Objective functions for comparing simulations with insect trap catch data

Guichard, S.[1,2], A. Leriche[2], D.J. Kriticos[2], J.M. Kean[3] and S.P. Worner[4]

[1] SCION, PB 3020 Rotorua, New Zealand
[2] CSIRO Entomology and the Cooperative Research Centre for National Plant Biosecurity, GPO Box 1700, Canberra, Australia 2601
[3] AgResearch, Private Bag 4749, Christchurch, New Zealand
[4] Lincoln University, Bio-Protection and Ecology Division, Canterbury, New Zealand
Email:sylvain.guichard@csiro.au

Abstract: Targeted surveillance of high risk invasion sites using insect traps is becoming an important tool in border biosecurity, aiding in early detection and subsequent monitoring of eradication attempts. The mark-release-recapture technique is widely used to study the dispersal of insects, and to generate unbiased estimates of population density. It may also be used in the biosecurity context to quantify the efficacy of surveillance and eradication monitoring systems.

Marked painted apple moths were released at three different locations in Auckland, New Zealand over six weeks during a recent eradication campaign. The results of the mark-release-recapture experiment were used to parameterise a process-based mechanistic dispersal model in order to understand the moth dispersal pattern in relation to wind patterns, and to provide biosecurity agencies with an ability to predict moth dispersal patterns. A genetic algorithm was used to fit some model parameters. Different objective functions were tested: 1) Cohen's Kappa test, 2) the sum of squared difference on trap catches, 3) the sum of squared difference weighted by distance from the release site, 4) the sum of squared difference weighted on distance between best-fit paired data. The genetic algorithm proved to be a powerful fitting method, but the model results were highly dependant on the objective function used.

Objective functions for fitting spatial data need to characterise spatial patterns as well as density (*ie*. recapture rate). For fitting stochastic models to datasets derived from stochastic spatial processes, objective functions need to accommodate the fact that a perfect fit is practically impossible, even if the models are the same.

Applied on mark-release-recapture data, the Cohen's Kappa test and the sum of squared difference on trap catches captured respectively the distance component of the spatial pattern and the density component adequately but failed to capture both requirements whereas the sum of squared difference weighted by distance from the release site did. However, in order to integrate the stochastic error generated by the model underlying stochastic process, only the sum of squared difference weighted on distance between best-fit paired data was adequate.

The relevance of each of the fitting methods is detailed, and their respective strengths and weaknesses are discussed in relation to their ability to capture the spatial patterns of insect recaptures.

Keywords: *biosecurity, dispersal, individual based model, insect, goodness of fit*

1. INTRODUCTION

Biosecurity aims to protect economic and biodiversity values of a region from unwanted exotic organisms. One important tool for insect biosecurity is a targeted surveillance system comprising a network of attractive traps that are monitored regularly for early detection of the target species. This approach is a mainstay of gypsy moth and bark beetle biosecurity systems worldwide, and was particularly useful in the combating a recent major incursion of painted apple moth in New Zealand (Suckling et al., 2007). In order to better respond to the detection of an unwanted species in a surveillance trap, a simulation model was developed to reproduce the likely flight path that led the insect into the trap, and in so doing, indicate where to search for the proximal source of the incursion. The simulation model uses an individual-based dispersal modelling approach, and is based on the well documented behaviour of insects in response to pheromone attractants. The model draws heavily on pheromone anemotaxis theory, which describes the optimal strategy used by male insects to track the pheromone plume emitted either by a female of the species or by a pheromone dispenser.

The dispersal model was partly parameterised using experimental data obtained from a mark-release-recapture experiment using the painted apple moth, *Teia anartoides*, in Auckland, New Zealand, in 2003 (Suckling et al., 2005). A genetic algorithm was used to fit inferentially some model parameters for which there was little knowledge available to inform their values. Genetic algorithms are methods inspired by natural selection processes in the evolution of biological organisms (Beasley et al., 1993). A set of parameters represents a conceptual "genotype", within which, each parameter is a "gene". Using an appropriate fitness function, the algorithm scores the genotypes from a population, initially randomly generated, and selects the best genotype in order to "breed" a new population using cross-over and mutation processes. The selection of the best genotypes is crucial for the convergence of the population to a "perfect match" and is highly dependant on the fitness function. The rank of the scores for different genotypes reflects the goodness of fit of the simulation.

In this paper, we present a comparison of different objective functions for comparing simulated insect trap catches with observed recaptures in the mark-release-recapture experiment.

2. MATERIAL AND METHODS

2.1. Observed recaptures

We chose a typical example of the recapture patterns observed from painted apple moth releases in Auckland, New Zealand in 2003. This example was from a release made on 6 March 2003 at Waikumete Cemetery (Figure 1). The spatial recapture pattern was characterised by three important features. Firstly, distance: most insects were recaptured close by, with recapture rates declining with distance from the release site. Secondly, direction: as a consequence of the particular meteorological patterns of the time, recaptures tended to occur to the north of the release site. Thirdly, stochastic error: since both insects and traps are discrete units in time and space, the observed outcomes from both simulation and experiment represent single realisations of the underlying stochastic processes. Importantly, this means that the simulation results will practically never match exactly the observed field results, even when the same mechanisms and parameters are responsible for both. The performance of any parameter fitting process for dispersal models such as this one depends on the ability of its objective function to assess adequately the distance, direction, and stochastic error of simulation results in relation to the observed recapture data.
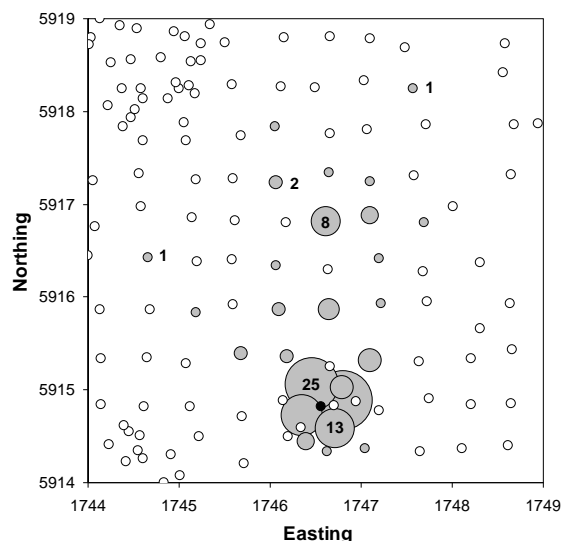
Figure 1. Recapture pattern observed mark-release-recapture of painted apple moth at Waikumete Cemetery, Auckland, on 6 March 2003. Black dot represents the release site, grey dots indicate recaptures, with circle size proportional to the number captured and open dots indicate traps with no recaptures. The NZTM coordinate system has been used here with Easting and Northing expressed in kilometres.

2.2. Objective functions

We evaluated the strengths and weaknesses of four different objective functions for adequately ranking the goodness of fit of different spatial patterns of insect recaptures:

- Cohen's Kappa statistic (Cohen, 1960). The score K evaluates the agreement between observed and simulated recaptures for each trap independently of the number of recaptures per trap (1). Trap catches are transformed to qualitative data and Pa is the proportion of agreement; Pc, the proportion of agreement expected by chance given the observed marginal totals. The higher the Kappa statistic, the better the agreement between observations and simulations is.

$$K = \frac{P_a - P_c}{1 - P_c} \qquad (1)$$

- Sum of squared differences (SSD) between observed and simulated recaptures for each trap (2). This quantifies the errors generated by differences in trap catches between observations and simulations. In trap i the number of observed recaptures is O_i; the number of simulated recaptures is S_i for the trap i and n is the total number of traps. The lower the sum of square difference, the greater the similarity between observations and simulations.

$$SSD = \sum_{i=1}^{n} \left[(O_i - S_i)^2 \right] \qquad (2)$$

- Distance-weighted sum of squared differences (DSSD). This is based on the sum of squared differences where each square difference is weighted against traps further from the release location (3). D_i is the distance in kilometres between each trap i and the release location.

$$DSSD = \sum_{i=1}^{n} \left[(O_i - S_i)^2 * D_i \right] \qquad (3)$$

- Distance-weighted sum of squared differences on paired results (DSSDP). This calculates a sum of squared differences between observed and simulated recaptures where each simulated recapture is paired with the closest observed recapture and weighted according to equation (3) except that Di is now the distance between the paired traps. If the distance between paired traps is greater than the distance between the trap of the observed recapture and the release location, then the traps are not paired and the regular DSSD is applied. Furthermore, when the total numbers of simulated and observed recaptures are different, the DSSD is applied to the data remaining after pairing the best fits. Hence, each observed and simulated data point contributes only once to the sum of square differences. With a regular trapping grid, if the observed recaptures occur simultaneously in different traps separated by the same distance to one simulated recapture, the simulated recapture is randomly paired with one of the closest observed recaptures. Figure 2 and Table 1 illustrate the calculation method. Compared to the SSD and DSSD, a recapture occurring in a trap close to an observation is penalized less: a near-miss in spatial terms is treated as almost a perfect match. This acknowledges that dispersal behaviour is a stochastic process, and that both simulated and observed recaptures are single realizations of similar stochastic processes. In this context, a recapture occurring in the closest trap could be considered as an approximation of a perfect match. This new objective function is inspired by the geostatistical principle of spatial inter-relationship between values of the same variable (for example, Rossi et al. 1992) but applied to two variables with known underlying stochastic processes.
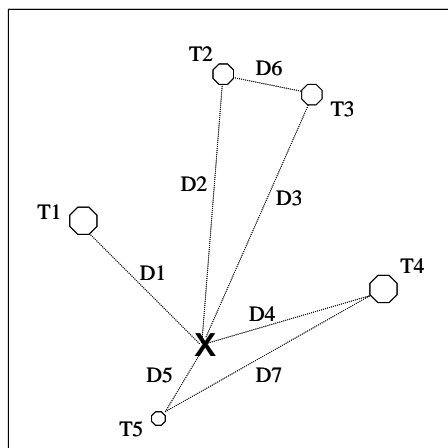
Figure 2. Example map of trap locations to illustrate the DSSDP calculation. The cross denotes the release location. Traps Tn are separated by distances Dn.

Table **1**. Details of the calculation of each components of the sum of square differences on paired trap catches weighted. In this example, trap T2 in which 4 recaptures were simulated and trap T3 in which 3 recaptures were observed, are paired. The calculation of the goodness of fit for T2 and T3 is: 3 simulated recaptures were paired with 3 observed recaptures and we applied Equation 3 with the distance D6; for the remaining recaptures in the trap T3 Equation 3 was used with the distance D3 as this observed recapture could not be paired with any simulated recapture. See Figure 2 for trap location and distance between traps.

Trap	Observed recaptures	Simulated recaptures	Distances between traps	Results
T1	4	4	D1	$(4-4)^2 * D1$
T2	0	3	D2 > D6	$(0-3)^2 * D6$
T3	4	0	D3 > D6	$+ (1-0)^2 * D3$
T4	4	0	D4 < D7	$(4-0)^2 * D4$
T5	0	1	D5 < D7	$(0-1)^2 * D5$

2.3. Trap catch scenarios

The different fitness functions were evaluated against a set of trap catch scenarios (Figure 3) designed to represent different aspects of the distance, direction, and stochastic error of simulation results compared to the observed recapture data illustrated by Figure 1. These scenarios were, from least to most accurate:

- *Scenario 1 (null case)*: No moths were recaptured. This scenario represents one of the lower limits of each fitness function; although it doesn't represent the worst matching case which is when recaptures occurred only in traps where no recapture was observed.
- *Scenario 2 (inverted dispersal pattern)*: The overall direction of recaptures is the same as observed, but the count of recaptures in individual traps is inverted, so that simulated recaptures tend to be distant from the release location. The primary source of error here is distance from the release point.
- *Scenarios 3 and 4 (truncated dispersal)*: Close to the release location, simulated results produce the observed pattern but no recaptures are observed far from the release location. In scenario 3 the errors begin closer to the release site than in scenario 4, so that the distance error is greater in the former scenario.
- *Scenario 5 (near-miss)*: This scenario is, for all practical purposes a desirable fit. The overall spatial pattern of abundance is similar between observation and simulation except for the traps far from the release location where simulated recaptures occurred in traps close to, but not coinciding with the observed recaptures. Although the errors appear greater here than in scenario 4, this is misleading since the individual recapture data are actually spatially closer to the observed recaptures. In scenario 5 there is very little direction or distance error but mostly just variation in the exact trap that catches an insect.
- *Scenario 6 (perfect match)*: All the simulated recaptures occurred where recaptures were observed. This scenario represents the upper limit of each fitness function.
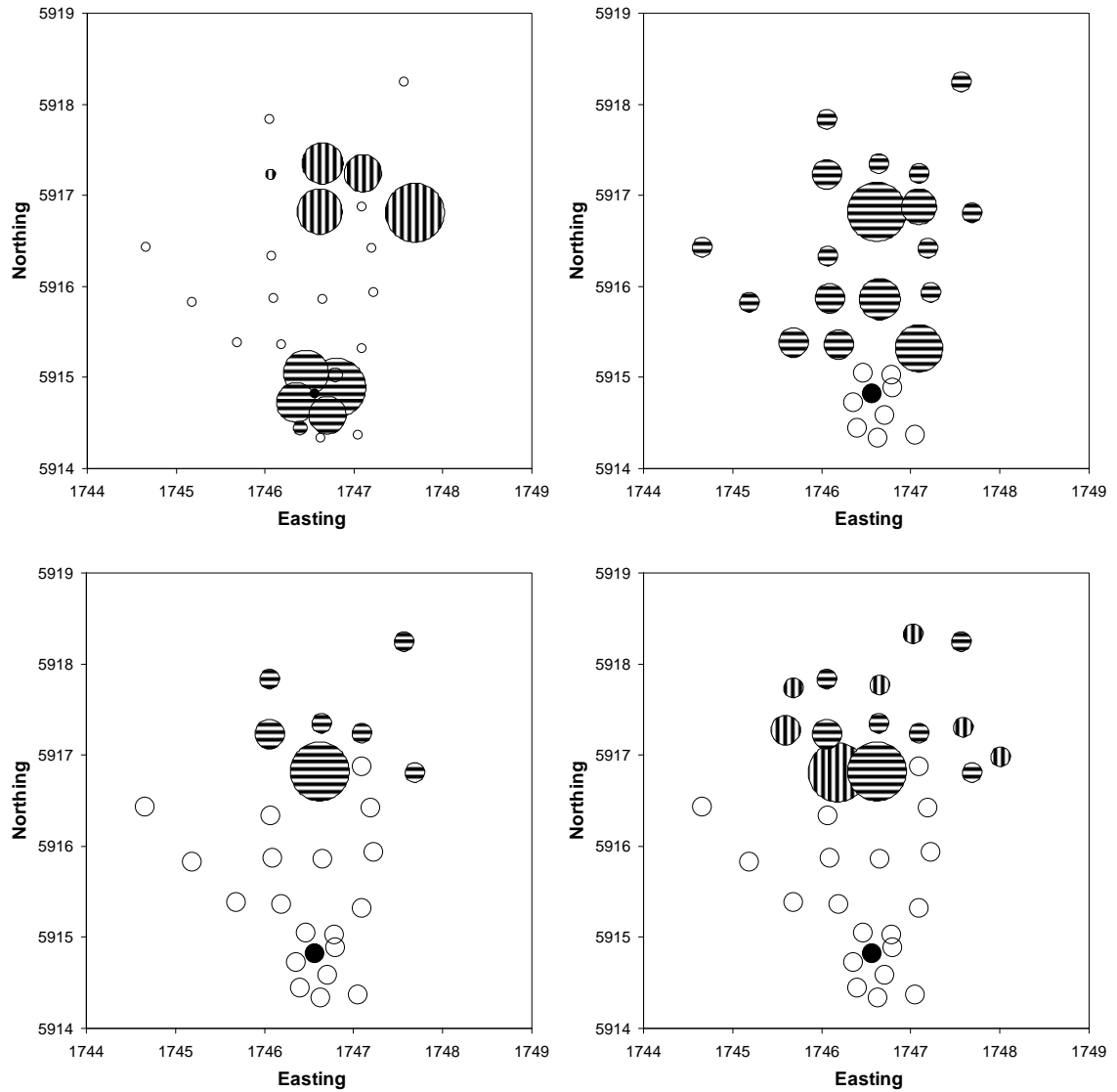
Figure 3. Illustration of the magnitude of differences (size of hatched circle) between observed and simulated trap catches following scenarios 2 (top left), 3 (top right), 4 (bottom left) and 5 (bottom right). The horizontal cross-hatching indicates omission errors (simulated recaptures < observed) and vertical cross-hatching indicates commission errors (simulated recaptures > observed). Open circles indicate that simulated recaptures = observed.

2.4.   Rescaling of the objective functions

Scenarios 1 and 6, the null and perfect match cases respectively, were used to calibrate each of the fitness functions. We expressed the scores as a percentage between these two limits. Cohen's Kappa was multiplied by 100 as scenarios 1 and 6 gave rise to scores of 0 and 1 respectively. For the other fitness functions the raw scores (S_{raw}) were 0 for the perfect match scenario and they increased in value with decreasing similarity between the observed and simulated datasets. These scores were transformed using equation (4) in order to adjust the score of null case scenario S_{null} to 0 and the score of the perfect match scenario to 100.

$$S = 100 - \left(\frac{S_{raw} * 100}{S_{null}} \right)$$

(4)

where S is the score after the rescaling, and $S < 0$ ranks the simulation worse than the null case. As the fitness function was designed to be used by a fitting algorithm, the critical aspect of each of the fitness

scores was not its actual value but rather its rank in comparison to other simulations within the "population" of simulations.

3. RESULTS

The results (Table 2) showed that because of its qualitative nature, Cohen's Kappa was unable to discriminate against scenario 2, which is correct in terms of direction but maximises error in distance. In contrast, the fitness functions based on sums of squared differences all ranked scenario 2 below the null case. The SSD detected the weakness of scenario 2 but scored all of the remaining scenarios very close to perfect, and provided little discrimination between scenarios 3 to 5. The SSD is highly influenced by the traps with the greatest recapture rates, and its performance was limited by the fact that recaptures were not uniformly distributed in space (Figure 1). Its inability to discriminate clearly between scenarios 3, 4, and 5 reflected the fact that the majority of the recaptures were found close to the release point. The DSSD correctly penalized scenario 2 and seemed to account adequately for the difference between scenarios 3 and 4. By weighting with distance from the release location, DSSD effectively accounted for the rapid decline in observed recaptures with distance from the release location. However, it also ranked scenario 5 lower than scenario 4, failing to capture the stochastic spatial error between simulation results and observed recapture data.

Table 2. Scores of each fitness function evaluated on a set of scenarios after rescaling of their score in order to set one lower limit, the null case scenario, to 0 and the upper limit, the perfect match scenario, to 100.

	Cohen's Kappa (K)	Sum of square difference on trap catches (SSD)	Distance-weighted sum of squared differences (DSSD)	Distance-weighted sum of squared differences on paired results (DSSDP)
Null case	0	0	0	0
Scenario 2 (inverted distance)	100	-39	-419	-419
Scenario 3 (truncated dispersal)	46	93	68	68
Scenario 4 (mildly truncated dispersal)	83	96	78	78
Scenario 5 (near-miss pattern)	71	93	57	95
Perfect match	100	100	100	100

The DSSDP gave the same results as DSSD for scenarios 2 to 4, since the functions are by definition equal when simulated recaptures all occurred in the same traps as observed recaptures. However, DSSDP was also able to recognise stochastic error and accordingly ranked scenario 5 highly. DSSDP was clearly the best of the functions tested for fitting the model for insect dispersal against the observed mark-release-recapture data in terms of direction and distance, but also recognising that observation and simulation will necessarily differ even when the dispersal functions are the same or similar.

4. CONCLUSIONS AND RECOMMENDATIONS

The choice of a fitness function is critical in the process of fitting model parameters. For modelling biosecurity trap recaptures, the direction and distance of recaptures relative to the release location are both important, and the fitness function had to capture both aspects effectively. Cohen's Kappa captured the qualitative pattern but failed to discriminate based on distance. In contrast, SSD only considers the difference in distance from release to recapture sites, but takes no account of the direction of dispersal.

The DSSD performed well when simulated and observed recaptures all occurred within the same subset of available traps, as would be the case if presence-only data were available, if only a subset of traps were effective, or if limited traps were used so that all traps caught insects. However, the DSSDP method, based on sum of squared differences for paired data and weighted by the distance between traps or between traps and the release location, proved to be the best fitness function for our purpose. It correctly ranked the test scenarios, successfully accounting for distance and direction from the release site, as well as recognising the inherent stochastic nature of insect dispersal. More generally, simulation modelling of any stochastic

spatial process could benefit from the use of objective functions that account for the inherent stochasticity in the process being modelled.

ACKNOWLEDGMENTS

REFERENCES

Beasley, D., Bull, D. R., and R. R. Martin (1993), An overview of genetic algorithms: part 1, fundamentals. *University Computing*, 15(2), 58-69.

Cohen, J. (1960), A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1), 37-46.

Rossi, R.E., Mulla, D.J., Journel, A.G. and E.H. Framz, (1992), Geostatistical tools for modeling and interpreting ecological spatial dependence. *Ecological Monographs*, 62(2), 277-314.

Suckling, D.M., Barrington, A.M., Chhagan, A., Stephens, A.E.A., Burnip, G.M., Charles, J.G. and S.L. Wee (2007), Eradication of the Australian painted apple moth *Teia anartoides* in New Zealand: Trapping, inherited sterility, and male competitiveness. *Area-wide control of insect pests: from research to field implementation*. (ed. by Vreysen, M.J.B., Robinson, A.S. and J. Hendrichs), 603-615. Springer, Dordrecht, Netherlands.

Suckling, D. M., Charles, J., Allan, D., Chaggan, A., Barrington, A., Burnip, G. M., and A. M. El-Sayed (2005), Performance of irradiated *Teia anartoides* (Lepidoptera: Lymantriidae) in urban Auckland, New Zealand. *Journal of Economic Entomology*, 98(5), 1531-1538.