

Modelling the evolution of periodicity in the periodical cicadas

Jaakko Toivonen and Lutz Fromhage

Department of Biological and Environmental Science, University of Jyväskylä, Jyväskylä, Finland

ABSTRACT

Background: Periodical cicadas (*Magicalcadas* spp.) have a life cycle that ends with the entire underground nymph population exhibiting a synchronized mass emergence to mate above ground. Previous studies have hypothesized that the periodical cicadas evolved from non-periodical cicadas by switching from a life-cycle length determined by body size to one determined by age.

Questions: When can a mutation coding for fixed life-cycle length invade a resident population in which life-cycle length is variable? What determines the length of the fixed cycle?

Methods: Numerical analysis of a mathematical model and simulations of an individual-based model.

Results: If there is a sufficiently strong predation intensity affecting the pool of individuals emerging to reproduce, a non-periodical population may become proto-periodical such that reproductive success varies yearly. Then, an emergence strategy with a fixed life-cycle length targeting years of high emergence density can invade.

Keywords: Allee effect, individual-based simulation, numerical analysis, periodical cicadas, semelparity, structured population model.

INTRODUCTION

The periodical cicadas (*Magicalcadas* spp.) are famous for their long, prime-numbered life cycles of 13 and 17 years. Periodical cicadas live almost all of their life underground as nymphs before they exhibit a tremendous synchronized mass emergence above ground to mate, reproduce, and then die. There exists a large body of work on the biology, ecology, and evolution of the periodical cicadas (Alexander and Moore, 1962; Dybas and Davis, 1962; Dybas and Lloyd, 1962, 1974; Lloyd and Dybas, 1966a, 1966b; White and Lloyd, 1975; White *et al.*, 1979; Karban, 1982, 1984; Martin and Simon, 1990; Heliövaara *et al.*, 1994; Marshall and Cooley, 2000; Sota *et al.*, 2013), with an excellent review given by Williams and Simon (1995).

It has been suggested that the periodical cicadas evolved from non-periodical cicadas (Martin and Simon, 1990) by switching from a size-based to an age-based emergence strategy

Correspondence: J. Toivonen, Department of Biological and Environmental Science, University of Jyväskylä, 40014 Jyväskylä, Finland. email: jaakko.toivonen@alumni.helsinki.fi
Consult the copyright statement on the inside front cover for non-commercial copying policies.

(Yoshimura, 1997; Ito *et al.*, 2015). With size-based emergence, individuals emerge from the ground when they have reached a threshold body size. Since individual growth rates vary, the population then is necessarily non-periodical, as individuals reach the threshold body size at different times. In the case of age-based emergence, individuals wait until they reach a certain age and then emerge irrespective of size. In this paper, we model the cicada life cycle and examine under what circumstances a switch from size-based to age-based emergence may occur.

Generally, a population is called periodical ‘if the life cycle has a fixed length of k years ($k > 1$) and if the adults do not appear every year but only every k th year’ (Bulmer, 1977). A number of studies (Hoppensteadt and Keller, 1976; Bulmer, 1977; Behncke, 2000; Davydova *et al.*, 2003, 2005; Diekmann *et al.*, 2005; Kon and Iwasa, 2007) have shown that if a population consists of k different age classes, ecological interactions can result in the extinction of all but one age class, thus making the population periodical. However, typically it is assumed that life-cycle length is already fixed so that each individual emerges exactly after k years. In this paper, we demonstrate how a population where individuals have random life-cycle lengths (size-based emergence strategy) can evolve into a periodical population where all individuals have the same fixed cycle (age-based emergence) and all emerge synchronously. First, we derive and numerically analyse a mathematical population model that highlights the basic ecological mechanisms promoting invasion of an age-based strategy in a resident population with the size-based strategy. This model makes a number of simplifying assumptions – for example, it assumes clonal reproduction, a simplified age structure, a particular limit case of size-dependent fecundity, and density-independent predation. Despite the simplifications, the model allows us to study when a size-based population can become proto-periodical and then be invaded by mutant age-based emergence strategies with different cycle lengths.

In the second part of the paper, we build a more realistic individual-based model (IBM) that relaxes all of the simplifications of the mathematical model. We run simulations of the IBM to verify and expand the results obtained from the analysis of the simplified mathematical model. Although our IBM is similar to the model of Ito *et al.* (2015), there are two key differences: we allow both short (1–9 years) and long (10–20 years) cycle lengths for age-based emergence, and we let the size threshold for size-based emergence evolve (instead of choosing an arbitrary fixed value). Our approach has the advantage of removing a confounding mechanism that could otherwise affect the results. Namely, if a size-based emergence strategy is based on an arbitrarily chosen size threshold, then this threshold is unlikely to be optimal with regard to the trade-off between nymph survival and fecundity, where the former decreases and the latter increases the longer an individual remains in the ground. Then, an arbitrarily chosen size threshold is generally evolutionarily unstable and potentially allows age-based emergence strategies to invade for reasons that have nothing to do with synchrony of emergence.

MATHEMATICAL MODEL

Here we assume that individuals live underground for a year while competing for resources, resulting in mortality due to interference competition. At the end of the year, individuals may emerge from the ground depending on their emergence strategy to produce new offspring. The emergence strategy is given by $\eta = (\eta_1, \eta_2)$, where η_1 is the proportion of the first year age class to emerge and η_2 is the proportion of the second year class to emerge. Individuals who do not emerge remain in the ground for another year and may emerge at

the end of the following year. Here we assume that every individual emerges at least by the end of their third year.

Every year, the emerging adult population suffers mortality due to predation. Adults that avoid predation produce offspring that are deposited in the ground. After producing new offspring, the adults die. If all individuals who emerge are predated, no new offspring are produced.

The above considerations lead us to the following discrete-time model:

$$n_1(t+1) = \beta_\eta \cdot \max\{0, (\eta_1 n_1(t) + \eta_2 n_2(t) + n_3(t)) \Gamma(\bar{n}(t)) - \gamma\} \quad (1)$$

$$n_2(t+1) = (1 - \eta_1) n_1(t) \Gamma(\bar{n}(t)) \quad (2)$$

$$n_3(t+1) = (1 - \eta_2) n_2(t) \Gamma(\bar{n}(t)), \quad (3)$$

where n_1 , n_2 , and n_3 are the population densities of individuals in their first, second, and third year in the ground, respectively, and the factor β_η is the fecundity of an individual with strategy η . The term

$$\Gamma(\bar{n}(t)) := \frac{1}{1 + \alpha(n_1(t) + n_2(t) + n_3(t))} \quad (4)$$

is the probability of surviving competition in the ground (Beverton and Holt, 1957) and $\bar{n} = (n_1, n_2, n_3)$ is a vector of population densities of all the age classes. Survival probability Γ is the same for each individual irrespective of age. The term γ is the reduction in adult population density due to predation. This is an extremely simple model of predation, which assumes that predation is independent of cicada density and the predator population density is constant from year to year. However, it is not a completely unreasonable assumption considering the conspicuousness of the *Magicicada* emergence and their complete lack of predator avoidance [Lloyd and Dybas (1966b) even describe the *Magicicada* as ‘predator fool-hardy’]. The discrete-time models of predation and survival above can be derived from continuous time processes and we include a simple derivation of both in the Appendix [for further discussion on derivations of discrete-time dynamics from continuous time processes, see, for example, Geritz and Kisdi (2004) and Eskola and Geritz (2007)].

Finally, we note that it is possible to simplify the model (1–3) slightly by scaling the factor α out. Choosing $\tilde{n}_1 := \alpha n_1$, $\tilde{n}_2 := \alpha n_2$, and $\tilde{\gamma} := \alpha \gamma$ (and then dropping the tilde-signs) gives the model (1–3) with $\alpha = 1$. This is the model (for a single phenotype) that we use in the following.

Invasion fitness

When the population consists of k phenotypes, the probability to survive competition in the ground in a given year t is

$$\Gamma_k(\bar{n}_1, \dots, \bar{n}_k) = \frac{1}{1 + \sum_{j=1}^k n_{j,1} + n_{j,2} + n_{j,3}}, \quad (5)$$

where $\bar{n}_i$ is the vector containing the densities of all age classes $n_{i,1}$, $n_{i,2}$, and $n_{i,3}$ of phenotype i . The density of phenotype- i individuals that emerge in a given year is

$$E_i(\bar{n}_1, \dots, \bar{n}_k) = \Gamma_k(\bar{n}_1, \dots, \bar{n}_k) (\eta_{i,1} n_{i,1} + \eta_{i,2} n_{i,2} + n_{i,3}), \quad (6)$$

where $\eta_{i,1}$ and $\eta_{i,2}$ are the probabilities of emergence from the first and second age classes, respectively, for phenotype i . The density of all individuals that emerge is

$$E_{all}(\bar{n}_1, \dots, \bar{n}_k) = \Gamma_k(\bar{n}_1, \dots, \bar{n}_k) \left(\sum_{j=1}^k \eta_{j,1} n_{j,1} + \eta_{j,2} n_{j,2} + n_{j,3} \right). \quad (7)$$

Then, the model (1–3) generalizes for k phenotypes to

$$n_{i,1}(t+1) = \beta_i \cdot \max \left\{ 0, E_i(\bar{n}_1(t), \dots, \bar{n}_k(t)) - \frac{E_i(\bar{n}_1(t), \dots, \bar{n}_k(t))}{E_{all}(\bar{n}_1(t), \dots, \bar{n}_k(t))} \gamma \right\} \quad (8)$$

$$n_{i,2}(t+1) = \Gamma_k(\bar{n}_1(t), \dots, \bar{n}_k(t)) ((1 - \eta_{i,1})n_{i,1}(t)) \quad (9)$$

$$n_{i,3}(t+1) = \Gamma_k(\bar{n}_1(t), \dots, \bar{n}_k(t)) ((1 - \eta_{i,2})n_{i,2}(t)), \quad (10)$$

where β_i is the fecundity of phenotype i . The predation experienced by the emerging individuals of phenotype i is assumed to be proportional to the fraction of individuals of phenotype i in the whole of the emerging population, which is reflected by the term $E_i(\bar{n}_1, \dots, \bar{n}_k)/E_{all}(\bar{n}_1, \dots, \bar{n}_k)$.

We rewrite the model for phenotype i shorthand as

$$\bar{n}_i(t+1) = f_i(\bar{n}_1(t), \dots, \bar{n}_k(t)), \quad (11)$$

where $f_i = (f_{i,1}, f_{i,2}, f_{i,3})^\top$ and $f_{i,1}$, $f_{i,2}$ and $f_{i,3}$ are given by the right-hand side of (8)–(10), respectively. Note that $f_{i,1}$ is continuously differentiable if

$$\left| E_i(\bar{n}_1(t), \dots, \bar{n}_k(t)) - \frac{E_i(\bar{n}_1(t), \dots, \bar{n}_k(t))}{E_{all}(\bar{n}_1(t), \dots, \bar{n}_k(t))} \gamma \right| > \delta \quad (12)$$

for some $\delta > 0$. With this restriction, the Jacobian of f_i is well-defined.

Let $\bar{n}$ denote the density of a resident population, $\bar{m}$ the density of a mutant population, and f_m the mutant population dynamical equations. If the resident population is at an equilibrium, then the invasion fitness of a rare mutant is given by the leading eigenvalue of the Jacobian $J(f_m(\bar{n}, \bar{m}))|_{\bar{n}=\hat{n}, \bar{m}=0}$, where $\hat{n}$ is the resident equilibrium (Metz *et al.*, 1992). The mutant invades if and only if the leading eigenvalue of the Jacobian is larger than one. If the resident is on a periodic 2-cycle, then invasion fitness can be determined from the stability of the fixed point $f_m^2(\hat{n}, 0)$, where the mapping f_m has been taken twice and $\hat{n}$ is one of the points of the 2-cycle (Alligood *et al.*, 1997). Then, invasion fitness is given by the Jacobian $J(f_m^2(\bar{n}, \bar{m}))|_{\bar{n}=\hat{n}, \bar{m}=0}$. Similarly, if the resident is on a periodic 3-cycle, the invasion fitness is given by the leading eigenvalue of the Jacobian $J(f_m^3(\bar{n}, \bar{m}))|_{\bar{n}=\hat{n}, \bar{m}=0}$.

Size- and age-based emergence strategies

Consistent with the hypothesis that age-based emergence evolved from an ancestral population with the size-based emergence trait, we assume that the resident population follows a size-based emergence strategy. We assume that there exists some threshold body size so that any individual with the size-based trait who grows above that threshold size emerges at the first opportunity. A size-based emergence strategy is characterized by $\eta = (\hat{\eta}_1, \hat{\eta}_2)$, where

$\hat{\eta}_1$ and $\hat{\eta}_2$ are the proportions of the first and second year age classes, respectively, that reach the threshold body size. In order to study the invasion of a rare, periodic mutant with an age-based emergence trait, we define two possible mutant strategies. First, a mutant with a fixed 2-year cycle that is characterized by the strategy $\eta = (0, 1)$: the mutant individuals remain in the ground for exactly 2 years before emerging irrespective of body size. Second, a mutant with a fixed 3-year cycle that is characterized by the strategy $\eta = (0, 0)$: the mutant individuals remain in the ground for exactly 3 years before emerging irrespective of body size.

We assume that fecundity is some increasing function of body size with maximum fecundity attained at the threshold for size-based emergence. Therefore, the fecundity of an individual with the size-based emergence trait is $\beta_{(\hat{\eta}_1, \hat{\eta}_2)} = \beta_{\max}$, where $\beta_{\max}$ is the maximum fecundity. The fecundity of an individual with the age-based emergence trait and a 3-periodic cycle is also $\beta_{(0,0)} = \beta_{\max}$, since all 3-year-old nymphs are assumed to have reached the threshold body size. An individual with the age-based emergence trait and a 2-periodic cycle typically should have an average fecundity that is less than $\beta_{\max}$. This is because typically not all individuals reach the threshold body size by the end of 2 years. Here, we assume a limit case such that only individuals who have reached the threshold body size actually reproduce, i.e. we assume that $\beta_{(0,1)} = (\hat{\eta}_1 + (1 - \hat{\eta}_1)\hat{\eta}_2)\beta_{\max}$. This ensures that the age-based emergence trait does not have any intrinsic advantage over the size-based trait due to some age-based individuals reproducing earlier than their size-based counterparts would have. Then, if we nevertheless discover that an age-based trait can evolve, it is a very robust result, which underestimates the propensity of periodic emergence to evolve rather than overestimating it.

Results

We numerically iterated the model (1–3) for 100,000 rounds to determine the resident population attractor while assuming that the resident population follows the size-based emergence strategy. We repeated the process for different values of predation γ and the growth parameter $\hat{\eta}_2$ while keeping $\hat{\eta}_1$ fixed to produce a bifurcation plot of resident attractors (Fig. 1a). We numerically calculated the Lyapunov exponent of each attractor to determine whether the attractor is an equilibrium, a cycle, an invariant loop, or a strange attractor (Caswell, 2001).

From Fig. 1a we see that if predation γ is low, the resident population always settles to an equilibrium. However, for higher values of γ , the resident population attractor becomes a 2-cycle or a 3-cycle, if $\hat{\eta}_2$ is sufficiently high or low, respectively. For intermediate values of $\hat{\eta}_2$, we do not observe cycles. Naturally, the population is not viable, if predation γ is too high.

In the case shown in Fig. 1a, we assume a low value for $\hat{\eta}_1$. Then, if $\hat{\eta}_2$ has a high value, most individuals emerge at the age of 2 years. Conversely, if $\hat{\eta}_2$ is low, most individuals emerge at the age of 3 years (Fig. 1c). When the population becomes a 2- or 3-cycle, predation suppresses other year classes except for one: since the resident population follows the size-based emergence strategy, some individuals will always emerge each year, but now within a given cycle there is only one year when the emerging population is large enough not to be entirely predated. If $\hat{\eta}_2$ has an intermediate value, any given individual has a reasonably high probability of emerging either at age 2 or age 3. Then, the emergence of a single cohort of newborns is spread across 2 years instead of most of the cohort emerging at the

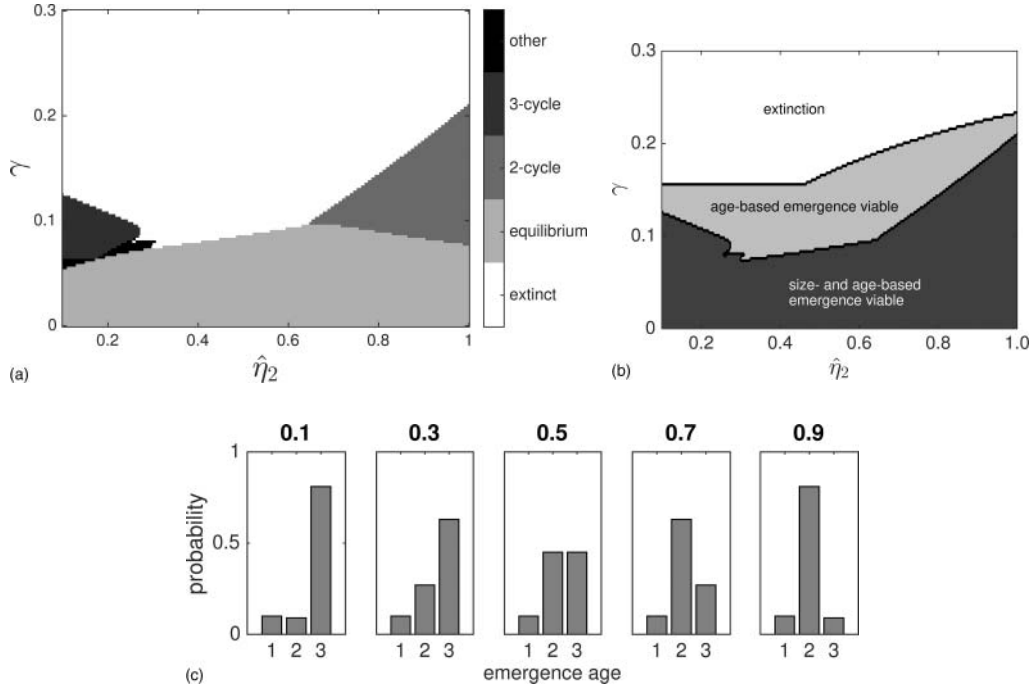


Fig. 1. Numerical analysis of the model (1–3) with $\hat{\eta}_1 = 0.1$ and $\beta_{\max} = 10$. (a) Bifurcation plot of population dynamical attractors for the resident population. (b) Viabilities of size-based and age-based emergence strategies. The viability region of the size-based emergence strategy is given by the dark grey region. Age-based strategies can exist both in the dark grey and light grey regions. No strategy is viable in the white region. (c) Probability distributions of the age at which an individual reaches the threshold body size for different values of $\hat{\eta}_2$ (indicated above each plot).

same time. Thus, the population goes extinct without becoming cyclic, if predation increases sufficiently.

We performed local stability analysis to get more insight into the bifurcations away from equilibria, i.e. for each equilibrium $\hat{n}$, we calculated the leading eigenvalue of the Jacobian $J(f_{\text{res}}(\hat{n}))|_{\hat{n}=\hat{n}}$, where $f_{\text{res}}(\hat{n})$ gives the resident population dynamics. We see that when $\hat{\eta}_2$ is sufficiently high, the transition from an equilibrium to a 2-cycle happens via a flip bifurcation: the equilibrium loses its stability with the leading eigenvalue exiting the unit ball via point $(-1, 0)$. This bifurcation is known to produce a stable 2-cycle (Caswell, 2001).

When $\hat{\eta}_2$ is sufficiently low, the transition from an equilibrium to a 3-cycle is interesting. First, the equilibrium loses stability through a Hopf bifurcation (a complex conjugate pair of leading eigenvalues leaves the unit ball) that creates an invariant loop (Caswell, 2001). However, a sufficient increase in γ then typically causes the dynamics to settle on a stable 3-cycle. Note that in Fig. 1a the black area contains invariant loops and also some long but fixed cycles.

We calculated the invasion fitness for the two periodical mutants for each combination of γ and $\hat{\eta}_2$ for which the resident was at an equilibrium, a 2-cycle or a 3-cycle. We found that the mutants were never able to invade an equilibrium. When the resident is at an

equilibrium, the size of the emerging population each year is constant and therefore there is no benefit in delaying emergence for the 3-cycle mutant. Similarly, for the 2-cycle mutant there is no benefit of delaying emergence for those who reach the threshold body size in the first year. Additionally, we assumed that 2-cycle individuals that did not reach the threshold body size by the end of the second year do not reproduce (but their size-based counterparts would have done so after the third year). However, we found that the 2-cycle mutant was able to invade precisely when the resident was on a 2-cycle and the 3-cycle mutant was able to invade precisely when the resident was on a 3-cycle. When emergence of the resident population is successful in some years only, a mutant strategy that waits to emerge precisely in those years is naturally beneficial: an individual emerging in the wrong year cannot reproduce, whereas an individual who waits for the correct year endures nymphal competition for an extended period, but still has a positive probability of successful reproduction.

We also iterated the model assuming that the resident population follows an age-based emergence strategy and found that these populations can persist under higher levels of predation than populations with the size-based emergence strategy (Fig. 1b).

INDIVIDUAL-BASED MODEL

Here we construct an individual-based model (IBM) to study the evolution of periodicity in the *Magicicada*. In contrast to the mathematical model studied in the previous section, here we incorporate sexual reproduction, explicit genetics, density-dependent predation, and an explicit model for individual growth resulting in variable (and possibly long) life-cycle lengths. In the following we describe the model in detail. A summary of the general flow of the IBM is given in Fig. 2.

A nymph living in the ground undergoes density-dependent competition. The probability to survive a given year is $1/(1 + an_0)$, where a is a coefficient that measures the intensity of competition, and n_0 is the size of the population (total number of individuals) at the beginning of the year (see Appendix for a derivation). Based on the general survival probability, a biased coin-flip is performed for each member of the nymph population to determine whether the individual survives competition or not. Death due to density-independent effects is assumed to be negligible. All surviving nymphs grow by a positive increment, which is drawn independently for each individual from a log-normal distribution

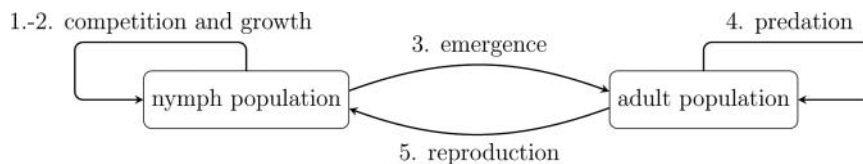


Fig. 2. The IBM goes through five main steps during each iteration of the program corresponding to one year in the cicada life-cycle. (1) Density-dependent competition among nymphs reduces population size. (2) All surviving nymphs grow in size. (3) Each nymph that meets its condition for emergence (threshold size or age) moves from the nymph population to the adult population. (4) Density-dependent mortality for the adult population due to predation. (5) Each individual remaining as part of the adult population reproduces. The offspring are added to the nymph population. All adults are removed from the population after reproduction.

$X = \exp(\mathcal{N}(\mu, \sigma^2))$, where $\mathcal{N}(\mu, \sigma^2)$ is the normal distribution with mean μ and standard deviation σ . The log-normal distribution was chosen as it is common in natural processes (Koch, 1966; Grönholm and Annala, 2007) and has the property of producing only positive values. A log-normal distribution of yearly growth values would follow, e.g. if each year individuals go through a period of exponential growth and the individual rates of growth vary yearly depending on a normally distributed amount of resources.

We assume that nymphs with the size-based emergence trait remain in the ground and feed until they reach a threshold body size, and only then do they emerge. In contrast, nymphs with the age-based emergence trait will wait and feed in the ground until a certain number of years has passed after which they emerge irrespective of their body size. In other words, if an age-based nymph reaches the threshold body size before the year in which it is supposed to emerge, it will remain in the ground and wait until the correct year to emerge.

The emerging population of adult cicadas suffers mortality due to predation. We assume that the probability of death by predation for a given individual is dependent on the size of the emerging population. Specifically, we assume that predation follows a Holling type II functional response (Holling, 1959) that models predator satiation (see Appendix for more details). Based on the general survival probability, a biased coin-flip is performed for each member of the adult population to determine whether the individual survives predation or not. We assume that the effect of feeding on the cicadas is negligible on the production of predator offspring and that the density of the predator population p remains constant throughout a given simulation.

Each adult that survives predation produces offspring. We model sexual reproduction but only keep track of female individuals and assume that enough males exist for reproduction with male genotype ratios identical to the female population. We assume a one-locus, two-allele genetic system where one allele codes for size-based emergence and the other for age-based emergence. We look at both cases of dominance, where either the allele coding for size-based emergence or that for age-based emergence is dominant. If an individual has two alleles coding for age-based emergence but with different cycle lengths, we assume that the shorter cycle dominates. Each new offspring receives one randomly chosen allele from both parents. We assume a fixed mutation probability for both alleles. If an allele coding for age-based emergence mutates, it becomes an allele coding for size-based emergence. If an allele coding for size-based emergence mutates, it becomes an allele coding for age-based emergence for which the life-cycle length is a random integer between 1 and 20. The offspring are deposited back into the ground, where they enter the juvenile (nymph) population. After reproduction, all the adults die. Then, the yearly cycle is repeated from the beginning.

For the simulations we assumed that fecundity increases linearly with respect to size so that the expected number of offspring was equal to the individual's size, e.g. an adult with size 2.5 would have two offspring and have a 50% chance of having a third offspring. Generally, if fecundity is an increasing function of body size, then there is a trade-off between fecundity and probability to survive until emergence, i.e. the longer an individual waits in the ground growing, the more her fecundity increases while the probability to survive until emergence decreases (due to intraspecific competition). Then, if we were to choose an arbitrary size threshold value for the size-based emergence strategy, age-based emergence might prevail simply because it allowed individuals to better optimize their emergence with regard to the fecundity–survival trade-off. However, this is not what we are

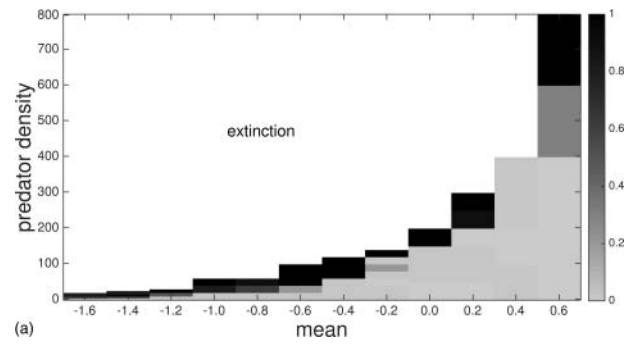
interested in here. Since our hypothesis here is that size-based emergence was an ancestral trait from which age-based emergence evolved, it is likely that size-based emergence would have had ample time to evolve into an evolutionarily stable threshold size before the introduction of age-based emergence (optimizing the size threshold with regard to the fecundity–survival trade-off). Therefore, we let the size-based emergence strategy first evolve such that we can determine an evolutionarily stable size threshold value. Then, we investigate whether a population with the size-based emergence trait and an evolutionarily stable size threshold value can be invaded by a mutation coding for age-based emergence.

We used a computer cluster provided by the University of Jyväskylä to run all of the simulations. We used Matlab (R2015b release) for writing and running the code for the IBM.

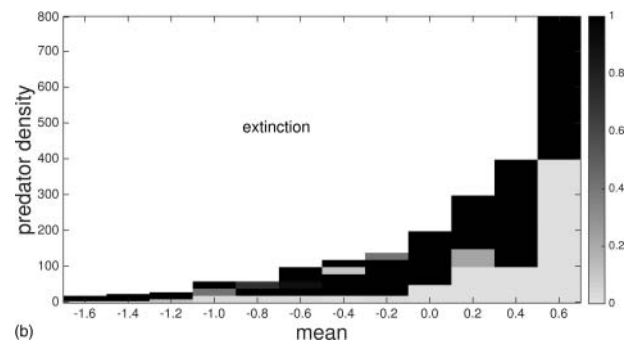
Results

We investigated how different levels of predation and different nymph growth rates (and thus different life-cycle lengths) affect the evolution of periodicity. More specifically, between different simulation runs we varied predator density p and the growth parameters μ and σ (mean and standard deviation, respectively, of the normal distribution underlying the log-normal distribution of individual yearly growth rates) as defined above. First, we determined the evolutionarily stable threshold size for size-dependent emergence. We ran five independent simulations for each combination of p and μ (with σ fixed) and instead of letting mutations change a size-based trait into age-based emergence, mutations affected the size threshold value of newborns. We started each simulation with a polymorphic initial population that had a uniform distribution of a range of threshold values represented in the population. The simulations typically converged so that at the end of each simulation all individuals had similar size threshold values. Then, we calculated the average end value for the size threshold for each parameter combination. We observe that generally increasing predator density p causes the evolutionarily stable size threshold value to increase. Similarly, smaller growth rates and thus longer life cycles resulted in a larger evolutionarily stable body size threshold value.

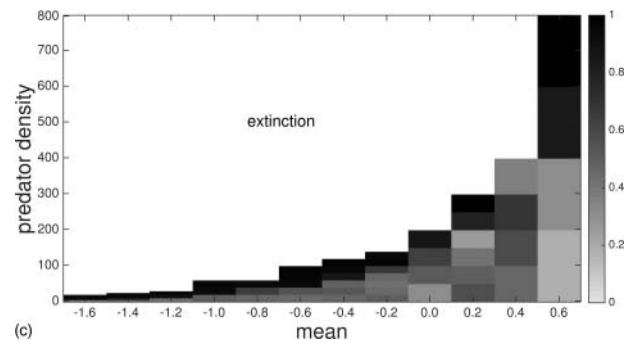
Next, we investigated whether age-based emergence could invade a resident population with the size-based emergence trait. We performed ten simulation runs for each combination of predator density p and the nymph growth parameter μ . We began each simulation run with an initial population that consisted of homozygote individuals with the size-based emergence trait with the specific size thresholds obtained as described above. The initial population was given an even distribution of all age classes $\{0, \dots, k\}$, where k was one less than the average age of emergence. An individual of a particular age class was given a starting body size equal to the expected body size at that age. Since our hypothesis here is that age-based emergence evolved from an ancestral population with the size-based emergence trait, we assume that the initial population has reached a population dynamical attractor before the age-based emergence trait is introduced. Therefore, we first ran each simulation for 2000 rounds with the mutation probability set to zero (so that the resident population had time to settle to a population dynamical attractor before mutations could occur). After 2000 rounds, we set the mutation probability to a fixed positive value and then the simulation was continued for another 98,000 rounds. We also started each simulation with predator density p set to zero and then progressively increased that density so that after



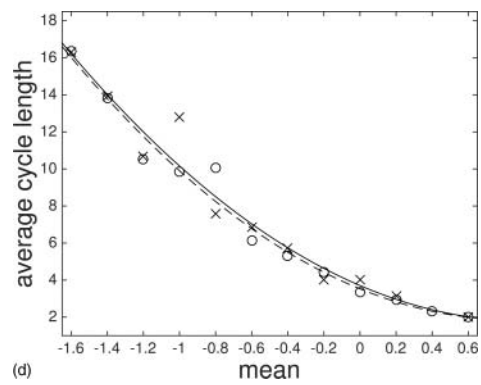
(a)



(b)



(c)



(d)

1000 rounds it had reached the desired level (this allowed the resident population to settle smoothly into a population dynamical attractor).

The simulation results are shown in Fig. 3. Generally, we observe that if there is no predation, the age-based emergence trait cannot invade a resident population with size-based emergence. However, if predator density is increased sufficiently, then invasion becomes possible. If predator density is increased too much, the resident population with a size-based emergence trait becomes non-viable. With slow growth rates population density decreases, because longer life cycles require individuals to endure intraspecific competition for longer periods (but this might change, for example, if fecundity was a non-linear function of size). Then, also reducing growth rates can make the invasion of age-based emergence possible. However, if growth rates are too small, then again the resident population with a size-based emergence trait becomes non-viable. We performed test simulations starting with an age-based resident population and found that generally populations with the age-based emergence trait were able to survive under higher levels of predation than their size-based counterparts.

We make the following observations about the dynamics of the resident size-based emergence population before mutations for the age-based emergence are allowed. We observe that if there is no predation, then the density of emerging adults is roughly constant each year with some natural variation due to the growth process being stochastic (Fig. 4a). If predator density is increased sufficiently, typically this causes the resident size-based emergence population to become proto-periodical, i.e. the density of the emerging adult population begins to oscillate from year to year in a distinguishable but not yet perfect pattern (Fig. 4b). To measure the level of proto-periodicity in the resident population, we measured the autocorrelation of the adult emergence density for the final 200 rounds before mutations were allowed to occur, i.e. the measurement was done after the resident population had reached a population dynamical attractor. To measure the autocorrelation, we used the autocorr function provided by Matlab. Figure 3c shows the average value of the highest correlation frequency for each μ, p combination.

Fig. 3. Results of the IBM simulations. Parameter values used for the simulations: $\sigma = 0.3$, $a = 0.0001$, $\alpha = 1$, $h = 1$. Initial population size was set to 2000 and the mutation probability (per allele) was 0.001. (a, b) Average proportion of the population with an age-based emergence trait at the end of the simulations is shown for each combination of p (predator density) and μ (mean). Low mean values indicate slow average growth whereas high values indicate fast growth. Black indicates that all simulations ended with the entire population having evolved to have age-based emergence. Light grey indicates that all simulations ended with the entire population having the size-based emergence trait. The white region indicates where the initial, size-based population is not viable. Panel (a) shows simulation results for the case when age-based emergence is assumed to be a recessive trait while panel (b) shows results for the case when it is assumed to be a dominant trait. (c) Correlation value of the highest autocorrelation frequency of the resident population emergence size measured for the final 200 simulation rounds before mutations were allowed to occur. Black regions indicate that the resident population has become highly proto-periodical, while the light grey regions indicate that the resident population is at a (stochastic) equilibrium. (d) Average cycle length (number of years) of the age-based trait in cases when the population evolved to have age-based emergence. Crosses indicate results for the case when age-based emergence is recessive; open circles indicate the case when age-based emergence is dominant. The solid line shows a quadratic curve fitted to the data of the recessive case and the dashed line shows a quadratic curve fitted to the data of the dominant case.

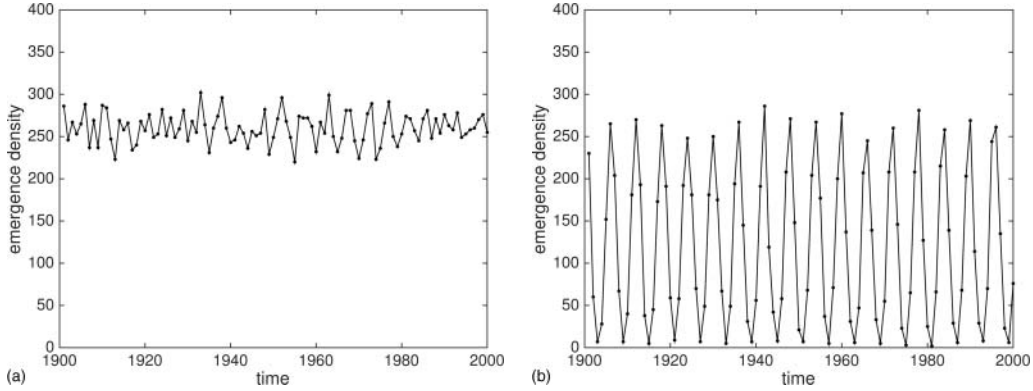


Fig. 4. Examples of resident population emergence densities over time with different levels of predation. Parameter values used for the simulations: $\mu = -0.6$, $\sigma = 0.3$, $\alpha = 0.0001$, $a = 1$, $h = 1$. (a) No predation ($p = 0$). Emergence density is roughly constant each year with some natural variation. (b) Relatively high predator density ($p = 80$). Emergence densities fluctuate periodically.

We observe that when the age-based emergence allele is recessive, invasion occurs only when the resident population already exhibits strong proto-periodicity (see Figs. 3a and 3c). However, when the age-based emergence allele is dominant, then invasion becomes much more common (Fig. 3b). Once the invasion of age-based emergence is complete, size-based emergence cannot invade back: individuals that do not emerge in synchrony with the rest of the population perish due to predation before having an opportunity to reproduce. We also observe that when the age-based emergence trait invades, its cycle length is naturally dependent on the nymph growth rates: when growth is slower and thus life cycles are longer, the resident proto-periodical cycles become longer and therefore also the invading age-based trait has a corresponding, longer cycle length (Fig. 3d).

Finally, we also tested the effect of varying the growth parameter σ on the evolution of periodicity. Here, we use a fixed value for the growth parameter μ and vary both σ and predator density p . As before, we first determined the evolutionarily stable threshold body size for each parameter combination and then ran simulations to determine whether an age-based trait would be able to invade a size-based resident population. The results are shown in Fig. 5a. We see that increasing σ reduces the ability of age-based traits to invade. Similar to the mathematical population model studied above, where periodicity did not evolve for intermediate values of $\hat{\eta}_2$, we believe the reason for this is that higher values of σ result in greater variance in the probability distribution of emergence ages for the size-based trait (Figs. 5c–e). Then, the size-based resident population simply cannot become (strongly) proto-periodical (Fig. 5b) as the emergence of the newborns of any given year is spread out over several years and no set of newborns is able sufficiently to remain as a single cohort. Thus, invasion of the age-based trait is not possible for sufficiently large σ .

DISCUSSION

In this paper, we show how the periodical cicadas (*Magicicada* spp.) may have evolved from non-periodical cicadas by switching from size-based emergence to age-based emergence. To do this, we develop two models of the cicada life cycle that complement each other: one a

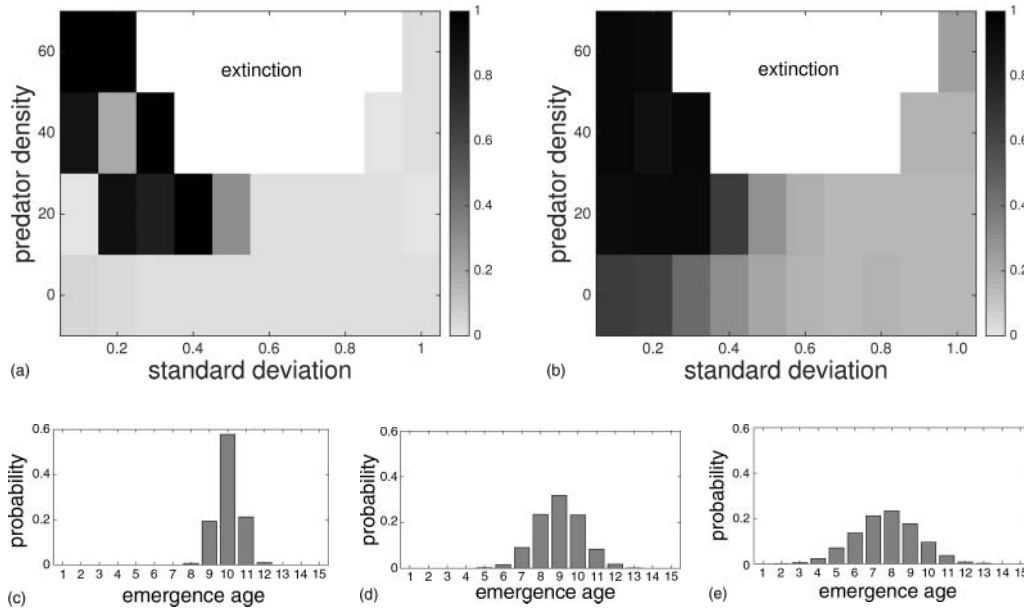


Fig. 5. Results of the IBM simulations. Parameter values used for the simulations: $\mu = -1.0$, $\alpha = 0.0001$, $a = 1$, $h = 1$. (a) Average proportion of the population with an age-based emergence trait at the end of the simulations shown for each combination of p (predator density) and σ (standard deviation). Age-based emergence is here assumed to be a recessive trait. In the figure, black indicates that all simulations ended with the entire population having evolved to have age-based emergence. Light grey indicates that all simulations ended with the entire population having the size-based emergence trait. The white region indicates where the initial, size-based population is not viable. (b) Correlation value of the highest autocorrelation frequency of the resident population emergence size measured for the final 200 simulation rounds before mutations were allowed to occur. Black regions indicate that the resident population has become highly proto-periodical, while the light grey regions indicate that the resident population is at a (stochastic) equilibrium. (c–e) Probability distributions of the age at which an individual reaches the threshold body size for different values of σ : (c) $\sigma = 0.2$, (d) $\sigma = 0.4$, (e) $\sigma = 0.6$.

simplified mathematical population model and the other a more realistic individual-based model (IBM). Our analysis of the mathematical model allows us to demonstrate that sufficiently high levels of predation relative to cicada population size can trigger a switch from size-based to age-based emergence. Simulations of the IBM concur with this finding and provide further evidence for the robustness of the result.

With both models, we start with an initial population where individuals emerge when they reach some predefined body size threshold (size-based emergence). We assume in both models that the yearly growth of individual nymphs in the ground is a stochastic process with some variance. Due to this variance in growth rates, individuals reach the threshold body size and emerge at different times, and thus the resident population is generally non-periodical. However, if predation increases sufficiently relative to the size of the emerging adult population, then the whole population can become proto-periodical, i.e. population density oscillates with reproduction repeatedly being reduced or completely suppressed in

some years. In these circumstances, a mutation coding for periodic emergence (age-based emergence) in years of high emergence density can invade. Although here we explicitly model predation, other mechanisms that create an Allee effect on the emerging adult cicada population, such as density-dependent fecundity or mating success (Karban, 1981), could also be responsible for inducing proto-periodicity.

We find that increasing predator density relative to the size of the cicada population can lead to the evolution of periodicity through a switch from size-based to age-based emergence. Generally, this switch is possible near the boundary of extinction (Ito *et al.*, 2015); that is, a switch to age-based emergence occurs when predator density is sufficiently high that the size-based population would not be able to exist if predator density increased (or cicada population size decreased) much further. In contrast, the age-based population typically can be sustained under higher levels of predation because it exhibits larger population sizes on emergence. However, it is important to understand that proto-periodicity of the resident size-based population is a necessary condition for a successful invasion of the age-based trait (but see also the next paragraph). If there is too much variance on average growth rates and the emergence ages are too widely spread over several years, it may not be possible for the size-based population to become proto-periodical even under increasing predation pressure and the population may become extinct without any possibility to evolve periodical emergence.

Contrary to Ito *et al.* (2015), we find that assumptions on the underlying genetics determining the trigger for adult emergence make a difference for when the evolution of periodicity is possible. If age-based emergence is triggered by a recessive allele, then generally it can only invade a size-based resident population, when the resident population has already become strongly proto-periodical. However, if the age-based emergence allele is dominant, then it can invade under a wider variety of circumstances. In particular, the resident size-based population does not necessarily have to be strongly proto-periodical. However, generally the age-based emergence trait cannot invade if there is no or very little predation relative to the size of the cicada population.

We assume that fecundity is an increasing function of body size. Then, each year there are conflicting pressures on whether an individual should emerge: staying in the ground to grow increases fecundity further, but it also reduces the probability of survival until reproduction due to intraspecific competition. In other words, typically the expected value (expected number of offspring produced) of staying in the ground is initially positive, but the bigger an individual becomes, the more she stands to lose by risking staying an extra year in the ground. Then, generally there exists an optimal threshold size for emergence so that up until that size an individual has a positive expected value to stay in the ground growing and beyond that size the expectation becomes negative (so that one should emerge). If the body size threshold for the resident size-based emergence population is not optimal, then age-based emergence could invade simply due to there being a better balance between fecundity and mortality. Moreover, the body size threshold must be under selection itself: any mutant individual with a size threshold closer to the optimum relative to the resident population threshold would be positively selected. Since the size-based emergence trait is assumed to be the ancestral trait that existed before age-based emergence, it would likely have had ample time to evolve into an evolutionarily stable threshold size. This is an aspect that was not considered in the paper by Ito *et al.* (2015), in which they assume an arbitrary size threshold for maturity, but then allow age-based individuals to produce offspring in proportion to their body size even if they have not reached the threshold body

size. Then, the arbitrary choice of the threshold body size is typically not the optimal threshold with regard to the survival–fecundity trade-off and it is generally evolutionarily unstable.

Based on previous concerns, it is not clear whether in the model of Ito *et al.* (2015) age-based emergence evolves because of periodicity *per se* or because it allows for a better fecundity–survival trade-off. In particular, Ito *et al.* find that periodicity can evolve even when there is no Allee effect affecting the emerging adult cicada population (see, for example, Figure S4 of Ito *et al.*, 2015), which is not possible in our model. Furthermore, we suspect that Ito *et al.* only find periodicity evolving in cold environments (when nymphal growth is slow) because they only consider periodic cycle lengths of at least 10 years. We find that periodicity can evolve in all environments and the cycle length of the invading periodical population depends on the environment: short cycles are possible in warm climates (when nymphal growth is fast) and long cycles are possible in cold environments (when nymphal growth is slow).

In order to study the benefits of periodicity itself, we have set up our models so that age-based emergence is never intrinsically favoured. This is done in two ways: for the mathematical model we assume that reproduction is possible only for individuals that have reached the threshold body size required for size-based emergence, and for the IBM we first establish what the optimal body size threshold is for each environment before we run simulations that allow mutations for age-based emergence. This allows us to demonstrate that age-based emergence is selected for in environments where the resident population following size-based emergence becomes proto-periodical. Furthermore, it is worth noting that because of our model setup, we are more likely to underestimate the evolutionary potential of periodicity than to overestimate it.

A periodical population following the age-based emergence strategy is robust in the sense that a size-based emergence strategy cannot invade if the level of predation remains sufficiently high (and this level may be lower than what was needed for periodicity to evolve originally): any individual not in synchrony with the main population will be eliminated due to predation when it emerges. For the same reason, periodically emerging mutants with a wrong cycle length would not be able to invade. We also find that periodical populations can persist under higher levels of predation pressure than non-periodical populations because periodical populations typically exhibit larger emergence sizes. However, if predation intensity was to decrease sufficiently, then it could become possible for the size-based emergence strategy to invade.

In this paper, we do not consider why the *Magicicada* life cycles are precisely the prime numbered 13 and 17 years. However, understanding how periodicity may evolve is a first step towards understanding why precisely those periods evolved in the *Magicicada*. We show that non-periodical cicadas can evolve into periodical cicadas with any cycle length, and the precise length is mostly determined by average nymph growth rates. In our model, cicada population densities are higher in warm climates (when nymphal growth is fast) and lower in cold climates. Also, the level of predation required to initiate the evolution of periodicity is high in warm climates and it becomes progressively smaller in colder climates. This is a potential explanation for the long periodic cycles of the *Magicicada* (and why contemporary short-lived annual cicadas remain non-periodical), i.e. the long periods developed in a cold climate and the reduced population densities made cicadas more vulnerable to predation. This resulted in proto-periodicity of the ancestral size-based cicada population and allowed for the invasion of age-based emergence. Our results are consistent

with the hypothesis that periodicity in the *Magicicada* developed during the Pleistocene glacial period (Cox and Carlton, 1988; Yoshimura, 1997).

Some authors have suggested that periodicity in the *Magicicada* evolved independently in several different populations and the populations with prime numbered cycles were selected for due to hybridization (Cox and Carlton, 1988, 2003; Yoshimura *et al.*, 2009): hybrid offspring of two different cycles would be lost due to their emergence not being in synchrony with either parent cycle and prime-numbered cycles co-emerge less often with other cycles. While we believe that hybridization could be an important factor selecting for prime-numbered cycles, we do not think that the issue has been fully resolved. For example, if there are two periodic populations with even cycle lengths such that one emerges in even years and the other in odd years, then the two populations never co-emerge (Lehmann-Ziebarth *et al.*, 2005)! Some type of environmental stochasticity causing occasional temporal shifts in emergence could possibly resolve the issue, and we encourage further research into the matter.

ACKNOWLEDGEMENTS

This research was funded by the Academy of Finland (grant #283486). We acknowledge grants of computer capacity from the Finnish Grid and Cloud Infrastructure (persistent identifier urn:nbn:fi:research-infras-2016072533).

REFERENCES

- Alexander, R.D. and Moore, T.E. 1962. The evolutionary relationships of 17-year and 13-year cicadas, and three new species (Homoptera, Cicadidae, *Magicicada*). *Misc. Publ. Mus. Zool. Univ. Mich.*, **121**: 1–59.
- Alligood, K.T., Sauer, T.D. and Yorke, J.A. 1997. *Chaos: An Introduction to Dynamical Systems*. New York: Springer.
- Behncke, H. 2000. Periodical cicadas. *J. Math. Biol.*, **40**: 413–431.
- Beverton, R.J.H. and Holt, S.J. 1957. *On the Dynamics of Exploited Fish Populations*. Caldwell, NJ: The Blackburn Press, 1993 (reprint).
- Bulmer, M.G. 1977. Periodical insects. *Am. Nat.*, **111**: 1099–1117.
- Caswell, H. 2001. *Matrix Population Models: Construction, Analysis and Interpretation*, 2nd edn. Sunderland, MA: Sinauer Associates.
- Cox, R.T. and Carlton, C.E. 1988. Paleoclimatic influences in the evolution of periodical cicadas (Insecta: Homoptera: Cicadidae: *Magicicada* spp.). *Am. Midl. Nat.*, **120**: 183–193.
- Cox, R.T. and Carlton, C.E. 2003. A comment on gene introgression versus en masse cycle switching in the evolution of 13-year and 17-year life cycles in periodical cicadas. *Evolution*, **57**: 428–432.
- Davydova, N.V., Diekmann, O. and van Gils, S.A. 2003. Year class coexistence or competitive exclusion for strict biennials? *J. Math. Biol.*, **46**: 95–131.
- Davydova, N.V., Diekmann, O. and van Gils, S.A. 2005. On circulant populations. I. The algebra of semelparity. *Linear Algebra Appl.*, **398**: 185–243.
- Diekmann, O., Davydova, N. and van Gils, S. 2005. On a boom and bust year class cycle. *J. Differ. Equ. Appl.*, **11**: 327–335.
- Dybas, H.S. and Davis, D.D. 1962. A population census of seventeen-year periodical cicadas (Homoptera: Cicadidae: *Magicicada*). *Ecology*, **43**: 432–444.
- Dybas, H.S. and Lloyd, M. 1962. Isolation by habitat in two synchronized species of periodical cicadas (Homoptera: Cicadidae: *Magicicada*). *Ecology*, **43**: 444–459.
- Dybas, H.S. and Lloyd, M. 1974. The habitats of 17-year periodical cicadas (Homoptera: Cicadidae: *Magicicada* spp.). *Ecol. Monogr.*, **44**: 279–324.

- Eskola, H.T.M. and Geritz, S.A.H. 2007. On the mechanistic derivation of various discrete-time population models. *Bull. Math. Biol.*, **69**: 329–346.
- Geritz, S.A.H. and Kisdi, E. 2004. On the mechanistic underpinning of discrete-time population models with complex dynamics. *J. Theor. Biol.*, **228**: 261–269.
- Grönholm, T. and Annala, A. 2007. Natural distribution. *Math. Biosci.*, **210**: 659–667.
- Heliövaara, K., Väisänen, R. and Simon, C. 1994. Evolutionary ecology of periodical insects. *Trends Ecol. Evol.*, **9**: 475–480.
- Holling, C.S. 1959. Some characteristics of simple types of predation and parasitism. *Can. Entomol.*, **91**: 385–398.
- Hoppensteadt, F.C. and Keller, J.B. 1976. Synchronization of periodical cicada emergences. *Science*, **194**: 335–337.
- Ito, H., Kakishima, S., Uehara, T., Morita, S., Koyama, T., Sota, T. *et al.* 2015. Evolution of periodicity in periodical cicadas. *Sci. Rep.*, **5**: 14094 (doi: 10.1038/srep14094).
- Karban, R. 1981. Effects of local density on fecundity and mating speed for periodical cicadas. *Oecologia*, **51**: 260–264.
- Karban, R. 1982. Increased reproductive success at high densities and predator satiation for periodical cicadas. *Ecology*, **63**: 321–328.
- Karban, R. 1984. Opposite density effects of nymphal and adult mortality for periodical cicadas. *Ecology*, **65**: 1656–1661.
- Koch, A.L. 1966. The logarithm in biology. I. Mechanisms generating the log-normal distribution exactly. *J. Theor. Biol.*, **12**: 276–290.
- Kon, R. and Iwasa, Y. 2007. Single-class orbits in nonlinear Leslie matrix models for semelparous populations. *J. Math. Biol.*, **55**: 781–802.
- Lehmann-Ziebarth, N., Heideman, P.P., Shapiro, R.A., Stoddart, S.L., Hsiao, C.C.L., Stephenson, G.R. *et al.* 2005. Evolution of periodicity in periodical cicadas. *Ecology*, **86**: 3200–3211.
- Lloyd, M. and Dybas, H.S. 1966a. The periodical cicada problem. I. Population ecology. *Evolution*, **20**: 133–149.
- Lloyd, M. and Dybas, H.S. 1966b. The periodical cicada problem. II. Evolution. *Evolution*, **20**: 466–505.
- Marshall, D.C. and Cooley, J.R. 2000. Reproductive character displacement and speciation in periodical cicadas, with description of a new species, 13-year *Magicalcica neotredecim*. *Evolution*, **54**: 1313–1325.
- Martin, A. and Simon, C. 1990. Temporal variation in insect life cycles: lessons from periodical cicadas. *Bioscience*, **40**: 359–367.
- Metz, J.A.J., Nisbet, R.M. and Geritz, S.A.H. 1992. How should we define ‘fitness’ for general ecological scenarios? *Trends Ecol. Evol.*, **7**: 198–202.
- Sota, T., Yamamoto, S., Cooley, J.R., Hill, K.B.R., Simon, C. and Yoshimura, J. 2013. Independent divergence of 13- and 17-y life cycles among three periodical cicada lineages. *Proc. Natl. Acad. Sci.-Biol.*, **110**: 6919–6924.
- White, J.A. and Lloyd, M. 1975. Growth rates of 17- and 13-year periodical cicadas. *Am. Midl. Nat.*, **94**: 127–143.
- White, J.A., Lloyd, M. and Zar, J.H. 1979. Faulty eclosion in crowded suburban periodical cicadas: populations out of control. *Ecology*, **60**: 305–315.
- Williams, K.S. and Simon, C. 1995. The ecology, behavior, and evolution of periodical cicadas. *Annu. Rev. Entomol.*, **40**: 269–295.
- Yoshimura, J. 1997. The evolutionary origins of periodical cicadas during Ice Ages. *Am. Nat.*, **149**: 112–124.
- Yoshimura, J., Hayashi, T., Tanaka, Y., Tainaka, K. and Simon, C. 2009. Selection for prime-number intervals in a numerical model of periodical cicada evolution. *Evolution*, **63**: 288–294.

APPENDIX

Nymph mortality due to competition

In both the ‘Mathematical model’ and ‘Individual-based model’ sections of this paper, we assume that yearly mortality of nymphs in the ground follows a simple model of continuous-time interference competition

$$\frac{dn}{dt} = -an^2. \quad (\text{A1})$$

Assuming $n(t) > 0$ for all t , the solution to (A1) is

$$n(t) = \frac{n_0}{1 + an_0 t}, \quad (\text{A2})$$

where $n(0) = n_0$ is the density of the nymph population at the beginning of the year. Then, the fraction of individuals surviving after one (arbitrary) unit of time is

$$\frac{n(1)}{n(0)} = \frac{1}{1 + an_0}, \quad (\text{A3})$$

which is the form of Γ used in (4) and which gives the yearly survival probability for each individual in the IBM.

Density-independent predation

In the ‘Mathematical model’ section, we assume that mortality due to predation is independent of cicada density and the change in cicada density over time is given by

$$\frac{dn}{dt} = -ap, \quad (\text{A4})$$

where a is the attack rate of the predator and p is predator density, which is assumed to be constant. The solution to (A4) is

$$n(t) = n_0 - apt, \quad (\text{A5})$$

where $n(0) = n_0$ is the initial density of the nymph population above ground immediately after emergence. Then, the density of surviving individuals after one (arbitrary) unit of time is

$$n(1) = \max\{0, n_0 - ap\}, \quad (\text{A6})$$

which is the form for the density of reproducing individuals used in (1) with $(\eta_1 n_1 + \eta_2 n_2 + n_3) \Gamma(\bar{n}) = n_0$ and $\gamma = ap$.

Density-dependent predation

In the ‘Individual-based model’ section, we assume that the density of emerging adult cicadas is reduced according to the following equation:

$$\frac{dn(t)}{dt} = -\frac{a n(t) p}{1 + a h n(t)}, \quad (\text{A7})$$

where n is the cicada population density, a is the attack rate of the predators, h is the handling time (time spent processing a captured cicada), and p is the predator population density. To calculate the probability of survival for each individual cicada, we take equation (A7) with initial condition $n(0)$ equal to the emerging cicada population size and then numerically solve the dynamics for one (arbitrary) unit of time. Then, we take the fraction of surviving cicadas $n(1)$ over the original emergence size as the probability of survival for any given individual, i.e. the probability of a given adult cicada to survive predation is $n(1)/n(0)$. Effectively, this process introduces an Allee effect on the emerging adult cicada population, whereby small (relative to predator density) emerging populations are likely to be wiped out by predation before they have the opportunity to mate.

