ aj, Ak)$ is the proportion of Ai_t agents performing aj to become Ak_{t+1}. The derivation of this term is fully determined (see description below) by the context of the action carried out (e.g. part of a probabilistic choice, or part of a communication) and the composition of the population (i.e. how many of each different agent there are). Where Ai evolves to the same agent Ak irrespective of which action it performs a single row is used for that agent which is labelled $Ai\ *$. An example is the $F1$ agent in Fig. 4. The mean field equation for Ak_{t+1} is obtained by summing the terms in the column Ak.

Some auxiliary definitions are required. Processes can be classified by syntactic features as: *communicating* (having an action enabled that is involved in a communication), *probabilistic* (having only actions enabled that are not involved in communication), and *priority* (communicating and using ω weights). Given a *serial* process $A = w_1.a_1 : A1 + w_2.a_2 : A2 + ... + w_n.a_n : An$ define $transitions(A) = \{w_1.a_1{:}A1, w_2.a_2{:}A2, ..., w_n.a_n{:}An\}$. Given a *parallel* process $A = A1 \times A2 \times ... \times An$ define $components(A) = \{A1, A2, ...An\}$. For a process communicating on action a, we define two groups of agents involved in the collaboration: *collaborators* are those processes with the matching action $\overline{a}$, and *competitiors* are those processes with the same action a.

```
function calculateTerm (A, w, a): String {
  case A in {
    probabilistic(A): return w * A_t;
    communicating(A) and priority(A):
      term = (A_t * collaborators(A))/(A_t + competitors(A));
      if a equals √ return (A - term) else return term;
    communicating(A) and not priority(A):
      term = (A_t * collaborators(A))/(A_t + collaborators(A) + competitors(A));
      if a equals √ return (A - term) else return term;
}}
```

Fig. 3. Pseudo code to calculate proportion of agents at time $t + 1$

The pseudo code to compute the terms in the table is given in Fig. 3. For probabilistic choice, the semantics of WSCCS (Table 1) specifies that over a number of experiments the different branches will be taken in numbers consistent with their weights. For convenience, the weights in such choices sum to 1 in the models in this paper hence the term is simply wA_t. For communication, McCaig enumerates the possible outcomes based on a classification of modes of communication (prioritised or not, single action a or multiple actions e.g. $a\#a\#a$). This results in complex formulae based on the weighted multinomial choice of those outcomes giving the average number of communications. For single actions, as used in this paper, these formulae can be simplified. These are the formulae used in the *calculateTerm* function of Fig. 3. The full version of the approach [13, 14] assumes weights do not have to sum to 1, and also gives the formulae for multiple action communications.

Derivation of MFE for a Simple Population Growth Model. Consider again the simplistic model of population growth given in Fig. 1. The actions in Fig. 1 are simply $\surd$. That is, activities are of no interest, only the evolution of numbers of agents is significant. As in all of our models, the system as a whole is described by the system equation *Population*, comprising multiple copies of each kind of agent in parallel.

The transition table for this system is as follows:

	0	N_{t+1}
$(N_t, \surd)$	$p_d N_t$	$(1 - p_b - p_d)N_t + 2p_b N_t$

Each column leads to a MFE for that agent, but 0 is ignored here since this is not of interest to us. The method outlined above generates the following MFE

$$N_{t+1} = (1 + p_b - p_d)N_t \tag{1}$$

where N_{t+1} represents the number of N agents at time $t + 1$ expressed in terms of N_t, the number of N agents at time t. Since this model has no communication between agents, and a single step with probabilistic choice, the derived MFE can be easily verified manually.

3 Density Dependent Growth

Equation (1) describes a simple recurrence relation. With $p_b > p_d$ the population will become infinitely large; $p_b < p_d$ will lead to the population dying out, while $p_b = p_d$ will lead to an equilibrium state for any initial population size, $N_0 = n$. The probabilities p_b and p_d are fixed, therefore the average behaviour of this model is similar to that of the simple exponential growth model described by Malthus [1]. Biologically, it is more realistic to consider a model in which the probability of birth and death vary depending on the size of the population at each instant in time (density dependence). For example, as the population grows food and shelter become scarce, therefore individuals become weaker and are more likely to die. Alternatively this weakness may manifest itself as a reduced fecundity and a reduction in the birth rates. This section presents several ways of modelling these notions in WSCCS, obtaining more realistic models of population growth.

3.1 Functional Probabilities

What is required is the ability to modify p_b and/or p_d on the fly as the population changes. The first method described here is to add assumptions about how probability of birth and death depend on population size using *functional probabilities* [13]. Functional probabilities add considerable convenience and elegance of expression to complex models, while adding no new semantic concepts to WSCCS. Functional probabilities are implemented by encoding population size as part of the agent name, a technique [19] commonly used in process algebra. The size of the resultant model is much increased, and the translation itself is unremarkable: the interested reader is referred to [13] for the full details.

Instead of fixed probabilities, a functional definition is given. For example, probability p_x can be made a function f of the number of A agents (denoted $[A]$) by

$$p_x \stackrel{\text{def}}{=} min(max(0, f([A])), p_L) \ .$$

The function may take any format required, since it appears directly in the MFEs and is often not computed numerically. The probability p_L is the upper limit for p_x, chosen to ensure that all probabilities in the model are always in the range $0 \leq p \leq 1$. The *min* and *max* expressions may be required to ensure that p_x is in the allowed range, but these terms make mathematical analysis of the MFEs more complex. Often, in our further analysis we assert $p_x = f([A])$ based on very low likelihood of reaching a state where *min* and *max* are not satisfied by f (therefore those states make little contribution to the *average* behaviour captured by the MFEs).

Density Dependent Birth. Density dependent birth can be added to the model in Fig. 1 by making the probability of birth p_b inversely proportional to $[N]$.

$$p_b \stackrel{\text{def}}{=} min(max(0, p_{b_0} - k * [N]), p_L) \ ,$$

where p_{b_0} is the probability of birth in the absence of crowding and k is a measure of the strength of the effect of crowding, $0 < k \ll 1$.

Using the method of Sect. 2.2, the MFE derived is

$$\begin{aligned} N_{t+1} &= N_t + (p_{b_0} - kN_t - p_d)N_t \\ &= N_t + (p_{b_0} - p_d)N_t\left(1 - \frac{kN_t}{p_{b_0} - p_d}\right) \ . \end{aligned} \tag{2}$$

This is our first realistic model of population growth, derived from an expression of individual behaviour. Compare this to the discrete time version of Verhulst's logistic equation

$$N_{t+1} = N_t + rN_t\left(1 - \frac{N_t}{K}\right) \ . \tag{3}$$

where r represents reproductive rate and K the carrying capacity of the population. Simple substitution of $r = (p_{b_0} - p_d)$ and $K = (p_{b_0} - p_d)/k$ in (3) yields (2). The logistic equation is the most commonly used equation for describing population dynamics and is frequently included as a self limiting growth term in models of disease spread. This gives confidence in our approach, and endorses Verhulst's equation.

Density Dependent Death. Density dependent death can similarly be added to Fig. 1 by choosing probability of death p_d directly proportional to $[N]$ with

$$p_d \stackrel{\text{def}}{=} min(max(0, p_{d_0} + k * [N]), p_L) \ ,$$

where p_{d_0} is the probability of death in the absence of crowding. The MFE, derived once again using our method,

$$\begin{aligned} N_{t+1} &= N_t + (p_b - (p_{d_0} + kN_t))N_t \\ &= N_t + (p_b - p_{d_0})N_t\left(1 - \frac{kN_t}{p_b - p_{d_0}}\right) \ , \end{aligned}$$

is equivalent to the logistic equation with $r = (p_b - p_{d_0})$ and $K = (p_b - p_{d_0})/k$.

Summary. The results above are pleasing: we have shown that it is possible to derive the logistic equation from an individual based model of population growth. This contradicts the findings of Brännström and Sumpter [15] who did not find the logistic equation for any of their models. Our results should not be surprising: in the functional probabilities we are making the probabilities linearly proportional to the population size, effectively encoding the same assumptions which lead to the logistic equation in the traditional population level models. It would have been more surprising if we had not derived the logistic equation.

$$
\begin{aligned}
N1 &\stackrel{\text{def}}{=} 1.eat : (N2 \times N2) + 1.\surd : N2 \\
F1 &\stackrel{\text{def}}{=} 1.\overline{eat} : F2 + 1.\surd : F2 \\
N2 &\stackrel{\text{def}}{=} p_d.\surd : 0 + (1 - p_d).\surd : N1 \\
F2 &\stackrel{\text{def}}{=} 1.\surd : F1 \\
Population &\stackrel{\text{def}}{=} N1\{n\} \times F1\{f\} \lceil \{\surd\}
\end{aligned}
$$

Fig. 4. Density dependence on births with non-prioritised communication

3.2 Food as an Explicit Resource

The advantage of individual based modelling techniques is that population level assumptions can be avoided, to be replaced by population level behaviours arising from the explicit individual interactions. To the models seen so far we add agents representing "food", i.e. some finite resource required by individuals to survive, and for which there is competition. Any other similar resource, e.g. space, can be modelled in exactly the same way. Access to this resource can be used to determine the likelihood of either birth or death.

Acquiring a resource is modelled in WSCCS by communication between food agents and individuals, requiring the use of more complex language features than seen in the models so far. Two forms of communication are available: prioritised and non-prioritised. Using prioritised communication between the food agents and the population agents forces individuals to eat; however, in a population it is likely that some individuals, while foraging, may fail to find food which is present. Using non-prioritised communication models the possibility that individuals fail to eat even when food is present and is therefore more biologically plausible. As above, models exploring density dependence on births and density dependence on deaths are considered separately.

Density Dependence on Births. The model given in Fig. 4 has individuals in the population competing for the available food resource (the *eat* action), giving birth after eating, and dying probabilistically.

The agents $N1$ and $N2$ represent the members of the population at the different stages of the model. The $N1$ agents can eat and become the parallel agent $N2 \times N2$, representing birth. If they do not eat the $N1$ agents become a single $N2$ agent. In the second stage of the model the $N2$ agents make a probabilistic choice to die or survive. The total number of food agents is constant therefore the F agents $(F1, F2)$ should be thought of as units of food which the environment can produce in a time step rather than discrete portions of food which are consumed by the population.

For such models, the method generates MFE for all agents, i.e. $N1, N2, F1, F2$, where $N1$ is expressed in terms of $N2$ and vice versa. Similarly for $F1$ and $F2$. Generally we are interested only in a complete cycle of behaviour. That is, starting with agents $N1$, evolving to agents $N2$, then back to $N1$ (two stages here).

We take the $N1$ equation, substitute to remove occurrences of $N2$ and obtain an equation only in $N1$ (and $F1$). Finally, we rename $N1$ as simply N. The fact that the number of food agents remains constant means that the derived MFE for $F1$ can be simplified to f in the MFE for N.

Deriving the terms of the MFEs for this model is more complex: although the definition of $N1$ suggests the choice to *eat* or not is equally weighted, in fact this choice is also influenced by availability of $F1$ agents with which to synchronise. This is reflected in the *calculateTerm* function described in Sect. 2.2. For example, here it is possible that no individuals eat (with very low probability), or that all do (assuming $[N1] \leq [F1]$) (also with low probability), or all of the possibilities inbetween. As explained earlier, the *calculateTerm* function yields a formula based on the weighted multinomial choice of those possible outcomes. The method yields the following transition table. Note that the term for the communicating action (*eat*) reflects that $N1$ *collaborates* with $F1$ but has no *competitors* for the action.

	0	$N1_{t+1}$	$N2_{t+1}$	$F1_{t+1}$	$F2_{t+1}$
$(N1_t, eat)$			$2\frac{N1_t * F1_t}{N1_t + F1_t}$		
$(N1_t, \surd)$			$N1_t - \frac{N1_t * F1_t}{N1_t + F1_t}$		
$(F1_t, *)$					$F1_t$
$(N2_t, \surd)$	$p_d N2_t$	$(1-p_d)N2_t$			
$(F2_t, *)$				$F2_t$	

Summing the columns and simplifying as described above leads to the MFE

$$N_{t+1} = (1-p_d)N_t + \frac{(1-p_d)fN_t}{f+N_t} \quad . \tag{4}$$

Here the term $(1-p_d)N_t$ represents the mean proportion of the existing population which survives the probabilistic death stage. The term $fN_t/(f+N_t)$ represents the mean number of new births with the factor $(1-p_d)$ representing the proportion of new births which survive the probabilistic death stage. We find the steady state of this model by setting $N_{t+1} = N_t = N^*$:

$$N^* = (1-p_d)N^* + \frac{(1-p_d)fN^*}{f+N^*} \quad .$$

Solving for N^* we get

$$N^* = \frac{(1-2p_d)f}{p_d} \quad .$$

For biological realism the steady state should be positive, therefore $p_d < 0.5$. Note that this fact is not obvious from the WSCCS model, but becomes clear in the MFE. The values of these probabilities can be observed in the field, but an important factor is the length of timestep. If we need to reduce p_d to meet the above requirement we can reduce the timestep represented by our models and adjust all other parameters accordingly.

$$
\begin{aligned}
N1 &\stackrel{\text{def}}{=} 1.eat : N2 + 1.\surd : 0 \\
F1 &\stackrel{\text{def}}{=} 1.\overline{eat} : F2 + 1.\surd : F2 \\
N2 &\stackrel{\text{def}}{=} p_b.\surd : (N1 \times N1) + p_d.\surd : 0 + (1 - p_b - p_d).\surd : N1 \\
F2 &\stackrel{\text{def}}{=} 1.\surd : F1 \\
Population &\stackrel{\text{def}}{=} (N1\{n\} \times F1\{f\})\lceil\{\surd\}
\end{aligned}
$$

Fig. 5. Density dependence on deaths with non-prioritised communication

Sumpter [10] developed a mechanism for describing self limiting growth in a population which made use of food as an agent. He derived the following MFE using an heuristic

$$N_{t+1} = (1 - p_d)N_t + \min[(1 - p_d)N_t, f] \ ,$$

where p_d is the probability of death in any timestep and f is the number of food agents. The underlying assumptions of this model are undesirable biologically: individuals are guaranteed to find food if it is available because prioritised communication is used. Therefore, every member of the population will give birth at each step of time until the size of the population is larger than the number of food agents, after which the number of births will be equal to the number of food agents. This model has a stable steady state of $N^* = f/p_d$, when $p_d \leq 0.5$, which is larger than for our model.

Density Dependence on Deaths. In Fig. 5 the $N1$ agents can once again eat, becoming the agent $N2$, but here if they do not eat they die, becoming the null agent 0. The $N2$ agents then give birth probabilistically and, to be realistic, can also die probabilistically. That is, in each step of time a proportion of the population die, for instance, due to age and some die due to a lack of food. The MFE for this model is

$$N_{t+1} = (1 + p_b - p_d)\frac{fN_t}{f + N_t} \ , \tag{5}$$

where term $fN_t/(f + N_t)$ represents the proportion of the population which eat and therefore survive the competition for food, with the factor $(1 + p_b - p_d)$ representing the increase in the population due to births and the decrease due to probabilistic death. Equation (5) can be rearranged to give

$$N_{t+1} = \frac{aN_t}{1 + bN_t} \ , \tag{6}$$

where $a = (1 + p_b - p_d)$ and $b = 1/f$. Equation 6 is the Beverton-Holt model [3], originally proposed as a model of salmon populations displaying density dependent birth; however, we have derived this equation from an individual

$$
\begin{aligned}
N1 &\stackrel{\text{def}}{=} p_b.\surd : (N2 \times B2) + p_d.\surd : D2 + (1 - p_b - p_d).\surd : N2 \\
F1 &\stackrel{\text{def}}{=} 1.\surd : F2 \\
N2 &\stackrel{\text{def}}{=} 1.eat : N1 + 1.\surd : 0 \\
F2 &\stackrel{\text{def}}{=} 1.\overline{eat} : F1 + 1.\surd : F1 \\
B2 &\stackrel{\text{def}}{=} 1.\surd : N1 \\
D2 &\stackrel{\text{def}}{=} 1.eat : 0 + 1.\surd : 0 \\
Population &\stackrel{\text{def}}{=} (N1\{n\} \times F1\{f\})\lceil\{\surd\}
\end{aligned}
$$

Fig. 6. Density dependence on deaths, with choice followed by communication

based model featuring density dependent death. Our derivation endorses the plausibility of the Beverton-Holt model, which is commonly used in models of plant populations but not so widely used for animal populations.

Setting $N_{t+1} = N_t = N^*$ in (5) and solving for N^* yields the steady state

$$N^* = (p_b - p_d)f \ .$$

In this case to ensure the steady state is positive we require $p_b > p_d$.

Order Matters? Clearly, changing the WSCCS model can affect the MFEs derived, but even a relatively small, intuitively negligible, change can make a difference. In the models considered in Figs. 4 and 5 the focus is on the two-stage behaviour of the $N1$ agents. This means that the communicative (eating) step is followed by the probabilistic step with births and deaths. We may naively assume that considering the two-stage behaviour of the $N2$ agents, thus reversing the order of the communicative and probabilistic steps, would lead to the same overall long term behaviour of the model. However, the derived MFE for the behavior of the $N2$ agents in Fig. 5 is

$$N_{t+1} = (1 + p_b - p_d)\frac{fN_t}{f + (1 + p_b - p_d)N_t} \ ,$$

where the denominator features a factor of $(1 + p_b - p_d)$ not present in (5).

This difference arises because changing the order in which the steps occur also changes the underlying biological assumptions of the model. The newborn individuals are now available to compete for the available food (leading to the $+p_b$ term) and the individuals which probabilistically die are not (leading to the $-p_d$ term). Generating a WSCCS model in which probabilistic choice is followed by a communicative phase is more complex than simply swapping these steps. A suitable model, which will lead to the MFE (5), can be seen in Fig. 6.

In Fig. 6 the agents which make the probabilistic choice to die enter a dying state, $D2$, where they compete for food and are then removed from the system,

irrespective of whether they eat or not. The newly born individuals are in the state $B2$ which does not compete for food and becomes $N1$ at the next stage. This means that the overall mean two-stage behaviour of the $N1$ agents in Fig. 6 is the same as for the $N1$ agents in Fig. 5.

This simple example illustrates the importance of thinking carefully about the biological interpretation of actions in the WSCCS model, highlighted by the derivation of MFEs. This is particularly important when considering more complex models such as that in Sect. 4 which adds population dynamics to a model of infectious disease.

4 Population Dynamics and Disease

While population dynamics are interesting in their own right they are also crucial in developing realistic models of disease spread. The model in Fig. 7 adds infectious disease spread, based on the models of Norman and Shankland [8], to the density dependent death population dynamics of Fig. 5. In a typical disease model the population is divided into 3 groups: susceptibles (S) have never had the disease, infecteds (I) currently have the disease, and recovereds (R) have previously had the disease and are immune to future infection.

The first stage in the model is the eating stage in which $S0, I0$ and $R0$ all compete for food. Those that do not eat will die. The second stage is a contact stage in which infected (*Trans*) agents come into contact with the population and potentially pass the disease to susceptibles. The infected individuals are represented by parallel agents with the *Trans* agents passing on the disease and the $T1$ agents able to be contacted by a *Trans* agent. Communication is prioritised so that all *Trans* make contact. Prioritised contact is plausible biologically since contact with the whole population is possible (not just the susceptibles) and contact is not guaranteed to result in infection (see $SI2$). $S1$ that are contacted become $SI2$, while $T1$ and $R1$ agents are not affected by contact since infected and recovered individuals cannot become infected again. After the contact stage the *Trans* agents all become the null agent 0 so that the infected individuals are once again represented by a single agent. The final stage is the probabilistic stage in which all individuals can give birth to a susceptible individual, with probability p_b, or die, with probability p_d. In addition the $SI2$ agents become infected with probability p_a and $I2$ agents can recover with probability p_r.

The system of MFEs derived from this model is

$$\begin{aligned} S_{t+1} &= \frac{f}{f+N_t}\Big((1-p_d)S_t + p_b N_t - \frac{p_a S_t I_t}{N_t}\Big) \\ I_{t+1} &= \frac{f}{f+N_t}\Big((1-p_d-p_r)I_t + \frac{p_a S_t I_t}{N_t}\Big) \\ R_{t+1} &= \frac{f}{f+N_t}\Big((1-p_d)R_t + p_r I_t\Big) \ , \end{aligned} \tag{7}$$

where $N_t = S_t + I_t + R_t$, the total population size at time t. These are similar to the standard SIR equations with frequency dependent transmission of

$$
\begin{aligned}
S0 &\stackrel{\text{def}}{=} 1.eat : S1 + 1.\surd : 0 & S1 &\stackrel{\text{def}}{=} \omega.infect : SI2 + 1.\surd : S2 \\
R0 &\stackrel{\text{def}}{=} 1.eat : R1 + 1.\surd : 0 & R1 &\stackrel{\text{def}}{=} \omega.infect : R2 + 1.\surd : R2 \\
I0 &\stackrel{\text{def}}{=} 1.eat : (T1 \times Trans) + 1.\surd : 0 & T1 &\stackrel{\text{def}}{=} \omega.infect : I2 + 1.\surd : I2 \\
Food0 &\stackrel{\text{def}}{=} 1.\overline{eat} : Food1 + 1.\surd : Food1 & Trans &\stackrel{\text{def}}{=} \omega.\overline{infect} : 0 + 1.\surd : 0 \\
Food1 &\stackrel{\text{def}}{=} 1.\surd : Food2 \\
Food2 &\stackrel{\text{def}}{=} 1.\surd : Food0 \\
S2 &\stackrel{\text{def}}{=} p_b.\surd : (S0 \times S0) + (1 - p_b - p_d).\surd : S0 + p_d.\surd : 0 \\
SI2 &\stackrel{\text{def}}{=} p_b.\surd : (S0 \times S0) + p_a.\surd : I0 + (1 - p_a - p_b - p_d).\surd : S0 + p_d.\surd : 0 \\
I2 &\stackrel{\text{def}}{=} p_b.\surd : (I0 \times S0) + p_r.\surd : R0 + (1 - p_r - p_b - p_d).\surd : I0 + p_d.\surd : 0 \\
R2 &\stackrel{\text{def}}{=} p_b.\surd : (R0 \times S0) + (1 - p_b - p_d).\surd : R0 + p_d.\surd : 0 \\
Population &\stackrel{\text{def}}{=} \Theta((S0\{s\} \times I0\{i\} \times Food0\{f\})\lceil\{\surd\})
\end{aligned}
$$

Fig. 7. SIR model with density dependence on deaths

disease [20], a form arising naturally from WSCCS models [8]. Here, however, there is an extra factor of $f/(f + N_t)$ on each equation that is the proportion of the population successfully eating. This is unconventional since in traditional models the transmission term (in this case $(p_a S_t I_t)/N_t$) is not affected by the density dependent birth or death term. We emphasise that the population dynamics of (7) come directly from explicit representation of individuals competing for food rather than any population level assumptions imposed on the model. These equations are therefore candidates for modelling population dynamics in disease systems, despite the differences to traditional models.

In contrast, if we had taken the population dynamics from Sect. 3.1, with functional probability of birth, and added disease as above, we would merely add a logistic term to the equation for S with each group also dying probabilistically. This result would be closer to the traditional ODE models. The advantage of this approach is that the nonlinear density dependent term only appears in one equation (S), therefore the equations are simpler and easier to analyse mathematically than (7) which contains nonlinear terms in all equations. The disadvantage of basing a disease model on the functional probability models of population growth is that the latter are based on assumptions about population growth which may be incorrect.

5 Conclusion

In this paper we have presented population dynamics models in which the population will, over time, tend to some steady state and will not display unbounded growth. There are two distinct types of model: those in which the effects of restricted resources are implicitly included by allowing more complex language fea-

tures in the model (functional probabilities) and those in which those resources are explicitly represented by agents. The introduction of functional probabilities allow us to succinctly take full advantage of the expressive capabilities of WSCCS. These models led naturally to the logistic equation [2], the classical expression used to describe population dynamics. This is in contrast to the results of Brännström and Sumpter [15] who found several other existing expressions could be derived from their individual based models but not the logistic equation. The logistic equation arises from our models because the assumptions used to introduce density dependence – functional probabilities which are linearly proportional to the population size – match the assumptions on which the logistic equation is based. If we use functional probabilities which are non-linearly proportional to the population size we would of course obtain different MFEs. It can be easily argued that adding functional rates is self-defeating for our objectives; if we allow inclusion of strong implicit assumptions, such as the nature of population growth, then we may as well simply write down the MFEs directly.

In order to reduce the number of population level assumptions in our models we have also developed models which feature agents to represent food, with the dynamics in the population arising from the competition between individuals for food. With density dependent death this model leads to the Beverton-Holt model [3] which was proposed for the population dynamics of fish stocks. The fact that this equation has naturally arisen here from the competition between individuals means we can consider the Beverton-Holt model a serious candidate to be used when modelling population dynamics. Further investigation including matching with data is required.

Lastly, our goal in population modelling is to incorporate models of disease to gain a more realistic individual based disease model. By adding a model of disease spread to population dynamics we have derived a system of equations (7) which differs from those which have previously been described in the literature. Because the population dynamics in our model naturally arise from the interactions between individuals and the environment, rather than any assumptions we have imposed on the population dynamics, we have well-founded reason to propose this model for a disease system featuring density dependence in deaths. As above, future work will include validating our models with disease data.

Acknowledgements This work was supported by EPSRC through a Doctoral Training Grant (CM, from 2004-2007), and through *System Dynamics from Individual Interactions: A process algebra approach to epidemiology* (EP/E006280/1, all authors, 2007-2010).

References

1. Malthus, T.: An essay on the principle of population (1798)
2. Verhulst, P.: Notice sur la loi que la population suit dans son accroissement. Corr. Math. et Phys. **10** (1838) 113–121
3. Beverton, R., Holt, S.: On the dynamics of exploited fish populations. Volume 19 of Fisheries Investigations, Series 2. H.M.S.O. (1957)

4. Gompertz, B.: On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. Philosophical transactions of the Royal Society of London Series B **115** (1825) 513–585
5. Hassell, M.: Density-dependence in single-species populations. Journal of Animal Ecology **45** (1975) 283–296
6. Ricker, W.: Stock and recruitement. Journal of the Fisheries Research Board of Canada **11** (1954) 559–623
7. Calder, M., Gilmore, S., Hillston, J.: Modelling the influence of RKIP on the ERK signalling pathway using the stochastic process algebra PEPA. In: Proceedings of BioCONCUR 2004. Electronic Notes in Theoretical Computer Science. Elsevier (2004)
8. Norman, R., Shankland, C.: Developing the use of process algebra in the derivation and analysis of mathematical models of infectious disease. In: Computer Aided Systems Theory - EUROCAST 2003. Volume 2809 of Lecture Notes in Computer Science. Springer-Verlag (2003) 404–414
9. Regev, A., Panina, E., Silverman, W., Cardelli, L., Shapiro, E.: Bioambients: an abstraction for biological compartments. Theoretical Computer Science **325** (2004) 141–167
10. Sumpter, D.: From Bee to Society: an agent based investigation of honeybee colonies. PhD thesis, UMIST (2000)
11. Tofts, C.: Using process algebra to describe social insect behaviour. Transactions of the Society for Computer Simulation **9** (1993) 227–283
12. Tofts, C.: Processes with probabilities, priority and time. Formal Aspects of Computing **6** (1994) 536–564
13. McCaig, C.: From individuals to populations: changing scale in process algebra models of biological systems. PhD thesis, University of Stirling (2008) Available from www.cs.stir.ac.uk/˜cmc/thesis.ps.
14. McCaig, C., Norman, R., Shankland, C.: Deriving mean field equations from process algebra models. Technical Report 175, University of Stirling (2008) Available from www.cs.stir.ac.uk/research/publications/techreps.
15. Brännström, A., Sumpter, D.: The role of competition and clustering in population dynamics. Proceedings of the Royal Society of London Series B **272** (2005) 2065–2072
16. Calder, M., Gilmore, S., Hillston, J.: Automatically deriving ODEs from process algebra models of signalling pathways. In: Proceedings of CMSB 2005 (Computational Methods in Systems Biology). (2005) 204–215
17. Hillston, J.: Fluid Flow Approximation of PEPA models. In: QEST'05, Proceedings of the 2nd International Conference on Quantitative Evaluation of Systems. IEEE Computer Society Press, Torino (2005) 33–42
18. Cardelli, L.: On process rate semantics. Theoretical Computer Science **391** (2008) 190–215
19. Milner, R.: Communication and Concurrency. Prentice Hall (1989)
20. Begon, M., Bennet, M., Bowers, R., French, N., Hazel, S., Turner, J.: A clarification of transmission terms in host-microparasite models: numbers, densities and areas. Epidemiology and infection **129** (2002) 147–153