

Modeled global invasive potential of Asian gypsy moths, *Lymantria dispar*

A. Townsend Peterson, Richard Williams* & Guojun Chen

Natural History Museum and Biodiversity Research Center, The University of Kansas, Lawrence, KS 66045, USA

Accepted: 6 May 2007

Key words: ecological niche modeling, GARP, Lepidoptera, Lymantriidae, natural history collections

Abstract

Asian populations of gypsy moths, *Lymantria dispar* (L.) (Lepidoptera: Lymantriidae), remain poorly characterized – indeed, they are not presently accorded any formal taxonomic status within the broader species. Their ecology is similarly largely uncharacterized in the literature, except by assumption that it will resemble that of European populations. We developed ecological niche models specific to Asian populations of the species, which can in turn be used to identify a potential geographic distributional area for the species. We demonstrated statistically significant predictivity of distributional patterns within the East Asian range of these populations; projecting the Asian ecological niche model to Europe, correspondence with European distributions was generally good, although some differences may exist; projecting the ecological niche model globally, we characterized a likely potential invasive distribution of this set of populations across the temperate zone of both Northern and Southern Hemisphere.

Introduction

Ecological variation and differentiation of invading species remain the subject of intense research; key questions include whether species interactions affect the invasion process, and possible effects of the invasion process on evolutionary dynamics of invading species (Lee, 2002). A first (and simpler) set of questions, however, revolves around ecological shifts in invading populations: in a simple ecological world, species would obey the same ecological ‘rules’ on invaded ranges as they do on native ranges, and the ecological characteristics of native ranges would be able to predict potential distributions on non-native areas. A sizeable body of evidence is accumulating to support the idea that invasive species often show such ‘simplicity’, that is, at least at coarse spatial scales, ecological shifts rarely accompany species’ invasions (Richardson & McMahon, 1992; Beerling et al., 1995; Higgins et al., 1999; Sutherst et al., 1999; Skov, 2000; Peterson, 2003; Iguchi et al., 2004; Robertson et al., 2004). Such ecological niche modeling (ENM) approaches have seen considerable exploration in the literature in recent years. Clearly, though, given the intense interest in these questions, further tests in invasive taxa are merited.

Gypsy moths, *Lymantria dispar* (L.) (Lepidoptera: Lymantriidae), are a pest species capable of defoliating large areas of forest (Leonard, 1974; Baranchikov, 1989). They have defoliated $>34 \times 10^6$ ha in the USA since 1924 (Tobin et al., 2007), and 80 000 ha in Germany in 1992–1994 (Reineke & Zebitz, 1998). The species is invasive in North America, having been introduced from France in 1869 (Goldschmidt, 1934). It was estimated that in 2001, the total area in the USA infested by *L. dispar* was 78.5×10^6 ha (Mayo et al., 2003). Despite the length of time since its introduction to North America, the range of the moth is still expanding and it occupies roughly one-third of the hypothesized North American range (Sharov, 2004). The relatively slow dispersal of gypsy moths in North America has been attributed to the fact that females of the colonizing populations cannot fly (Sharov, 2004).

Asian populations of gypsy moths are poorly known and poorly characterized in terms of distribution (Hunter & Lindgren, 1995). Although these populations are not currently accorded any formal taxonomic status, differences from European populations exist in life history and basic biology. Genetic differences between the two populations are well documented (Pfeifer et al., 1995; Garner & Slavicek, 1996). Asian populations feed and grow on a broader range of host plants, including conifers (Baranchikov, 1989), hatch at lower temperatures, and reach adulthood more rapidly than European populations

*Correspondence: E-mail: ricw@ku.edu

(Keena, 1996); perhaps most importantly, females from some Asian populations are capable of sustained flight (Schaefer et al., 1984; Baranchikov, 1989; Koshio, 1996) while females from European populations probably, and their North American descendents certainly, are incapable of sustained flight (Reineke & Zebitz, 1998). At the other extreme, some females from Siberian populations are reported to fly up to 100 km (Keena et al., 2001). Baranchikov (1989) described the reduction of ability to fly in females of this species from east to west, stating that a transition zone of occasional female flights exists in eastern Europe.

Although considerable speculation has centered on the ecological potential of gypsy moths drawn from Asian vs. European source populations (Wallner, 1993; Keena et al., 2001), the ecological characteristics of the former have been treated only via assumption of similarity to those of the latter (Matsuki et al., 2001; Gevrey & Worner, 2006). This study explores the ecological characteristics of Asian gypsy moth populations in coarse-scale climatic and topographic dimensions to identify a potential geographic distributional area for the species. Specifically, we characterize areas that share coarse-scale climatic and topographic characteristics with localities from which Asian gypsy moth specimens have been collected. The specimen localities on which we base our model are East Asian, where female gypsy moths are reported to cover long-distance flying (Baranchikov, 1989). Our characterization is based on regional-scale locality data and the model we build is projected onto the global scale. Therefore, it is incomplete in that it does not 'see' finer-scale effects (Soberón & Peterson, 2005), that might be detectable in a study based at a higher scale of resolution. However, it is a useful first step, particularly for such poorly known populations of such an economically important species.

Materials and methods

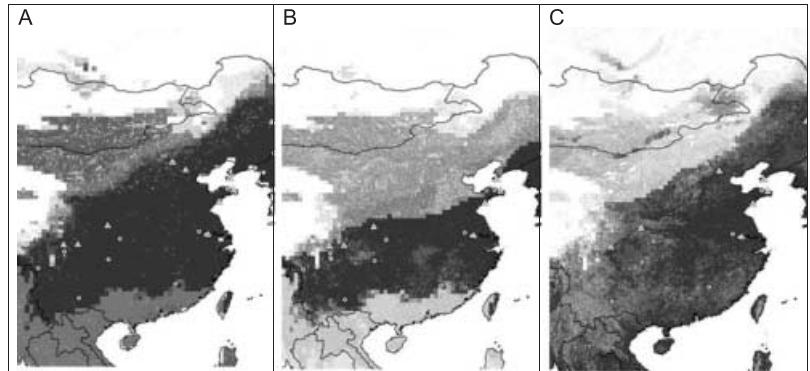
Occurrence data for Asian *L. dispar* populations were accumulated from specimens housed in the natural history collections listed in Acknowledgments. Specimens of this species from East Asia are rare, but at least identifications are relatively unambiguous. All data associated with each specimen were noted and textual collection locality descriptors later converted into geographic coordinates via consultation of Internet gazetteers (Alexandria Digital Library Gazetteer, <http://middleware.alexandria.ucsb.edu/client/gaz/adl/index.jsp> and GEONet Names Server, <http://gnswww.nga.mil/geonames/GNS/index.jsp>), on-line geographical indexes that provide latitude–longitude coordinates for the locations they list. In total, 43 voucher specimens for which geographic coordinates were available

were used to generate this model. These specimens are all from East Asia, where flight in females is well documented (Goldschmidt, 1934; Baranchikov, 1989; Koshio, 1996). Specimens from western and central Asia were excluded, as we were unable to assess whether they were equally capable of flight as more eastern populations.

Digital raster GIS datasets used to describe ecological landscapes included elevation, slope, aspect [from the US Geological Survey's Hydro-1K data set (USGS 2001), native resolution 1 × 1 km], annual precipitation, annual mean temperature, annual mean maximum monthly temperature, annual mean minimum monthly temperature, and solar radiation [for 1960–1990 from the Intergovernmental Panel on Climate Change (New et al., 1997), native resolution 50 × 50 km]. Note that because compound topographic index has not been developed for Australia in the Hydro-1K data set, that continent is included in our projections via an additional set of models developed omitting that data layer. Data sets for training models to characterize the Asian gypsy moth's distribution in the eastern portion of its native range (Schaefer et al., 1984; Baranchikov, 1989) were developed based on East Asian environments only, and tested for predictive accuracy. Once models were validated, we projected model rule sets to the entire world to visualize the global extent of the Asian gypsy moth's potential distribution. All data sets were resampled to 0.1° for analysis to match the approximate resolution of occurrence data.

Our approach to ENM is based on the idea of ecological niches defined as the set of conditions under which a species is able to maintain populations without immigration (Grinnell, 1917, 1924). Several avenues of research have demonstrated widespread evolutionary conservatism in niche characteristics, allowing accurate predictions of invasive species' potential distributions (Peterson, 2003). Our approach consisted of three steps: (i) model ecological niche requirements of the species based on known occurrences on native distribution areas; (ii) test accuracy of native-range predictions; and (iii) project niche model onto other regions to identify areas putatively susceptible to invasion. We used the Genetic Algorithm for Rule-set Prediction (GARP) (Stockwell & Peters, 1999) for ENM development. Genetic Algorithm for Rule-set Prediction uses an evolutionary-computing genetic algorithm to search for non-random associations between environmental variables and known occurrences of species, as contrasted with environmental characteristics across the overall study area. Previous tests of the GARP approach have shown successful predictions of distributional phenomena for numerous taxa and regions, including many applications to predicting the potential distributions of invasive species (Peterson, 2003).

Figure 1 Example of random subsetting of input data to provide a test of model predictivity on the native range of Asian *Lymantria dispar* – dotted circles indicate subset 1, and dotted triangles indicate subset 2. (A) Prediction developed using all occurrence data; (B) prediction based on subset 1, tested with subset 2; and (C) prediction based on subset 2, tested with subset 1. Darkest grey represents areas with complete agreement of 10 ‘best subsets’ of GARP models, lighter grey shows lower agreement, white areas were not predicted by our model.



For testing the predictive accuracy of models, we split available occurrence data for the Asian *L. dispar* populations by setting aside a random 50% of latitude–longitude pairs (Figure 1; these initial analyses were run twice – one with subset 1 predicting subset 2, and vice versa). Within the functioning of the GARP program, the occurrence points to be used to train ENMs are resampled (with replacement) once again to create a population of 1250 presence points; an equivalent number of points is resampled from the population of grid squares (pixels) from which the species has not been recorded. These 2500 points are divided equally into training (for creating models) and internal testing (for evaluating model quality) data sets.

Genetic Algorithm for Rule-set Prediction models consist of sets of conditional rules developed through an iterative process of rule selection, evaluation, testing, and incorporation or rejection (all within the program’s processing, and not subject to user options). First, a method is chosen from a set of possibilities (e.g., logistic regression, bioclimatic rules), and applied to the training data set. Then, a rule is developed using a number of inferential tools mimicking DNA evolution (point mutations, deletions, crossing over, etc.) to maximize predictive accuracy. Rule accuracy is evaluated via the testing data based on the percentage of points correctly predicted as present or absent by the rule. The change in predictive accuracy from one iteration to the next is used to evaluate whether a particular rule should be incorporated into the final ruleset (i.e., did accuracy increase or decrease?). The algorithm runs either 1000 iterations or until addition of new rules has no effect on predictive accuracy (‘convergence’; in this case set at 1%; most models converge at 100–120 iterations, so the maximum iteration parameter is never reached in practice). The final ruleset, or ecological niche model, is then projected onto digital maps to identify potential geographic distribution, either over the native distribution or worldwide. Spatial predictions of presence and absence

can include two types of error: omission (predicted absence in areas of actual presence) and commission (predicted presence in areas of actual absence; Fielding & Bell, 1997).

Because GARP is a random-walk procedure, it does not produce unique solutions; consequently, we followed recently published best-practices approaches to identify an optimal subset of resulting replicate models (Anderson et al., 2003) for final testing and interpretation. In particular, we developed 100 replicate models based on occurrence data for eastern populations of *L. dispar*; of these models, we retained the 20 that showed the lowest omission error rates. Then, we retained the 10 models with intermediate commission error (i.e., we discarded the 10 models with area predicted present showing greatest deviations from the overall median area predicted present across all models). This ‘best subset’ of models was summed to produce final predictions of potential distributions in the form of grids with values ranging from 0 (all models agree in predicting absence) to 10 (all models agree in predicting presence). Because this best-subsets procedure is based on independent subsamples of available occurrence data, it is an effective means of avoiding overfitting of models to occurrence data – the evaluation with distinct subsamples enforces generality in model predictive abilities.

Model predictions were tested initially by overlaying the random 50% of occurrence points that had been set aside prior to ENM development on the predictions developed. At each threshold in the prediction (i.e., 1, 2, 3 . . . 10), we calculated the proportional area predicted present and the success in predicting the independent test data set. We then used these two quantities to calculate cumulative binomial probabilities that the observed degree of predictive success would be achieved were predictions and test points to be random with respect to one another (Anderson et al., 2002). Once model predictions were validated, we combined all available occurrence data, and developed final East Asian

models to produce a worldwide picture of potential distributional areas.

Results

The ENMs developed based on the two random subsets of available occurrence data for *L. dispar* in East Asia were similar in their general predictions regarding the East Asian range of these populations – a broad swath from Japan and northeastern China south to southeastern China (Figure 1). Coincidence of test occurrence points that had been set aside prior to model development with areas of predicted presence was very good – indeed, at all thresholds for both sets of training and testing data, model predictions were better than random expectations (all $P < 0.05$; Figure 2). Indeed in all cases but one, probabilities of observing such close coincidence at random were very low ($P < 0.0001$). Given this statistical significance of ENM predictive power for this species, we explored the broader geographic implications, that is, the species' global potential distribution.

Projecting an all-Asian-data ENM globally (Figure 3), we can develop yet another level of testing of consistency of model predictions. Specifically, we can check to see if the ecological profile of the Asian populations can predict the geographic distribution of the much better known European *L. dispar* populations, which would provide an indication of ecological niche conservatism across the broad distribution of this species. Our Asian model predictions show good general concordance with earlier maps of the

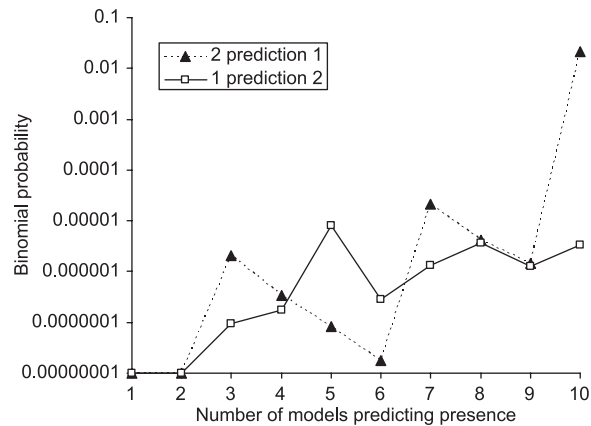


Figure 2 Summary of threshold-dependent binomial probabilities associated with each level of prediction in the two random-subset models of native range of Asian *Lymantria dispar* shown in Figure 1.

European distribution of the species (Schedl, 1936; Giese & Schneider, 1979), indicating that ecological characteristics of the species are uniform between the eastern and western portions of the species' distributional area. The major differences are that our predictions range somewhat further into northern Europe and Asia Minor (i.e., UK, Eire, Fennoscandia, Northern Russia, Northern Turkey) than the known distribution of the European gypsy moth. Giese & Schneider (1979) report that gypsy moths are virtually extinct in Britain, and that individuals found in Fennoscandia have been male immigrants. They were also

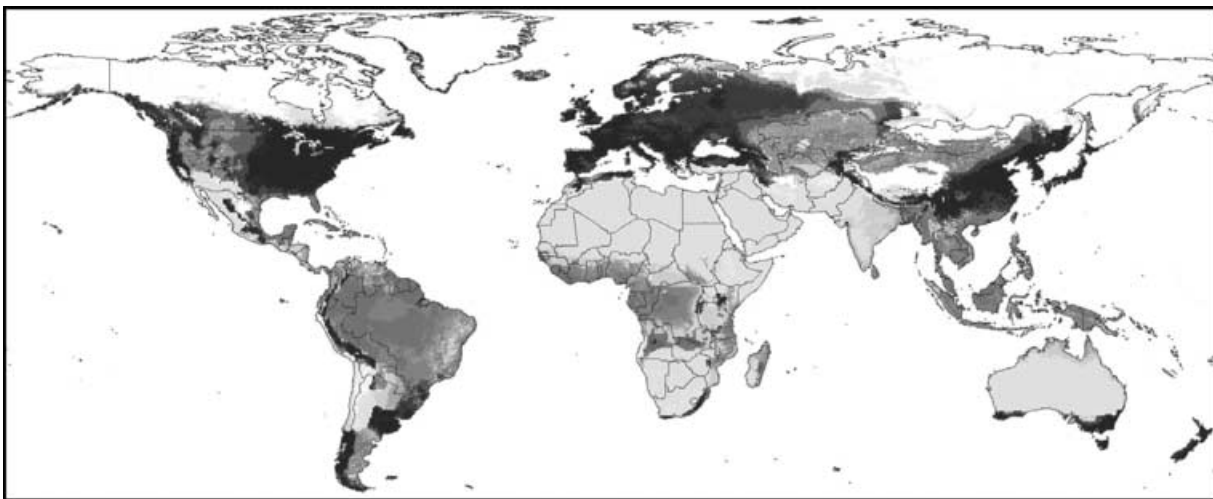


Figure 3 Worldwide projection of native-range ecological niche model of Asian *Lymantria dispar*, showing areas globally that fit the ecological niche profile of the species as characterized on its native range. Note that Australia is included based on a separate suite of models because the compound topographic index coverage is not available for that continent. Darkest grey represents areas with complete agreement of 10 'best subsets' of GARP models, lighter grey shows lower agreement, white areas were not predicted by our model.

unable to corroborate the presence of the gypsy moth in Northern Turkey, an area of strong prediction for us, although *L. dispar* is considered to be widespread and common in Turkey (Unal et al., 1998), and we found several specimens from the region (Southern Turkey, Levant, Caucasus) in museum collections. The global projection of these models indicates broad invasive potential for Asian gypsy moth populations – basically the entirety of North and South Temperate zones.

Discussion

In this article, we analyze the scanty distributional information that we were able to assemble regarding Asian populations of *L. dispar* in the context of climatic variation across Eurasia. These populations are poorly known and poorly characterized, in spite of their potentially great economic impacts. As such, if the natural history differences that have been discussed (Giese & Schneider, 1979; Baranchikov, 1989; Keena, 1996; Koshio, 1996) are real, then consequences of their introductions for natural landscapes and silviculture on other continents could be serious. The ENMs developed herein thus provide some first bits of information regarding the distribution of these populations and their invasive potential.

An interesting insight that emerged from our analyses is that Asian and European *L. dispar* populations appear generally to share ecological characteristics, and as such the Asian-based ENM was able to predict the major features of the geographic distribution of the European populations. The extension of our predictions into the areas mentioned above that are farther north than reported distributions could be a consequence of slight ecological differences that are coarse-scale manifestations of natural history differences, such as broader range of host plants (including conifers), more rapid growth, and greater ability to disperse as a result of female flight (Giese & Schneider, 1979; Koshio, 1996). It is impossible to distinguish which, if any, of these natural history differences translate into differences in distributional potential, especially as we did not include any of these factors in our perforce coarse-scale model. It may be that Asian and European gypsy moth populations have differentiated ecologically in these small details; alternatively our model may be overpredicting into suites of ecological conditions that are not actually usable by Asian gypsy moths (AGMs). Overall, though, AGMs and European gypsy moths (EGMs) appear to share the major features of their coarse-scale ecological distributions.

The global potential distribution of the Asian *L. dispar* populations is quite broad. Put very generally, these populations appear to have the potential to colonize almost all temperate-zone areas, except for montane regions and

deserts. The behavioral and fine-scale ecological differences between Asian and European *L. dispar* populations (Geise & Schneider, 1979; Koshio, 1996) should then be of some concern. Given more rapid maturation (Keena, 1996) and possibly better dispersal abilities (Baranchikov, 1989), were Asian *L. dispar* populations to colonize other regions worldwide (e.g., North America), their control may prove much more difficult.

More generally, this study demonstrates the anticipatory capacity of the ENM tools we used. Poorly known species, and species that have not seen detailed study, can be impenetrable to physiologically based approaches to forecasting distributional potential (Sutherland et al., 1999; Porter et al., 2002), but niche modeling can approach the challenge based only on occurrence information. Indeed, even species that have not yet invaded can be evaluated as to their invasive potential, and risks can thus be evaluated proactively (Peterson, 2003). Although certainly not 100% correct in its predictions, the ENM approach offers some information where little or none was available before. This predictive and proactive approach can be used on broad scales to prioritize use of limited resources in control and prevention of species' invasions.

Acknowledgements

We thank the staff of the following natural history collections for their kind provision of access to specimens and data under their control: American Museum of Natural History (New York), Museum für Naturkunde Berlin, Zoologisches Forschungsinstitut und Museum Alexander König (Bonn), Natural History Museum (London), Zoologische Staatssammlung München, and Naturhistorisches Museum (Vienna). This work was developed in part under a contract from the US Department of Defense.

References

- Anderson RP, Lew D & Peterson AT (2003) Evaluating predictive models of species' distributions: criteria for selecting optimal models. *Ecological Modelling* 162: 211–232.
- Anderson RP, Gómez-Laverde M & Peterson AT (2002) Geographical distributions of spiny pocket mice in South America: insights from predictive models. *Global Ecology and Biogeography* 11: 131–141.
- Baranchikov YN (1989) Ecological basis of the evolution of host relationships in Eurasian gypsy moth populations. *Proceedings, Lymantriidae: A Comparison of Features of New and Old World Tussock Moths* (ed. by WE Wallner & KA McManus), pp. 319–338. US Department of Agriculture, Broomall, PA, USA.
- Berling DJ, Huntley B & Bailey JP (1995) Climate and the distribution of *Fallopia japonica*: use of an introduced species to

- test the predictive capacity of response surfaces. *Journal of Vegetation Science* 6: 269–282.
- Fielding AH & Bell JF (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation* 24: 38–49.
- Garner KJ & Slavicek JM (1996) Identification and characterization of a RAPD-PCR marker for distinguishing Asian and North American gypsy moths. *Insect Molecular Biology* 5: 81–91.
- Giese RL & Schneider ML (1979) Cartographic comparisons of Eurasian gypsy moth distribution (*Lymantria dispar* L., Lepidoptera: Lymantriidae). *Entomological News* 90: 1–16.
- Gevrey M & Worner SP (2006) Prediction of global distribution of insect pest species in relation to climate by using an ecological informatics method. *Journal of Economic Entomology* 99: 979–986.
- Goldschmidt R (1934) *Lymantria*. *Bibliographia Genetica* 11: 1–185.
- Grinnell J (1917) Field tests of theories concerning distributional control. *American Naturalist* 51: 115–128.
- Grinnell J (1924) Geography and evolution. *Ecology* 5: 225–229.
- Higgins SI, Richardson DM, Cowling RM & Trinder-Smith TH (1999) Predicting the landscape-scale distribution of alien plants and their threat to plant diversity. *Conservation Biology* 13: 303–313.
- Hunter AF & Lindgren BS (1995) Range of gypsy moth in British Columbia: a study of climatic suitability. *Journal of the Entomological Society of British Columbia* 92: 45–55.
- Iguchi K, Matsuura K, McNyset K, Peterson AT, Scachetti-Pereira R et al. (2004) Predicting invasions of North American basses in Japan using native range data and a genetic algorithm. *Transactions of the American Fisheries Society* 133: 845–854.
- Keena MA (1996) Comparison of the hatch of *Lymantria dispar* (Lepidoptera: Lymantriidae) eggs from Russia and the United States after exposure to different temperatures and durations of low temperature. *Annals of the Entomological Society of America* 89: 564–572.
- Keena MA, Wallner WE, Grinberg PS & Cardé RT (2001) Female flight propensity and capability in *Lymantria dispar* (Lepidoptera: Lymantriidae) from Russia, North America and their reciprocal F₁ hybrids. *Population Ecology* 30: 380–387.
- Koshio C (1996) Pre-ovipositional behaviour of female gypsy moths, *Lymantria dispar* L. (Lepidoptera: Lymantriidae). *Applied Entomology and Zoology* 31: 1–10.
- Lee CE (2002) Evolutionary genetics of invasive species. *Trends in Ecology and Evolution* 17: 386–391.
- Leonard DE (1974) Recent developments in ecology and control of the gypsy moth. *Annual Review of Entomology* 19: 197–229.
- Matsuki M, Kay M, Serin J, Floyd R & Scott JK (2001) Potential risk of accidental introduction of Asian gypsy moth (*Lymantria dispar*) to Australasia: effects of climatic conditions and suitability of native plants. *Agricultural and Forest Entomology* 3: 305–320.
- Mayo JH, Straka TJ & Leonard DS (2003) The cost of slowing the spread of the gypsy moth (Lepidoptera: Lymantriidae). *Journal of Economic Entomology* 96: 1448–1454.
- New M, Hulme H & Jones P (1997) A 1961–1990 mean monthly climatology of global land areas. Climatic Research Unit, University of East Anglia, Norwich, UK.
- Peterson AT (2003) Predicting the geography of species' invasions via ecological niche modeling. *Quarterly Review of Biology* 78: 419–433.
- Pfeifer TA, Humble LM, Ring M & Grigliatti TA (1995) Characterization of gypsy moth populations and related species using a nuclear DNA marker. *Canadian Entomologist* 127: 49–58.
- Porter WP, Sabo JL, Tracy CR, Reichman OJ & Ramankutty N (2002) Physiology on a landscape scale: plant–animal interactions. *Integrative and Comparative Biology* 42: 431–453.
- Reineke A & Zebitz CPW (1998) Flight ability of gypsy moth females (*Lymantria dispar* L.) (Lep., Lymantriidae): a behavioural feature characterizing moths from Asia? *Journal of Applied Entomology* 122: 307–310.
- Richardson DM & McMahon JP (1992) A bioclimatic analysis of *Eucalyptus nintens* to identify potential planting regions in Southern Africa. *South African Journal of Science* 88: 380–387.
- Robertson MP, Villet MH & Palmer AR (2004) A fuzzy classification technique for predicting species' distributions: applications using invasive alien plants and indigenous insects. *Diversity and Distributions* 10: 461–474.
- Schaefer PW, Weseloh RM, Sun X, Wallner WE & Yan J (1984) Gypsy moth, *Lymantria* (= *Ocneria*) *dispar* (L.) (Lepidoptera: Lymantriidae), in the People's Republic of China. *Environmental Entomology* 13: 1535–1541.
- Schedl KE (1936) *Der Schwammspinner (Porthetria dispar L.) in Euroasien, Afrika und Neuengland*. Verlagsbruchhandlung, Berlin, Germany.
- Sharov AA (2004) The bioeconomics of managing the spread of exotic pest species with barrier zones. *Risk Analysis* 24: 879–892.
- Skov F (2000) Potential plant distribution mapping based on climatic similarity. *Taxon* 49: 503–515.
- Soberón J & Peterson AT (2005) Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodiversity Informatics* 2: 1–10.
- Stockwell DRB & Peters DP (1999) The GARP modelling system: problems and solutions to automated spatial prediction. *International Journal of Geographical Information Systems* 13: 143–158.
- Sutherst RW, Maywald GF, Yonow T & Stevens PM (1999) CLIMEX User Guide. Predicting the Effects of Climate on Plants and Animals. CSIRO Publishing, Collingwood, Victoria, Australia.
- Tobin PC, Liebhold AM & Roberts EA (2007) Comparison of methods for estimating the spread of non-indigenous species. *Journal of Biogeography* 34: 305–312.
- Unal I, Serez M, Eroglu M & Bilgili E (1998) Evaluation of wing traps baited with Disparlure for monitoring Gypsy Moth [*Lymantria dispar* (L.)] populations. *Turkish Journal of Agriculture and Forestry* 22: 329–331.
- USGS (2001) HYDRO1k Elevation Derivative Database. US Geological Survey, Washington, DC, USA. Available at <http://edcdaac.usgs.gov/gtopo30/hydro/>.
- Wallner WE (1993) Attractancy to light and flight behavior of Asian gypsy moth females. *Proceedings: Annual Gypsy Moth Review* (ed. by KE Kuhariz), pp. 30–33. Pennsylvania Department of Environmental Resources, Harrisburg, PA, USA.