Research

Ecography 33: 809–817, 2010
doi: 10.1111/j.1600-0587.2009.06018.x
© 2010 The Authors. Journal compilation © 2010 Ecography
Subject Editor: Pedro Peres-Neto. Accepted 30 November 2009

Comparing methods for measuring the rate of spread of invading populations

Marius Gilbert and Andrew Liebhold

M. Gilbert (mgilbert@ulb.ac.be), Biological Control and Spatial Ecology, Univ. Libre de Bruxelles CP160/12, Av FD Roosevelt 50, BE-1050 Brussels, Belgium, and Fonds National de la Recherche Scientifique, rue d'Egmont 5, BE-1000 Brussels, Belgium. – A. Liebhold, Northern Research Station, USDA Forest Service, Canfield St. 180, Morgantown, WV 26505, USA.

Measuring rates of spread during biological invasions is important for predicting where and when invading organisms will spread in the future as well as for quantifying the influence of environmental conditions on invasion speed. While several methods have been proposed in the literature to measure spread rates, a comprehensive comparison of their accuracy when applied to empirical data would be problematic because true rates of spread are never known. This study compares the performances of several spread rate measurement methods using a set of simulated invasions with known theoretical spread rates over a hypothetical region where a set of sampling points are distributed. We vary the density and distribution (aggregative, random, and regular) of the sampling points as well as the shape of the invaded area and then compare how different spread rate measurement methods accommodate these varying conditions. We find that the method of regressing distance to the point of origin of the invasion as a function of time of first detection provides the most reliable method over adverse conditions (low sampling density, aggregated distribution of sampling points, irregular invaded area). The boundary displacement method appears to be a useful complementary method when sampling density is sufficiently high, as it provides an instantaneous measure of spread rate, and does not require long time series of data.

Understanding the mechanisms operating during the introduction, establishment and spread of invading organisms is critical to prevent their sometimes dramatic consequences on agricultural, forest and urban ecosystems (Sakai et al. 2001, Liebhold and Tobin 2008). In many cases, however, biological invasions are detected at a time when eradication is no longer possible and efforts are directed toward reducing and/or delaying the invader's spread and ultimate impact (Sharov et al. 2002a, b, Tobin et al. 2004). These later efforts rely on a good understanding of the invading species' population biology, and in particular on the factors influencing spread into new environments (Sharov et al. 1999, Gilbert et al. 2004, Tobin et al. 2007a). An accurate measure of the spread rate is therefore important for evaluating the effect of different conditions and management activities on the rate of invasion, as well as providing a method for predicting where and when the invader will spread in the future.

Several methods have been used to measure invasion spread rates (Liebhold et al. 1992, Shigesada and Kawasaki 1997), but these methods have not been comprehensively compared for the accuracy and variability of their estimates. Recent studies in the dispersal ecology literature have highlighted that inference of parameter values may be influenced by the selection of the estimation method (Edwards et al. 2007). One particular problem of comparing methods for measuring spread rates is that even with abundant empirical data, the true spread rate is never known. For example, Tobin et al. (2007b) compared two methods to measure the rate of spread of the gypsy moth *Lymantria dispar*, and found no substantial difference in estimates, but since they had no knowledge of the true rate of spread, they are unable to assess whether both methods were similarly accurate, or similarly biased. Likewise, Evans and Gregoire (2007) compared three regression techniques to derive the rate of spread from empirical data on the spread of hemlock woolly adelgid *Adelges tsugae* but again the true rate of spread was unknown.

The approach taken in this study was to compare the performances of several spread rate measurement methods using a set of simulated invasion data sets with known theoretical spread rates. In these data sets, spread was simulated using Fisher's (1937) model of invasion (diffusion and population growth limited by a carrying capacity), with the incorporation of local, temporal and spatial stochasticity (Shigesada and Kawasaki 1997). We simulated invasion into a hypothetical geographical region, and simulated sampling the geographical extent of populations with various densities and distributions of sample locations (e.g. representing county-level observations, or insect traps),

and various spatial configurations of the invaded region. We compared how measures of the spread rate and their variability were affected by the choice of spread rate measurement method.

Methods and materials

Invasion models

1) Fisher stochastic model

We employed Fisher's (1937) invasion model to simulate spread with a known rate. Fisher's model differs from the more widely-known reaction-diffusion model of Skellam (1951), in that it accounts for a limited carrying capacity whereas Skellam's model assumes unlimited exponential growth. The deterministic formulation of the model is divided into part, 1) population growth following Ricker's (1954) formulation of a classic discrete logistic population model (eq. 1), 2) diffusion according to Fick (1855) (eq. 2).

$$N'_{j,t+1} = N_{j,t} e^{r(1-N_{j,t}/K_i)} \tag{1}$$

$$N_{i,t+1} = \sum_{j=1}^{n} N'_{j,t+1} \times \frac{e^{\frac{-d_{i,j}^2}{4D}}}{4\pi D} \tag{2}$$

where N_t and N_{t+1} are populations at time t and $t+1$, r is the growth rate, K_i is the carrying capacity in point i, $d_{i,j}$ is the distance between point i and j, and D is the diffusion coefficient. The model was discrete in time and space, with space as a spatial grid of 1000 by 1000 cells. For ease of interpretation, we set each cell as being 1 km^2. One of the computer processing limitations in implementing Fishers' model in a discrete space is the need to build the redistribution matrix used to estimate the population spreading from each cell to every other cell in the spatial grid (eq. 2). Performing a simulation over a 1000×1000 cell spatial grid would result in the need to build and store a $10^6 \times 10^6$ redistribution matrix. Even coded as one byte per cell, the redistribution matrix would have a size of one terabyte, which is far beyond the internal memory capacity of a personal computer such as employed in this study. However, there is a vast amount of wasted memory in the classic redistribution matrix because the majority of individuals dispersing from a given cell spread to the closest cells (depending on the diffusion coefficient), i.e. not to others more distant ones, and a large amount of cells are actually empty. So, we proposed a moving-window approach to circumvent the problem, in which the $<0.1\%$ of individuals spreading over distance larger than the 99.9% percentile (according to the Gaussian dispersal kernel, Kot et al. 1996) was ignored.

We added three types of stochasticity to the model: temporal stochasticity (to account for adverse/favorable conditions affecting the entire population in a given time step), spatially correlated stochasticity (to account for spatially correlated temporal variation in habitat suitability) (Hudson and Cattadori 1999), and uncorrelated random spatial noise (to account for local variation in population growth as well as measurement error). These different types of stochasticity were included in the model in eq. 3 and eq. 4

$$N'_{j,t+1} = N_{j,t} e^{r(1-N_{j,t}/K_i)} \times u \tag{3}$$

$$u = e^{\varepsilon_t + \varepsilon_l + \varepsilon_s} \tag{4}$$

where ε_t is temporal stochasticity (same value for all cells of the matrix, but different values for each time step) following a normal distribution $(0,\sigma_t)$, ε_l is the local stochasticity $\sim(0,\sigma_l)$ (different value in each cell and each time step), and ε_s is the spatially correlated stochasticity $\sim(0,\sigma_s)$ (same values for each time step, different values in each cell, but closer cells have closer values than those distant apart). The default values for standard deviations corresponding to temporal, local and spatial stochasticity were respectively 0.20, 0.20 and 0.4. Spatially correlated stochasticity was generated using unconditional Gaussian simulation that generates values distributed over space that follow a random normal frequency distribution with a specified spatial covariance matrix. A spherical model was used to model the spatial covariance structure of unconditional Gaussian simulation of autocorrelated noise, with a range of 100 km, a sill equal to 1, and no nugget effect (Supplementary material Fig. S1). More details on unconditional Gaussian simulation can be found in Pebesma (2004).

Invading populations were initialized at the centre cell of the grid with a population level of 100 individuals (a population level of zero was assigned to all other grid cells). We assumed a growth rate $r = 1.0986$, a carrying capacity $K_i = 1000$ for all locations, and a diffusion coefficient $D = 10$. The invasion model was iterated over 70 time steps. Default parameter values were chosen such that invading populations did not spread beyond the 1000×1000 km domain, exhibited noticeable irregularity in the spatial area of invasion, and exhibited population fluctuations around the carrying capacity yet did not lead to frequent local extinctions (Fig. 1). Supplementary material Fig. S2 illustrates the spatio-temporal distribution of simulated population in one realization of the simulation.

2) Circular deterministic model

One problem arising with the Fisher stochastic model was that the spread rate of simulations was known, but only approximately. The radial rate of spread produced by the Fisher model is only approximated by $2\sqrt{rD}$ (Shigesada and Kawasaki 1997). Consequently, the "expected" spread rate arising from our model had to be estimated for each simulation by regressing the square root of invaded area against time based on the simulations in the entire 1000×1000 cells matrix. Potentially, this could confuse the interpretation of results because of ambiguity about the baseline of expected spread rate to which each simulated spread rate measurement method should be compared to. We addressed this problem by also utilizing a separate spread rate model in which spread was idealized simply as circles of increasing radius over time. Population levels of cells inside the circle were set to the carrying capacity $K_i = 1000$ and population levels in cells outside the circle were equal to 0. At each time step, the circle radius was incremented by the theoretical spread rate $2\sqrt{rD} = 6.629$. Simulations using concentric circles obviously did not incorporate any stochasticity. All spread rate measurement methods were applied to sample values extracted from both invasion models.

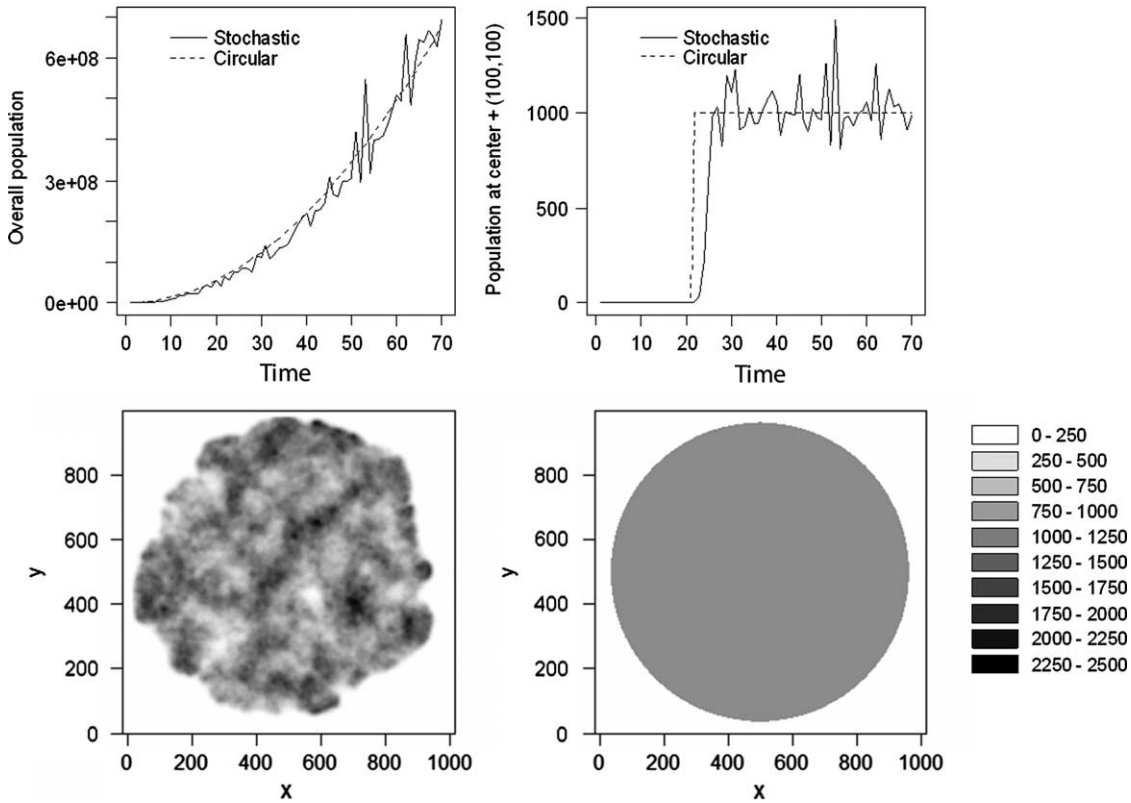

Figure 1. Change in total population over time (top left), change in population at point (+100, +100) from the centre (top right), distribution obtained by the stochastic (bottom left) and circular (bottom right) invasion models at the last time step. All figures are derived from one realization of the simulated invasion over 70 time steps in a 1000 × 1000 cells domain, with default parameters.

Idealized sampling conditions

Ideally, surveys to monitor the spread of an invading species should cover the entire region, with spaced sampling locations according to a particular design, and the sampling density should be high enough to adequately estimate the position of the invasion boundary. However, in practical situations, surveillance data are often clustered (e.g. more observations are located in highly populated areas), or sparsely distributed. Empirical data may also cover an area of irregular shape, spatially constrained by areas that can not be invaded (e.g. sea, mountain). We attempted to reproduce these sampling conditions so that comparison between spread rate measurement methods included comparisons of how different methods may accommodate adverse sampling conditions.

Four different shapes of the sampled area were tested. The first configuration that was tested (Fig. 2A) was the idealized situation of an invasion starting from the centre of an unconstrained study area, which is rarely encountered with empirical data. The second configuration simulated the establishment of a surveillance sampling network at the edge of an invaded area (Fig. 2B). This corresponded to a more realistic situation encountered in several studies. The third configuration simulated a situation in which the invasion spreads through a corridor (Fig. 2C), or a situation where the surveillance network is distributed along a transect. The fourth configuration corresponded to a sampling network with a very irregular shape (Fig. 2D), either constrained by the geographical distribution of suitable habitat, or by the distribution of sampling points in the hypothetical survey program. All survey areas had an equivalent surface area of 100 000 km^2. In addition, we evaluated the impact of different distributions (regular, aggregated and random).

Species distribution atlases are typically organized in the form of number of observation records per square, and regularly spaced data are therefore frequent. Aggregated and randomly distributed sampling is also often found in empirical data, either resulting from some design, or caused by field data collection constrains, or clumped distribution of observers. The random distribution of sampling points was generated by randomply selecting n points from a 1000×1000 grid. For the aggregated sampling, we generated a grid of 1000 by 1000 points assigned spatially correlated identification values using unconditional Gaussian simulation (range = 20 km, sill = 0.5, nugget = 0.5) and then selected the n points with the highest values.

Finally, we also varied the densities of sampling points to test the resistance of different method to reductions of the sampling effort (low, medium and high and high densities were tested, with respectively 200, 1000 and 5000 points/ 100 000 km^2, Supplementary material Fig. S3).

Due to practical computational constrains, it was not possible to test spread rate measurements based on sampling using all combinations of, spatial configuration (4 different configurations), point distribution (regular, random and aggregated) and density (3 different densities). Furthermore, some combinations were considered less important. For example, reducing the sampling density may have a significant impact on the accuracy of spread measurement,

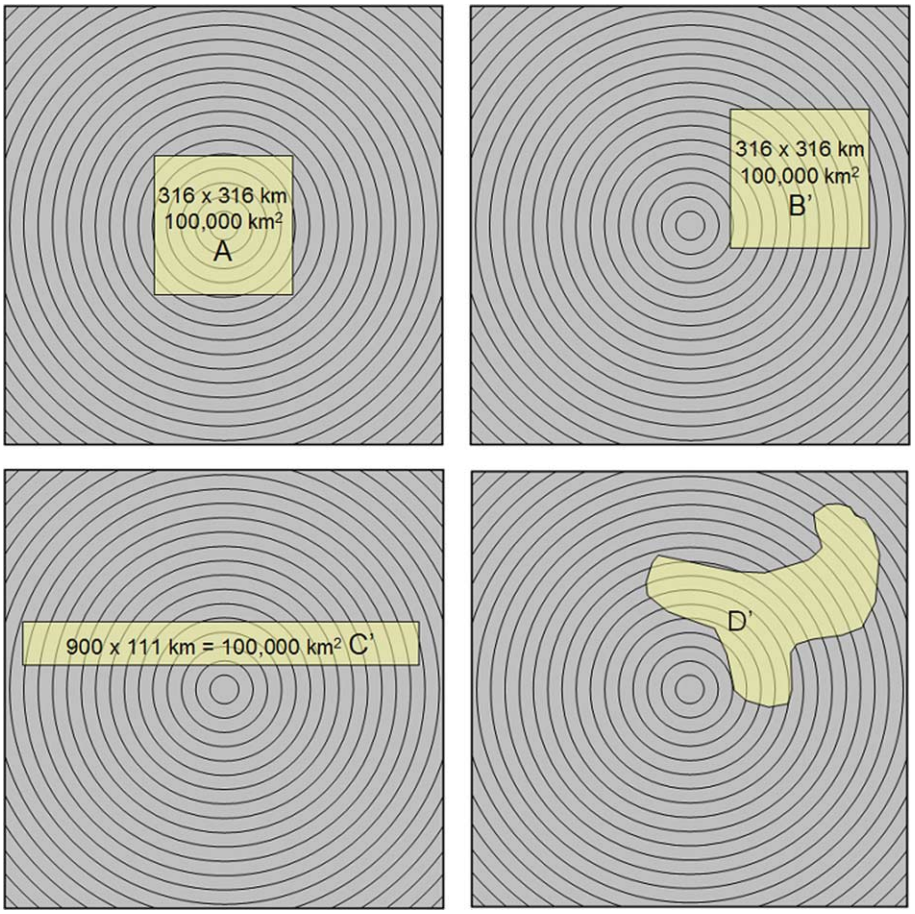

Figure 2. Spatial configuration of the sampled area in relation to the simulated invasion (represented here by the concentric circles): square – center invasion (A), square – lateral invasion (B), corridor invaded by one of its side (C), irregular shape invaded by its side (D).

which is interesting to establish and quantify, but this relationship is not likely to be affected by different spatial configurations. Therefore, we focused on a series of comparison where other factors were kept constant, and for each evaluation, we compared the results produced by the three spread rate measurement methods, as detailed in Table 1. For each set of parameters and stochasticity level (e.g. Fig. 1) invasions were simulated 25 times, and for each simulated invasion chronology, the distribution of sampling points was simulated 25 times. A total of 625 spread rate measurements were thus estimated per spread rate measurement method, and then used to quantify the accuracy and variability of their outcomes.

Spread rate measurement methods

Distance Regression. This method is perhaps the simplest to implement as it does not require determining an invaded area at different time steps, but rather, is based on the time (e.g. year) a species was first detected in a series of sampling points (Liebhold et al. 1992). The distance of these sampling points to the assumed point of origin of the invasion (generally the first point where a species was found) is then regressed as a function of the time of first detection, and the slope of the regression line provides the estimated rate of spread. This method is particularly suitable for observation data collected, for example, in a series of municipalities. It is generally assumed that the detection threshold is constant, i.e. that on average, the time when a species is first detected corresponds to the time when populations have exceeded some population threshold, reflecting local establishment. In our simulations, we extracted the population level in all sampling points at each time step, selected those that were newly infested (points where the population level for the first time exceeded a population threshold of 250), and computed their distance to the point of origin of the invasion in the study area. The data from all time steps were then pooled together in a single scatterplot of time vs distance, and the slope was estimated by linear regression (Fig. 3).

Square Root Area Regression. This method is generally used to derive the spread rate from historical maps of invaded areas in different time steps (Skellam 1951). The square root of the invaded area is plotted and regressed as a function of time to estimate the rate (regression slope) of spread. However, in many cases, surveillance data are point-based (trapping sites, observation point, etc.) which must be used to estimate invaded areas in each time step. We accomplished the estimation of infested area using two methods. In the first method, the infested area was estimated from the number of sample points where the local population level exceeded a threshold of 250 multiplied by the overall number of km^2 in the studied area/sampling

Table 1. Comparison of spread rate measurement conditions, and factors kept constant in the comparison of the performance of the different methods.

Varying conditions	Other conditions kept constant
Sampling points distribution (aggregated, random, regular)	Medium sampling density, centered square configuration.
Sampling density (low, medium, high)	Random sampling distribution, centered square configuration.
Invasion configuration (irregular shape, long rectangle, square – border, square – center)	Random sampling distribution, medium sampling density.

points. In the second method, the infested area was estimated via spatial interpolation from the sample point locations, a method employed frequently in empirical studies. Several methods are available to interpolate point-based observations (e.g. moving-window averaging, inverse-distance weighted, Kriging). However, in order to facilitate the automated interpolation, we chose inverse distance weighted method (IDW) so that there was no need to estimate and model the spatial autocorrelation structure for each simulation. In our simulations, we extracted the population value at each survey point in each time step, interpolated these values over all cells using IDW, defined invaded area as all cells where the interpolated population was >250, and calculated the area invaded at each time step. For both versions of this method, the square root of the invaded area was then plotted and regressed as a function of time to obtain the radial rate of spread (Fig. 3).

Boundary displacement. This method estimates the average displacement between pairs of consecutive invasion boundaries along axes radiating from the invasion origin, or perpendicular to the main invasion front (Sharov et al. 1997). The method also requires spatial delimitation of the invaded area, with a boundary separating the invaded vs. non-invaded area. In simulations, infested areas were estimated via interpolation of survey point data at each time step as described for the previous method. Then radial axes were superimposed on the boundaries corresponding to two consecutive distribution areas, and the distance between the intersections of each axis and each boundary was measured over all axes. The values from all pairs of consecutive boundaries along the study period were averaged to estimate the overall radial rate of spread (Fig. 4). We note that both the distance regression and boundary displacement methods require knowledge of the point of origin of the invasion.

Results

Under the most simple conditions (randomly located survey points, highest sample density, regular/circular invasion), all methods performed equally well (Fig. 4 bottom, 5000 survey points). Measures derived from the three methods were nearly identical and exhibited very little variation. However, sampling from simulations conducted using the stochastic model yielded slightly different results among the different sampling methods. The two methods utilizing interpolation (square root area regression no. 2 and boundary displacement) tended to overestimate the spread rate, but the square root area regression no. 1 method tended to underestimate it. Considering all four methods, only the distance regression method performed adequately at providing an unbiased measure of the spread rate from the stochastic model. However, all methods provide measures within 5% of the expected spread rate.

As would be expected, decreasing sampling density resulted in higher variability in spread rate estimates, but

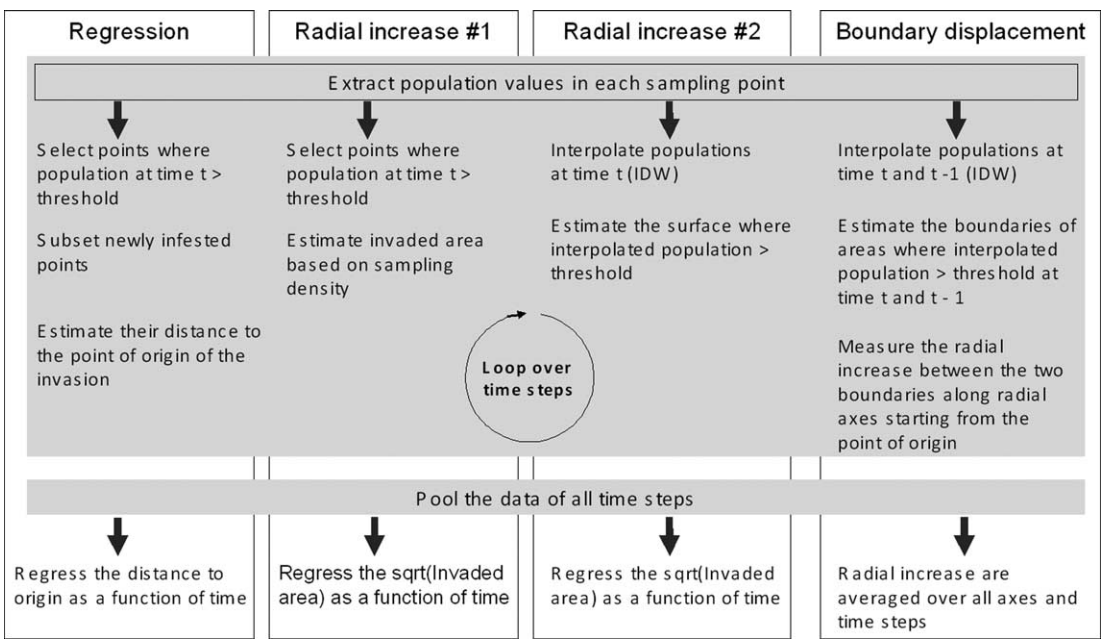

Figure 3. Steps involved in estimating the rate of spread according to the four methods tested in the study.

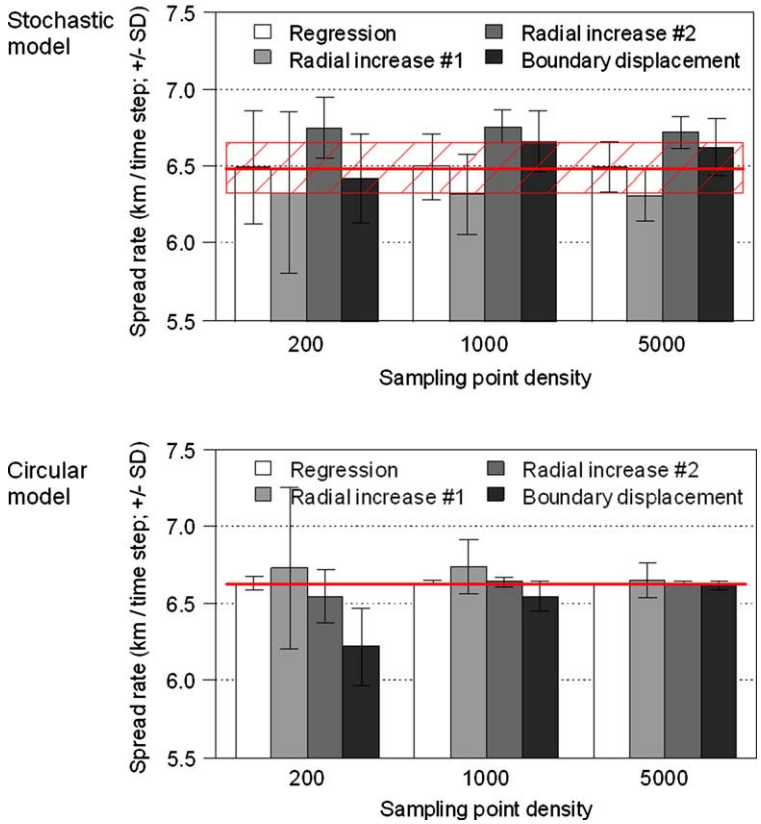

Figure 4. Spread rate measured for varying sampling point densities, based on the stochastic (top) and circular (bottom) spread models. These were obtained with a random distribution of points using the square – center invasion configuration. The hatched area corresponds to the distribution of spread rate estimates calculated from the exhaustive (all cells) data set.

all methods were not affected equally by sample size (Fig. 4). The distance regression method appears to the least adversely affected by small sample size, while the square root area regression no. 1 and boundary displacement methods were the most sensitive. For data generated using the circular spread model (the only model for which the "true" spread rate was precisely known), the boundary displacement method performed particularly poorly at low sample sizes.

The distance regression method was also the least sensitive to the distribution of survey points (Fig. 5). The mean estimate from the distance regression method was always very close to the expected spread rate under aggregative, regular and random distributions of survey points and for data generated from both the stochastic and the circular models. However, in the case of the circular model, all four methods produced little bias in estimates. In the case of the stochastic model, sampling via regular and random distributions produced similar results.

Aggregative distribution of sampling points resulted in a particularly high level of variability of estimates using the square root area no. 1 method applied to data from both the circular and stochastic models.

Irregularity in the shape of the invaded area had a large negative effect on the accuracy of rate estimates from all methods except distance regression and boundary displacement (Fig. 6). In the simplest configuration of spread from the center, all methods performed equally well. But when spread was measured over more irregular areas, both of the square root area methods provided gross under-estimates of the true rate of spread. The boundary displacement method was least affected by irregularities in the shape of the area, though the distance regression method was nearly as good in many cases. One should note that this may result from an incorrect assignment of the point of origin that is used by the distance regression method. In the algorithm, the point of origin is the first point in the study area where the population was higher than the threshold. Using the stochastic invasion model, there may have been considerable error in locating this point of origin, hence affecting the regression method, while in the deterministic circular model, this would not be a problem.

Discussion

Predictions of future spread are often in high demand in order to make management decisions in anticipation of a species' arrival. For example silvicultural decisions for managing forest stands are often predicated on the number of years remaining before a specific forest exotic insect or disease is expected to arrive (Waring and O'Hara 2005). Predictions of future spread may be based either on simulations using a mechanistic model or extrapolations from past patterns of spread (Hastings et al. 2005, Liebhold and Tobin 2008). While process-based models may be the only choice for some organisms, in many systems, there may exist historical data from which a rate of spread can be estimated and applied to predict future invasion.

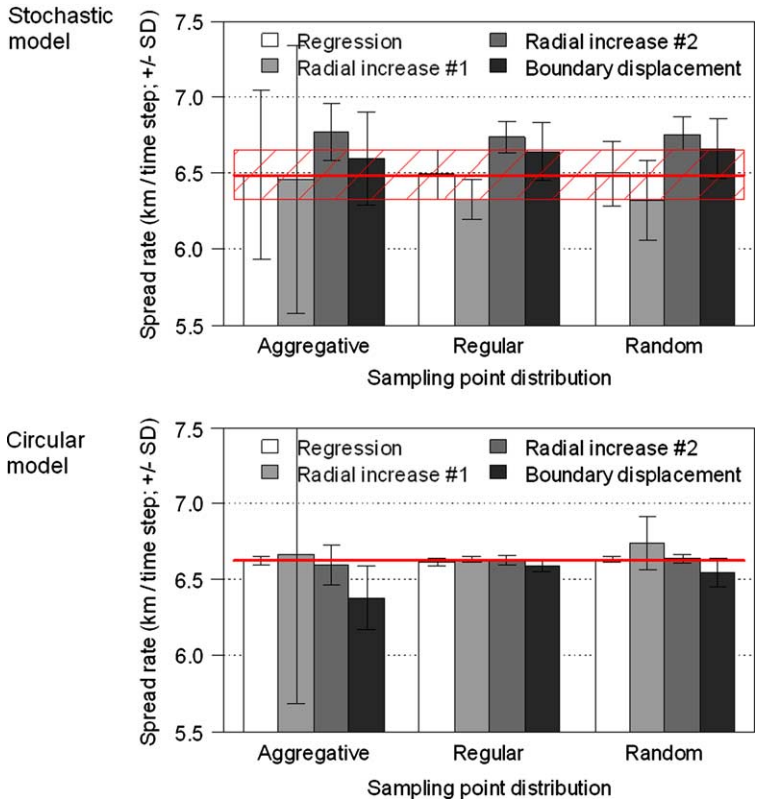

Figure 5. Spread rate measured from varying sampling point distributions, based on the stochastic (top) and circular (bottom) spread models. These were obtained with an intermediate density of sampling points (1000) using the square – center invasion configuration.

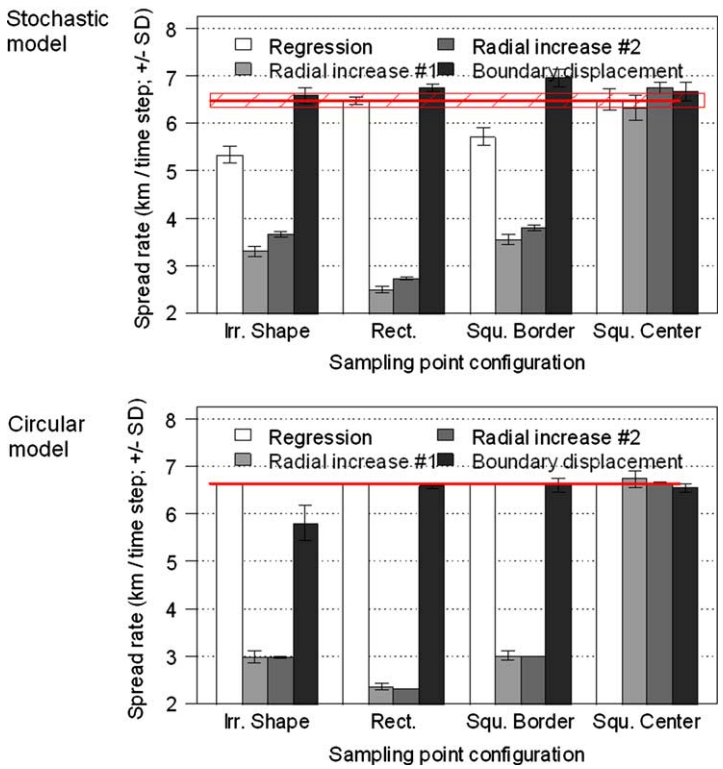

Figure 6. Spread rates measured for varying invasion configurations, based on the stochastic (top) and circular (bottom) spread models. These were obtained with an intermediate density of sampling points (1000) distributed randomly.

There are several issues that may complicate estimation of spread rates. First, there may be many situations where spread rates are not constant through time. For example, Weber (1998) found that in about half of plant invasions analyzed, the rate of spread increased through time, while in the other half spread was approximately constant. Simple reaction-diffusion models are known to asymptotically produce constant radial rates of spread, while more complex mechanisms are known to sometimes produce nonstationary rates of spread (Shigesada and Kawasaki 1997, Hastings et al. 2005). Despite the existence of more complex patterns and mechanisms of spread, the simple reaction-diffusion model provides a remarkably realistic representation of spread in many invading species (Andow et al. 1990, Shigesada and Kawasaki 1997). In addition, one may partly control for temporal variation in spread rates by constraining the measurement of spread rates to a limited number of years. Another complicating factor in the spread of invading species may be the presence of spatial anisotropy. As a result of landscape heterogeneity that affects either population growth or dispersal, the spread of some species may be unequal among different directions from the point of origin of the invasion (Andow et al. 1993, Evans and Gregoire 2007, Morin et al. 2009).

As described here, there are several approaches to estimating spread rates and there can be several complications in this process. As is the case in many applied ecological problems, the ability to make precise estimates may be limited both by the quantity and quality of data. In virtually all estimation procedures, more data yields more precise estimates, and this is certainly the case with estimation of spread rates from historical survey data (Fig. 4). Among the four methods evaluated here, the distance regression method was preferable when sample sizes were low; at the lowest sample size tested, it provided the most accurate estimates. The two methods that utilized spatial interpolation, the square root area regression no. 2 method and the boundary displacement method, both were adversely affected by the availability of a small sample size. The strong negative effect of small sample sizes on the accuracy of spatial interpolation is well known (Isaacs and Srivastava 1989) and it is therefore understandable that these methods would not perform well with small sample sizes.

Overall, the distance regression method appeared to provide the most reliable estimates of spread rate. As described above, it provided the most accurate and precise measure even with low sample sizes, both for the circular model and the stochastic model. Similarly, it's accuracy was not adversely affected by deviation from a regular or random distribution of sampling points. Even when the sampling was constrained to an irregular area, the distance regression method performed well, though the boundary displacement method performed as well, or in some cases better. The two methods based on area both performed poorly when the sampling configuration deviated from simple spread from the center. As previously noted by Shigesada and Kawasaki (1997), any irregularity in the habitat (e.g. the intrusion of a large lake or ocean) means that the change in the square root area will not be proportional to the increase in the radius of the same area. While controlling for those changes may be relatively easy if the invaded area has a non-circular but regular shape (e.g. a cone), it may become difficult to compensate for more complex shapes.

A further advantage of both the distance regression and boundary displacement methods is that they can be more easily adjusted to test for spatial anisotropies by measuring spread rate using different spatial subsets with different orientations to the origin. In conditions where sampling density is sufficient, the boundary displacement method has an additional advantage of allowing estimation of the instantaneous spread rate in different regions of the newly invaded area. As pointed out by Tobin et al. (2007b), the method is desirable when information on the temporal and spatial variation in spread rate is needed. For example, Tobin et al. (2007a) used the boundary displacement method to analyze spatial and temporal variation in the radial rate of spread of the gypsy moth in N. America and to relate this to spatiotemporal variations in Allee effects. Another example is the study Sharov et al. (1999) where the boundary displacement method was used to examine the correlation of the instantaneous spread rate of gypsy moth in N. America with climate and other landscape characteristics of the newly invaded area. In addition, the boundary displacement method does not require a long time-series of observation, and could be used to implement a quick measure of spread rate for an invasion that is in the first stages of invading a new region. For example, *Cameraria ohridella* first invaded Germany in the years 1996–1999 before entering France from the east (Gilbert et al. 2004, 2005). France needed to quickly measure the rate of spread of the species in its landscape and use these estimates to design a dense network of observations and pheromone traps along the expanding population front of *C. imicola* in order to monitor spread over three consecutive years, (Augustin et al. 2004).

Evans and Gregoire (2007) argued that the application of ordinary least squares regression in the distance regression method leads to an under-estimation of the true spread rate. Their argument was that because it is impossible to have negative distances, the scatter of points is constricted and they suggested applying Tobit regression in order to correct for this problem. Evans and Gregoire (2007) did not compare the performance of ordinary least squares regression with Tobit regression to confirm their theory. However, our empirical finding here that ordinary least squares regression provided an unbiased estimate of the true spread rate does not support their conclusion.

Thus, we may conclude that the distance regression method is generally the most robust and reliable method for estimating spread rates, especially when sample sizes are constrained, and the invaded area has an irregular shape. Nevertheless, there still may be some situations in which the boundary displacement method may be desirable, as long as large sample sizes are available. One such situation is when the origin of the invasion is unknown. In contrast to the boundary displacement method, the distance regression method requires the identification of the origin of the invasion. An exploratory analysis carried out using the default configuration (square – center invasion configuration, medium density, random sampling) and where error was introduced in the location of the invasion origin demonstrates a strong dependency of the distance regression method to a correct spatial allocation of the origin (Fig. 7).

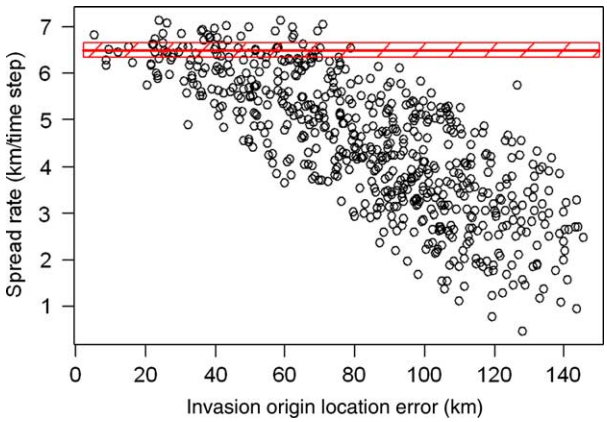

Figure 7. Spread rates measured using misallocated invasion origin. These were obtained with an intermediate density of sampling points (1000) distributed randomly using the square – center invasion configuration. The invasion origin was misallocated by simulating it's distribution following a uniform distribution of x and y coordinates.

Further work would be required to explore in more details the causes underlying the bias in the distance regression method estimates by examining the dependency on the selection of the correct starting point in varying conditions. Furthermore, the method by which we extract a sub-domain from an invasion taking place over a larger domain implicitly assumes that populations keep developing independently outside the sub-domain. Whilst this may be realistic in some circumstances (e.g. sampling was only in a sub-region), it may not be in others (e.g. spread is constrained to the sub-domain surrounded by non-suitable habitat) in which the regression method may appear more suitable even in conditions where the invaded area has an irregular shape. Finally, we could also use the same framework to test different adaptive sampling strategies aiming at optimizing the cost-benefit of sampling (e.g. densities varying as a function of distance to the invasion front).

Acknowledgements – The authors wish to thank the USDA Forest Service Northern Research Station for financing this work (Agreement 05-JV-11242328-054), and the Fonds National de la Recherche Scientifique for funding M. Gilbert.

References

Andow, D. A. et al. 1990. Spread of invading organisms. – Landscape Ecol. 4: 177–188.

Andow, D. A. et al. 1993. Spread of invading organisms: patterns of spread. – In: Kim, K. C. and McPheron, B. A. (eds), Evolution of insect pests: the pattern of variations. Wiley, pp. 219–242.

Augustin, S. et al. 2004. Monitoring the regional spread of the invasive leafminer *Cameraria ohridella* Deschka & Dimic (Lepidoptera: Gracillariidae) by damage assessment and pheromone trapping. – Environ. Entomol. 33: 1584–1592.

Edwards, A. M. et al. 2007. Revisiting Levy flight search patterns of wandering albatrosses, bumblebees and deer. – Nature 449: 1044–1048.

Evans, A. M. and Gregoire, T. G. 2007. A geographically variable model of HWA spread. – Biol. Invasions 9: 369–382.

Fick, A. 1855. Über Diffusion. – Poggendorf's Annalen Bd. 94: 59–81.

Fisher, R. A. 1937. The wave of advance of advantageous genes. – Ann. Eugenics 7: 355–369.

Gilbert, M. et al. 2004. Long-distance dispersal and human population density allow the prediction of invasive patterns in the horse-chestnut leafminer *Cameraria ohridella*. – J. Anim. Ecol. 73: 459–468.

Gilbert, M. et al. 2005. Forecasting *Cameraria ohridella* invasion dynamics in recently invaded countries: from validation to prediction. – J. Appl. Ecol. 42: 805–813.

Hastings, A. et al. 2005. The spatial spread of invasions: new developments in theory and evidence. – Ecol. Lett. 8: 91–101.

Hudson, P. J. and Cattadori, I. M. 1999. The Moran effect: a cause of population synchrony. – Trends Ecol. Evol. 14: 1–2.

Isaacs, E. H. and Srivastava, R. M. 1989. Applied geostatistics. – Oxford Univ. Press.

Kot, M. et al. 1996. Dispersal data and the spread of invading organisms. – Ecology 77: 2027–2042.

Liebhold, A. M. and Tobin, P. C. 2008. Population ecology of insect invasions and their management. – Annu. Rev. Entomol. 53: 387–408.

Liebhold, A. M. et al. 1992. Gypsy moth invasion in North America: a quantitative analysis. – J. Biogeogr. 19: 513–520.

Morin, R. S. et al. 2009. Anisotropic spread of hemlock woolly adelgid in the eastern United States. – Biol. Invasions, in press.

Pebesma, E. J. 2004. Multivariable geostatistics in S: the gstat package. – Comput. Geosci. 30: 683–691.

Ricker, W. E. 1954. Stock and recruitment. – J. Fish. Res. Board Can. 11: 559–623.

Sakai, A. K. et al. 2001. The population biology of invasive species. – Annu. Rev. Ecol. Syst. 32: 305–332.

Sharov, A. A. et al. 1997. Methods for monitoring the spread of gypsy moth (Lepidotperta: Lymantriidae) populations in the Appalachian Mountains. – J. Econ. Entomol. 90: 1259–1266.

Sharov, A. A. et al. 1999. What effects the rate of gypsy moth (Lepidoptera: Lymantirriae) spread: winter temperature or forest susceptibility? – Agric. For. Entomol. 1: 37–45.

Sharov, A. A. et al. 2002a. Evaluation of preventive treatments in low-density gypsy moth populations. – J. Econ. Entomol. 95: 1205–1215.

Sharov, A. A. et al. 2002b. "Slow the spread": a national program to contain the gypsy moth. – J. For. 100: 30–36.

Shigesada, N. and Kawasaki, K. 1997. Biological invasions: theory and practices. – Oxford Univ. Press.

Skellam, J. G. 1951. Random dispersal in theoretical populations. – Biometrika 38: 196–218.

Tobin, P. C. et al. 2004. Management of the gypsy moth through a decision algorithm under the STS Project. – Am. Entomol. 50: 200–209.

Tobin, P. C. et al. 2007a. Invasion speed is affected by geographic variation in the strength of Allee effects. – Ecol. Lett. 10: 36–43.

Tobin, P. C. et al. 2007b. Comparison of methods for estimating the spread of a non-indigenous species. – J. Biogeogr. 34: 305–312.

Waring, K. M. and O'Hara, K. L. 2005. Silvicultural strategies in forest ecosystems affected by introduced pests. – For. Ecol. Manage. 209: 27–41.

Weber, E. 1998. The dynamics of plant invasions: a case study of three exotic goldenrod species (*Solidago* L.) in Europe. – J. Biogeogr. 25: 147–154.

Download the Supplementary material as file E6018 from <www.oikos.ekol.lu.se/appendix>.