

Phenology predicts the native and invasive range limits of common ragweed

DANIEL S. CHAPMAN*, TOM HAYNES†, STEPHEN BEAL†, FRANZ ESSL‡§¶ and JAMES M. BULLOCK||

*NERC Centre for Ecology & Hydrology, Bush Estate, Edinburgh EH26 0QB, UK, †NatureBureau Ltd, 36 Kingfisher Court, Hambridge Road, Newbury RG14 5SJ, UK, ‡Environment Agency Austria, Spittelauer Laende 5, 1090 Vienna, Austria, §Department of Conservation Biology, Vegetation and Landscape Ecology, University of Vienna, Rennweg 14, 1030 Vienna, Austria, ¶Centre for Invasion Biology, Department of Botany & Zoology, Stellenbosch University, Private Bag ×1, Matieland 7602, South Africa, ||NERC Centre for Ecology & Hydrology, Benson Lane, Wallingford OX10 8BB, UK

Abstract

Accurate models for species' distributions are needed to forecast the progress and impacts of alien invasive species and assess potential range-shifting driven by global change. Although this has traditionally been achieved through data-driven correlative modelling, robustly extrapolating these models into novel climatic conditions is challenging. Recently, a small number of process-based or mechanistic distribution models have been developed to complement the correlative approaches. However, tests of these models are lacking, and there are very few process-based models for invasive species. We develop a method for estimating the range of a globally invasive species, common ragweed (*Ambrosia artemisiifolia* L.), from a temperature- and photoperiod-driven phenology model. The model predicts the region in which ragweed can reach reproductive maturity before frost kills the adult plants in autumn. This aligns well with the poleward and high-elevation range limits in its native North America and in invaded Europe, clearly showing that phenological constraints determine the cold range margins of the species. Importantly, this is a 'forward' prediction made entirely independently of the distribution data. Therefore, it allows a confident and biologically informed forecasting of further invasion and range shifting driven by climate change. For ragweed, such forecasts are extremely important as the species is a serious crop weed and its airborne pollen is a major cause of allergy and asthma in humans. Our results show that phenology can be a key determinant of species' range margins, so integrating phenology into species distribution models offers great potential for the mechanistic modelling of range dynamics.

Keywords: allergy, *Ambrosia artemisiifolia*, climate change, invasion, macroecology, process-based model, range shift

Received 8 February 2013 and accepted 14 August 2013

Introduction

Species' ranges are largely considered to be determined by the climate and so climate change is expected to have a major impact on biodiversity (Thuiller *et al.*, 2005). Among the most important documented impacts of recent warming are poleward and uphill range shifts (Kelly & Goulden, 2008; Lenoir *et al.*, 2008) and changing phenology (Menzel *et al.*, 2006; Sherry *et al.*, 2007; Chapman, 2013). Recent studies have suggested that these may be linked (Chuine, 2010) because the timing of development determines exposure to seasonal climatic variation, which will be a key determinant of individual demographic rates, population dynamics and distribution (Inouye, 2008). Spatial variation in

climate and phenology are therefore thought to interact in setting the position of species' range margins (Chuine, 2010).

Despite this, phenology has rarely been included in species distribution models (Chuine & Beaubien, 2001; Morin *et al.*, 2007). Instead, prediction is nearly always based on correlative models that do not explicitly represent biological mechanisms (Thuiller *et al.*, 2005; Dormann *et al.*, 2012). These use data-driven, statistical relationships between climate and species' occurrence to predict range shifts (Thuiller *et al.*, 2005) or forecast non-native species invasion (Petitpierre *et al.*, 2012). Their strength lies in their efficiency for modelling large numbers of species. However, correlative models have been criticized on several grounds. Model fitting and calibration is troubled by spatial autocorrelation (Chapman, 2010; Chapman & Purse, 2011), spurious correlations can arise from spatial bias in the distribution data

Correspondence: Daniel S. Chapman, tel. +44 131 4458 549, fax + 44 131 4453 943, e-mail: dcha@ceh.ac.uk

(Dormann *et al.*, 2012) and a lack of biological process impedes transfer or extrapolation to novel combinations of climatic drivers in potentially invaded regions or after climatic change (Gallien *et al.*, 2010).

This has motivated the development of mechanistic or process-based distribution models that are complementary to the correlative approach (Dormann *et al.*, 2012). These explicitly represent environmental effects on physiology, demography and/or dispersal and predict distributions as the regions in which population persistence is possible (Kearney & Porter, 2009). A useful distinction is often made between process-based models that are fitted vs. 'forward' models whose formulation, parameters and predictions are based on ecological knowledge rather than being tuned to reproduce a known distribution (Dormann *et al.*, 2012). Forward models are especially valuable for testing hypotheses about species' ranges and invasions because they have a much lower potential to predict the correct distribution for the wrong reason than do correlative or fitted mechanistic models (Dormann *et al.*, 2012). For example, if accurate range predictions can be made by projecting a phenology model in space then this will provide powerful evidence for phenological limitation of species' ranges (Chuine & Beaubien, 2001; Chuine, 2010) and yield a mechanistic and biologically informed basis for predicting range shifts and invasive spread. However, while there are several examples of fitted mechanistic models for alien species (e.g. Gallien *et al.*, 2010; Smolik *et al.*, 2010), very few forward process-based models of invasives have been developed (Kearney *et al.*, 2008), and we are unaware of any for invasive plants.

We used a forward model to test whether phenology imposes a limit on the native and invasive distributions of common ragweed (*Ambrosia artemisiifolia* L.). Ragweed is native to North America where it is a serious agricultural weed (Chikoye *et al.*, 1995) and its extremely allergenic pollen is a major cause of allergic rhinitis (hay fever) and asthma (Oswalt & Marshall, 2008). Outside of North America, ragweed has invaded temperate Europe, Asia, Australia and South America, bringing similar problems (Oswalt & Marshall, 2008). Predicting its potential distribution in the native and invaded continents is therefore very important for planning responses to ensure human health and well-being.

Previous modelling from correlative (Dullinger *et al.*, 2009; Essl *et al.*, 2009; Petitpierre *et al.*, 2012; Cunze *et al.*, 2013) and fitted process-based (Smolik *et al.*, 2010) perspectives has suggested that ragweed's invasive distribution is temperature-dependent. Phenological studies have also shown that warming is lengthening the pollen season (Ziska *et al.*, 2011). As

an annual species, we hypothesized that ragweed's poleward and high-elevation range limits would occur where thermal and photoperiod constraints mean that mature seeds rarely develop before winter frost. This study tests that hypothesis by assessing forward predictions of the native and invasive ranges made from a phenology model developed and parameterized from published growth experiments (Deen *et al.*, 1998a,b, 2001; Shrestha *et al.*, 1999). We use the model in three ways: (i) to compare its phenological predictions with observations of wild ragweed plants; (ii) to predict the native and European invaded range of ragweed as the region in which phenological development to reproductive maturity occurs; (iii) to project how climate change may expand this range. In so doing, we explicitly link the phenology and distribution of a highly damaging invasive weed and provide a mechanistic basis for projecting distribution shifts promoted by global climate change.

Materials and methods

Phenology model

We made several substantive adaptations to an existing model of *A. artemisiifolia* phenology (Deen *et al.*, 1998a,b, 2001) allowing it to be used in the novel context of mapping continental-scale phenology and predicting the range. The original model formulation and parameterization are based on growth trials that quantified the rate of ragweed development from stratified seeds to reproductive maturity at fixed temperatures and photoperiods (Deen *et al.*, 1998a,b, 2001).

In the model, phenology is simulated with an hourly time step where each hour contributes $r(T)\lambda(L)$ 'biological hours' of development, i.e. chronological hours at optimal temperature T and photoperiod L . The functions r and λ describe the responses of development rates to T and L , respectively. Each phenological phase has a characteristic duration in 'biological days' (BDs, i.e. 24 biological hours) quantified in the growth experiments (Deen *et al.*, 1998a,b, 2001). The sequence of non-overlapping phases in the model are germination (3.5 BDs), seedling emergence (1 BD, assuming a 1 cm burial depth (Fumanal *et al.*, 2008)), emergence to end of juvenile phase (7 BDs), appearance of main stem terminal bud (4.5 BDs), appearance of pistillate flowers (4.5 BDs), anthesis (4.5 BDs) and seed maturity (14.5 BDs) (Deen *et al.*, 2001). Therefore, a total of 39.5 BDs are needed to complete the lifecycle. By computing a cumulative BD sum through chronological time, the model estimates phenology as the day at which each stage is reached in a given location.

The original model used a triangular function for $r(T)$, which is biologically unrealistic. We replaced this with a generalized plant growth function based on minimum, optimum and maximum growing temperatures, T_{min} , T_{opt} and T_{max} , and a scaling parameter c (Yin & Kropff, 1996),

$$r(T) = \begin{cases} 0 & \text{if } T < T_{min} \\ \left(\frac{T - T_{min}}{T_{opt} - T_{min}} \left(\frac{T_{max} - T}{T_{max} - T_{opt}} \right)^{T_{max} - T_{opt}} \right)^c & \text{if } T_{min} \leq T \leq T_{max} \\ 0 & \text{if } T > T_{max} \end{cases}$$

We fitted $r(T)$ to growth rate data digitized from three published data sets (Deen *et al.*, 1998b; Shrestha *et al.*, 1999) by least squares ($R^2 = 0.956$) with parameters $T_{min} = 4.88$ °C, $T_{opt} = 30.65$ °C, $T_{max} = 42.92$ °C and $c = 1.696$ (Table 1, see Appendix S1 for details). These cardinal temperatures are close to those in the original model, but our function limits low temperature growth more strongly, as is consistent with other experimental data on development rates (Shrestha *et al.*, 1999) (Appendix S1).

From the growth experiments, the modelled photoperiod response $\lambda(L)$ delays flowering when the day is longer than 14.5 h, which occurs in summer at latitudes above 36.5°N. The photoperiod delay is controlled by a sensitivity parameter L_s , taking a value of 0.400 from the end of the juvenile phase to the appearance of pistillate flowers (see below), and a value of 0 (i.e. no sensitivity) at other stages of the life cycle (Deen *et al.*, 1998a, 2001) (Appendix S1),

$$\lambda(L) = \begin{cases} e^{(L-14.5)\ln(1-L_s)} & \text{if } L \geq 14.5 \\ 1 & \text{if } L < 14.5 \end{cases}$$

The original model was based on planted stratified seed and so gives no indication as to when seed dormancy is broken and BD accumulation should begin. Studies on other plants have successfully applied 'chilling degree day' models whereby the species must accumulate exposure to low temperature before breaking winter dormancy (Chuine, 2000). Since we had no data on which to model such an effect, we elected to break seed dormancy on the first day after the spring equinox when the average daily minimum temperature

exceeds T_{min} . The estimated T_{min} is close to the minimum known ragweed germination temperature (Shrestha *et al.*, 1999) and the temperature evaluated as the best of three alternatives for stratifying ragweed seed (Willemssen, 1975). The equinox constraint prevents unrealistically early germination in the far southern parts of the USA where average winter temperatures do not fall as low as T_{min} . We note that this will have no effect on predictions of the northern range margin, where winters are always cold enough.

During model testing, we found a positive correlation between latitude and predicted anthesis date. However, *Ambrosia* pollen season start dates reported for 10 North American locations between 30 and 52°N in 1995 and 2009 are not significantly correlated with latitude ($n = 20$, $r = -0.236$, $P = 0.316$) and had changed little (mean of 2.7 days earlier, within the start date estimation error) (Ziska *et al.*, 2011). The biological explanation for this is likely to be local adaptation of phenology (Hodgins & Rieseberg, 2011), possibly in the photoperiod response. Since there are insufficient data to model this, we enforced a minimum anthesis date of day 208 (27 July) which is the median pollen season start date across America (Ziska *et al.*, 2011). This predicts flowering to occur synchronously in warm low-latitude locations, but later in cooler and more northerly latitudes where sufficient BDs are not accumulated before day 208. As with the assumption about dormancy breaking, this does not affect predictions of northern range limits since these are in sufficiently cool locations to delay flowering.

Validation of predicted phenology

Validating the phenological predictions is a precursor to using the model to predict the species' range. As such, we tested the

Table 1 Phenology range model parameters, their sources (also see Appendix S1) and results of a sensitivity analysis. For the latter, the model was run with 250 random parameter draws (uniform distributions given in the table) over sample blocks from North America (−95–90° longitude, 40–60° latitude) and Europe (10–15° longitude, 45–65° latitude). Sensitivity was estimated as the t -value from a multiple linear regression predicting the modelled northern range limit (highest latitude where reproduction is possible) from the parameters ($R^2 = 0.900$ for North America and 0.854 for Europe). This gives a standardized measure of the parameter effect direction (sign) and importance (absolute magnitude)

Parameter	Default value	Source	Range for sensitivity analysis	Sensitivity for North America	Sensitivity for Europe
Seed dormancy breaking and minimum development temperature (T_{min})	4.88 °C	A	0–7.5 °C	−20.8	−16.5
Optimum development temperature (T_{opt})	30.65 °C	A	28–33 °C	−10.1	−7.60
Maximum development temperature (T_{max})	42.92 °C	A	40–46 °C	−0.08	2.18
Temperature shape parameter (c)	1.70	A	1–2	−15.5	−14.2
Threshold day length (L_0)	14.5 h	B	13–16 h	22.7	13.3
Photoperiod sensitivity (α)	0.40	B	0–1	−26.4	−23.8
Minimum day of anthesis	208	C	200–216	−1.3	0.91
Growing season termination temperature	0 °C	D	−7.5–7.5 °C	1.2	−0.02

A, Fitted to data on leaf appearance, shoot growth and seed germination rates from controlled environment growth experiments with populations from southern Canada (Deen *et al.*, 1998b; Shrestha *et al.*, 1999).

B, Values based on original phenology model, developed from growth experiments in southern Canada (Deen *et al.*, 1998b).

C, Median start of *Ambrosia* pollen season for 10 monitoring stations across North America (Ziska *et al.*, 2011).

D, Widespread observation that ragweed is killed by frost.

model's ability to predict ragweed phenology using data from across the native range in 2009–2012 (USA National Phenology Network, 2013). These data were not used for parameterization and so allow independent evaluation of the model's applicability to wild populations. Records from five observers reporting implausible flowering before the summer equinox were excluded, as these are not consistent with ragweed's short-day nature (Ziska *et al.*, 2011). We also removed leafing observations as leaves occur at all times between emergence and senescence, so are not informative for testing phenology model predictions.

This left 47 georeferenced observations, comprising a dated record of the phenophase of a ragweed individual, at latitudes of 28–46°N. Although the data set was small, it is the only available information for validating the model. NPN phenophases were assigned to the corresponding phases (or range of phases) represented in the model. To assess how well the model predicted the observed phenology, the day ranges when ragweed was predicted to be in those phases was estimated, and the range midpoint plotted against the observation day. As a test statistic for comparing these, the root-mean-square-error (RMSE) was calculated.

Range prediction

We used our refined phenology model to make a binary prediction of the native (North America) and invaded (Europe) cold range margins, by estimating the region where seed maturity was reached before autumn frost. We also mapped the region where ragweed could germinate and grow to anthesis, but was killed by frost before setting seed, as introduced ragweed plants growing in these areas may lead to occurrence records of the species. Termination of the growing season was modelled as the first day when minimum temperatures fell to 0 °C, when we expect frost to kill plants and terminate seed ripening.

Gridded long-term average (1960–1990) hourly temperatures were estimated from monthly average minimum and maximum temperatures in the 2.5 arc-min WorldClim database (Hijmans *et al.*, 2005). Monthly averages were projected onto 5 × 5 km equal area grids (Albers Equal Area Conic for North America and Lambert Azimuthal Equal Area for Europe – this resolution was chosen as a compromise between computational demand and strong topographic variation in temperature in mountainous regions near ragweed's range margin) and temporally downscaled to a daily resolution using a method based on bias-corrected regression splines. This involved fitting a thin plate regression spline (R package 'mgcv'; Wood, 2003, 2013) with one degree of freedom per month to the averages and computing predicted values for each day. Monthly means recovered from this were strongly correlated with the observed ($r > 0.999$ for every month, $r > 0.997$ for every grid cell), but overpredicted the coldest month and underpredicted the warmest month. To remove this bias, we refitted the spline to monthly data that was expanded or contracted about its annual mean by the transformation $b(T - \bar{T}) + \bar{T}$, where T is the monthly temperature and $b = 1.0247$ minimized the sum of squares between observed

and fitted recovered monthly mean temperatures for 1000 randomly chosen grid cells. Hourly temperature time series were created by assuming temperatures pass between the estimated daily minima and maxima following a transformed sine wave with 24-hour periodicity.

We investigated the effects of the model parameters on the range prediction by means of a sensitivity analysis using random parameterizations drawn within fixed limits (Table 1). Because of the model's computational demands, we restricted this to 250 parameterizations and a subset of the native and invaded range, centred on the margin.

Testing the range prediction

Range predictions of the phenology model were contrasted with the observed distribution in both continents. For this, we assembled a database of ragweed occurrences in North American counties, using the county-level United States Department of Agriculture Plants Database as a starting point. This was supplemented with records from reliable online sources and the literature (see Appendix S2). European occurrences from 1990 to 2010 were compiled on 50 × 50 and 10 × 10 km grids. Data were retrieved from online databases, published maps, literature references and databases held by herbaria, universities and individuals (Appendix S2). Data quality varied among countries due to different survey efforts. For example, no data were retrieved for Iceland, Russia, Belarus, Lithuania, Estonia, Bosnia and Herzegovina, Albania, FYR Macedonia, Kosovo or Turkey.

As the forward model prediction is entirely independent of the distribution, we calculated standard measures of agreement between the binary range prediction and the observed presence or 'absence' (lack of a record). These were sensitivity (proportion of presences correctly predicted), specificity (proportion of 'absences' correctly predicted), Cohen's kappa and the true skill statistic (TSS) (Allouche *et al.*, 2006).

We also conducted a more sophisticated test that accounted for two major limitations in the former measures, namely that kappa strongly depends on the ratio of presences to 'absences' and both kappa and TSS treat presence as equivalent to 'absence' (Allouche *et al.*, 2006). This latter is highly questionable since a lack of filling within a species' predicted range will occur because of limitation by nonmodelled factors, e.g. drought, land use, dispersal or under-recording (Petitpierre *et al.*, 2012). Instead we wished to assess whether ragweed can only persist within the region where phenology permits reproduction and whether this correlates with its range margin, indicating that phenology is an important determinant of the distribution (Kearney *et al.*, 2008).

To do this, we examined variation in sensitivity for two spatially- or climatically informed range scenarios. First the phenologically predicted range was expanded or contracted by fixed distances, and sensitivity was plotted as a function of the expansion/contraction distance. Second, we identified the monthly mean, minimum or maximum temperature isotherm most closely corresponding to the range prediction in each continent. We then plotted sensitivity against a range of isotherm values for that month. In both cases we expected an

optimal range prediction to lie at the transition between very high sensitivity (overly optimistic prediction with too large a region suitable) and a rapid drop-off in sensitivity (overly conservative model with many records beyond the margin).

Projection to 2050

To illustrate the use of the model in predicting climate change-driven range shifts, we predicted the area suitable for ragweed reproduction using downscaled projections of monthly mean temperatures in the 2050s for emissions scenario A2a of the Intergovernmental Panel on Climate Change Special Report on Emissions Scenarios (SRES). Spatially downscaled, gridded monthly minimum and maximum temperature estimates from the Hadley Centre Coupled Model (HADCM3) (Johns *et al.*, 2003) were obtained from the CCAFS-climate data portal (<http://www.ccafs-climate.org/>). This predicts mean increases in mean annual temperature of 2.6 °C in Europe and 3.1 °C in North America, mainly with greater warming during summer than winter. The largest predicted increases in mean annual temperature are in eastern and Arctic areas, with the least warming near the western coasts. Projected monthly data were converted to hourly time series equivalently to the current-day data for use in the phenology model.

Results

Phenology predictions aligned reasonably with independent observations from the USA National Phenol-

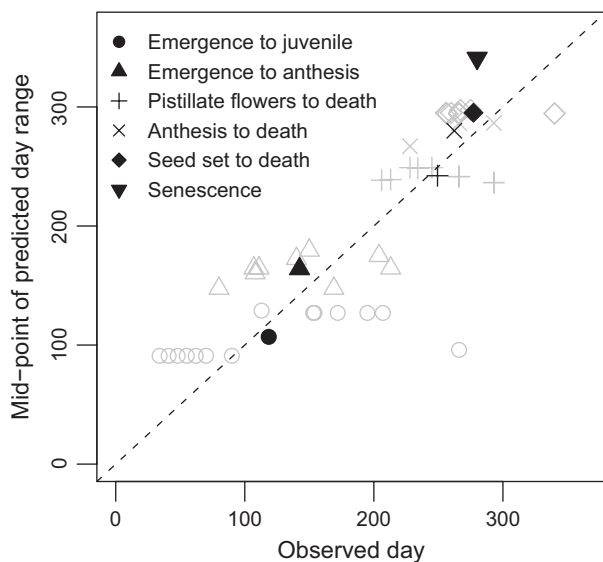


Fig. 1 Comparison of the phenology model predictions with phenological observations of *A. artemisiifolia* in 2009–2012 from the USA National Phenology Network (NPN). NPN phenophases were assigned to a corresponding range of phases in our model and the observed day plotted against the predicted day range midpoint. Predictions therefore show the long-term average expected day of the observations.

ogy Network (Fig. 1; $n = 47$, regression slope = 0.899, $R^2 = 0.708$). Calculation of RMSE showed an average absolute difference between the predicted phenophase midpoints and the actual observations of 46.6 days. This large difference can be attributed to several factors including the comparison of phase midpoints with actual days, the use of long-term average climate data rather than meteorological data and prediction of population averages vs. observations of individual plants, as well as error in model specification of ragweed's phenology. Indeed the phenophase group means (which average out much of the observation data noise) were much better predicted by the model (mean RMSE weighted by group size = 17.7 days).

At the landscape scale, the phenology model predicts that in an average year ragweed can reach maturity and produce seed in lowland USA and southern Canada, and in lowland Europe south of northern Britain, Estonia and Fennoscandia (Fig. 2a–c). The higher mountain ranges (e.g. Rocky Mountains, Sierra Nevada and Alps) are predicted too cold for successful reproduction (Figs. 2–3). Sensitivity analysis showed that these predictions were most strongly affected by the photoperiod response parameters and the minimum growing temperature (Table 1).

Visual comparison with the distribution shows that the reproductive boundary aligns to the poleward and high-elevation limits of the species in both the native North American and invaded European ranges (Figs. 2–3). A few ragweed occurrences lay beyond the predicted margin in both continents, mainly in the region with flowering but not reproduction (Fig. 2). No southern range limit was predicted because temperatures were not high enough to limit phenological development and other factors which may be more important here (e.g., drought or lack of winter chilling) did not feature in the model.

Formal measures of agreement showed very high sensitivity (excellent prediction of presences; 0.997 for North America and 0.920 for Europe) but poor specificity (most 'absences' within the predicted range; 0.055 for North America and 0.418 for Europe). As a result kappa and TSS were low. On first consideration this suggests the model performed badly. However, our testing of the range prediction based only on ragweed presences, which we believe to be more appropriate, gave more optimistic results.

The very high sensitivity shows that ragweed rarely occurs where the phenology model predicts that the species cannot complete its lifecycle. Furthermore, two tests showed that this high sensitivity was not due to the model predicting too large a region to be suitable. First, expansion or contraction of the predicted range caused an abrupt transition between a loss of sensitivity

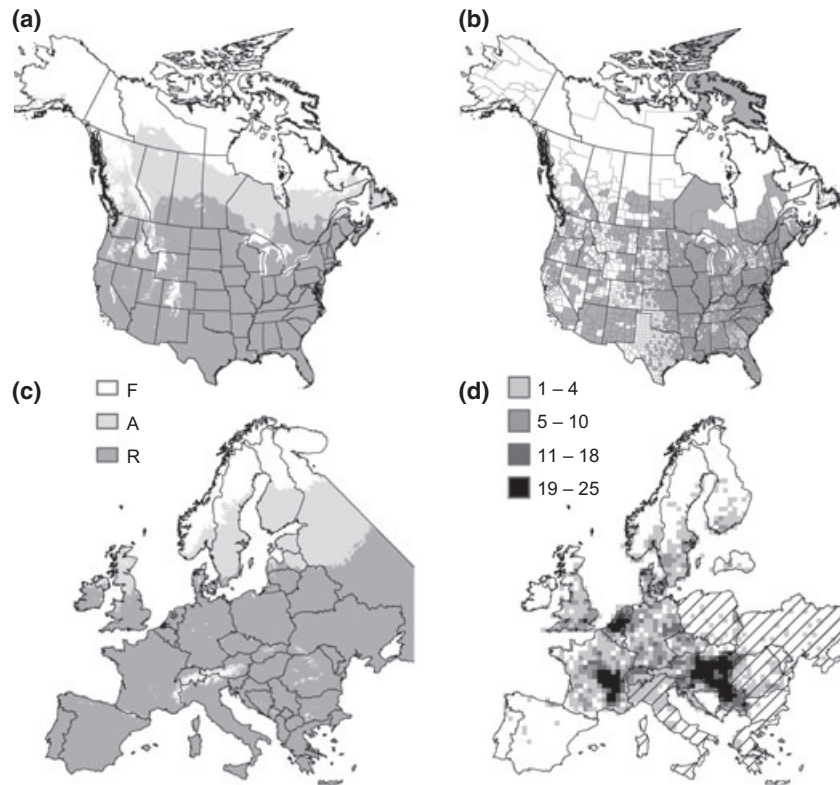


Fig. 2 (a and c) Phenology model prediction of the regions in which *A. artemisiifolia* fails to reach maturity (F), grows to anthesis but fails to produce mature seed (A) or successfully reproduces (R) before autumn frost strikes in an average year in North America and Europe. The predicted range margin is at the boundary of R and A. (b) North American counties in which ragweed has been recorded. The occupied islands in the Canadian Arctic are a single county with one record. (d) The European distribution at a 50×50 km grid scale, expressed as the number of constituent 10×10 km grid cells with a ragweed record (or 1 if only a 50×50 km resolution record is available). Hatched countries are considered to have poor quality distribution data, while countries where we obtained no data are omitted.

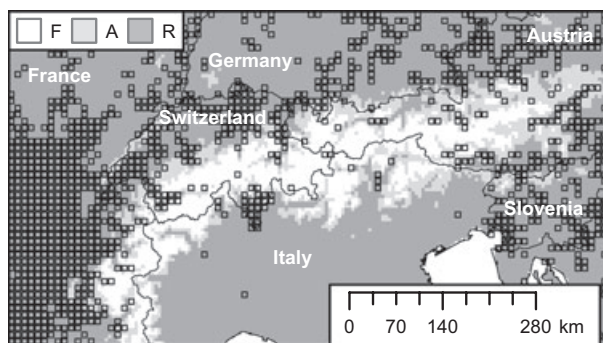


Fig. 3 Phenology model prediction around the European Alps equivalent to Fig. 2c, with 10×10 km records of *A. artemisiifolia* occurrence overlaid as open squares. Very few records were obtained for Italy despite widespread invasion (Déchamp *et al.*, 2009).

during contraction, and a negligible gain in sensitivity during expansion (Fig. 4a–b). Second, the isotherm most closely aligned to the range prediction was close to the transition between very high sensitivity and

sensitivity falling off rapidly (Fig. 4c–d). Both tests show that the northern and uphill range limit prediction of the phenology model approximately bounded but did not exceed the ragweed occurrences, both in geographic and climate space.

Applying the model to predicted temperatures in the 2050s showed substantial northwards and uphill shifts in the range margins (Fig. 5 compared to Fig. 2). The model predicted expansion into central and eastern Canada and northeast Europe (e.g., Sweden, Finland, Estonia and Russia). This reflects both increases in summer temperatures and delays in autumn frost (median of 10 days later in North America and 18 days later in Europe). No change in the southern part of the range was projected by the model for the same reason as for the current day.

Discussion

Using a forward and process-based phenology model, we showed a clear correspondence between the

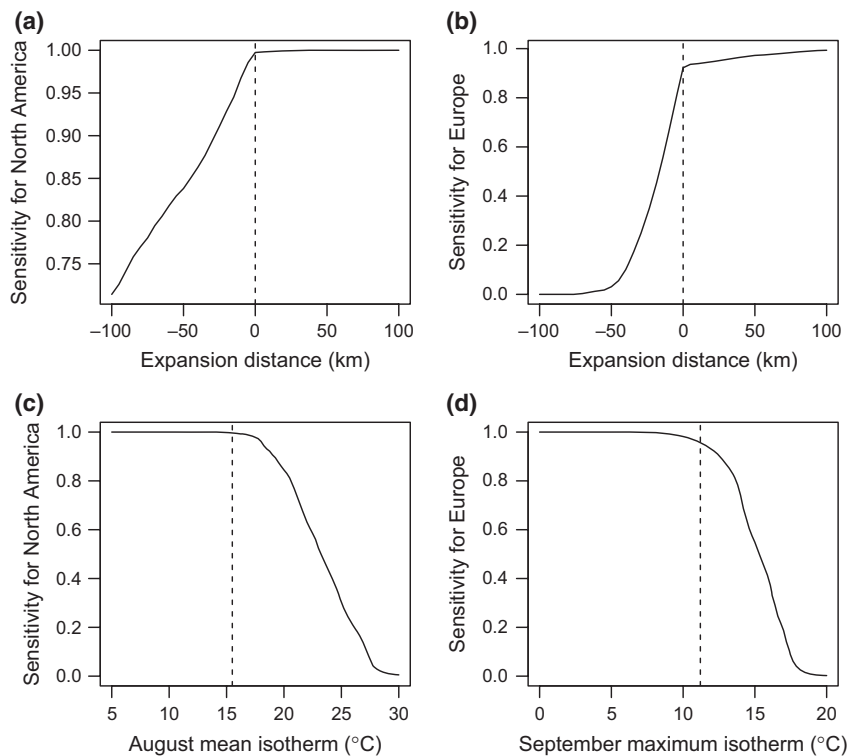


Fig. 4 (a–b) Expansion of the phenology model range limit prediction using Euclidean distance buffer functions has little effect on sensitivity (proportion of presences correctly predicted) in North America and Europe. By contrast, contraction sharply reduces sensitivity. (c–d) Temperature isotherms that most closely match the predicted range (dashed lines, as estimated by the kappa statistic between isotherm and predicted range) are also close to the transition between high and sharply falling sensitivity for ragweed occurrences in both continents. The type of isotherm used in each continent was the one that most closely matched the range prediction.

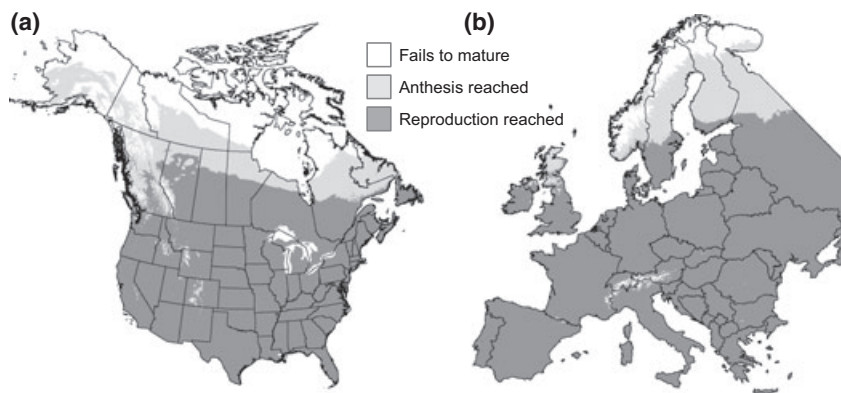


Fig. 5 Illustration of the use of the model for projecting future range expansion driven by climate change. Maps show predicted ragweed ranges in the 2050s according to the SRES A2a emissions scenario and HADCM3 climate model, equivalent to Fig. 2a,c.

predicted limit of ragweed life-cycle completion and its observed northern and high-elevation range limits in two continents. This suggests that thermal and photoperiod constraints on development are a key determinant of the 'cold' range margins, leading to the firm prediction that climatic warming will increase the area in which ragweed can reproduce. Ragweed has strong

human-aided dispersal ability (Lavoie *et al.*, 2007) and so range expansion seems almost certain. Predicting spread of this species is very important given its invasive nature and significant impacts on crops and human health (Chikoye *et al.*, 1995; Oswald & Marshall, 2008). Our process-based model and accurate 'forward' range prediction is therefore an important step towards

a biologically informed modelling of native and invasive species distributions.

Although correlative models would represent similar associations, they are fitted to the distribution and lack ecological process, so there will always be uncertainty over their functional significance and transferability in space or time (Dormann *et al.*, 2012). By contrast, our process-based model explicitly represents ragweed development and was formulated and parameterized from published phenological experiments (Deen *et al.*, 1998a,b, 2001; Shrestha *et al.*, 1999) and first principles, rather than being fitted. Therefore, it is interesting to compare how our model differs from correlative models. The most relevant example for *A. artemisiifolia* is by Cunze *et al.* (2013), who fitted several models to the native range to predict the invasive distribution in Europe. Their prediction for the range expansion up to 2080 for the same scenario as in Fig. 5 indicates a quite different pattern than was predicted by this model (albeit over a longer time period), with less northwards spread in western Europe and greater spread in eastern Europe. One reason for this difference may be in our model's depiction of photoperiodic limitation, which our sensitivity analysis (Table 1) shows could be very important in limiting latitudinal range expansions driven by climatic warming. We cannot conclude which model makes the better prediction, but this nevertheless highlights the potential for process-based models to make quite different predictions compared with correlative ones. Since nearly all predictive studies of climate change impacts on species distributions use correlative models (Thuiller *et al.*, 2005; Dormann *et al.*, 2012), the difference between both modelling strategies adds to their uncertainty.

To our knowledge, this is the first time phenology has been used to predict an invasive plant distribution. Phenological limitation of tree distributions has previously been demonstrated using a model that integrates phenology and mortality in winter and drought (Chuine & Beaubien, 2001; Morin *et al.*, 2007). Drought is undoubtedly also important for ragweed (Shrestha *et al.*, 1999), as is suggested by the thinning of the distribution towards southwest USA and Mediterranean Europe (Fig. 2; Dullinger *et al.*, 2009; Essl *et al.*, 2009; Petitpierre *et al.*, 2012). Including a drought effect in the model would increase the accuracy of the current day prediction and allow forecasts of changes in the southern range margin driven by future changes in precipitation. However, a lack of experimental data meant that we were unable to include drought in the model without fitting to the distributions (a factor also common to the tree studies).

Instead we concentrated on defining the thermal and photoperiodic phenological limits using a process-based

model parameterized from experimental studies, ensuring a strict 'forward' prediction entirely independent of the observed phenological observations and distribution patterns (Dormann *et al.*, 2012). Virtually all ragweed occurrences lay within the predicted range, leading to very high sensitivity and providing good evidence that ragweed cannot persist in areas where frost truncates development of its lifecycle. However, specificity (correct prediction of 'absence') was very low and so kappa and TSS, two standard measures of model agreement, were also low. We contend that in this analysis poor specificity is not necessarily a weakness as it can be explained by two factors that have nothing to do with the performance of the model in identifying areas phenologically suited to ragweed persistence. First, there is a lack of range filling within the predicted and observed range (e.g. due to limitation by nonmodelled abiotic or biotic factors, dispersal constraints or poor recording) (Petitpierre *et al.*, 2012). Therefore, many locations without records will be phenologically suitable, and potentially or actually inhabited. Second, the analysis was hampered by geography – the correct prediction of absence from most of Canada had little effect on the agreement statistics because of the large county size (Fig. 2b), while the prediction of a high latitude margin in Europe meant there was only a small region of (largely correct) predicted absence (Fig. 2c). As pointed out by Kearney *et al.* (2008), forward process-based models aim to map that part of the species' fundamental niche explicitly formulated in the model, i.e. the potential range with respect to the modelled processes, while the actual geographic range of the species is expected to be more restricted.

This raises an important issue over how similar forward process-based models should be properly validated. Our approach was to examine variation in sensitivity for two spatially or climatically informed range predictions (Fig. 4). The spatial test showed that the prediction geographically bounded but did not exceed the known ragweed occurrences. The climatic test showed that the predicted range approximately spanned the warmest region in which the ragweed occurrences could be contained. From this, we conclude that the high model sensitivity was achieved from an extremely conservative prediction, rather than by predicting too large a region to be suitable. This not only suggests that phenology contributes to the species' fundamental niche, but that it is the limiting factor determining ragweed's 'cold' range margins in both the native and invasive distribution.

Given the arguments presented above, the most serious inaccuracies of the model are when ragweed records occur beyond the predicted range. This was a bigger problem in Europe than the native region

(ignoring the single anomalous occurrence in the Canadian Arctic, Fig. 2b). However, the literature on *A. artemisiifolia* in northern Europe reveals that these occurrences represent cases where the species has been accidentally introduced as a contaminant of imported agricultural or bird seed, but failed to reproduce and persist. For example, consistent reproductive failure is reported from Norway, Finland, Sweden (away from the southern coast) and Estonia (Dahl *et al.*, 1999; Saar *et al.*, 2000; Déchamp *et al.*, 2009). By contrast, there are many references to seed production in the northern part of the predicted range, such as in Germany, Netherlands, southern UK, Poland, coastal Sweden and Lithuania (Rich, 1994; Dahl *et al.*, 1999; Saar *et al.*, 2000; Brandes & Nitzsche, 2006; Déchamp *et al.*, 2009; Sauliene *et al.*, 2011). These reports confirm the model predictions about where ragweed is able to set seed but suggest that ragweed can be recorded beyond the phenologically-suitable region because of repeated introductions (Gaudeul *et al.*, 2011), causing part of the apparent difference between the model and the data.

In addition many of the northern populations within the predicted range are considered casual, despite successfully reproducing (e.g., UK, northern Germany, the Netherlands; Déchamp *et al.*, 2009). Cool summers in these locations probably mean that although the lifecycle is completed, the number, survival or viability of seeds is too low to sustain long-term population growth. Without repeated introductions the observed invaded range would therefore probably be restricted to below approximately 50°N, where ragweed is most invasive (Déchamp *et al.*, 2009; Dullinger *et al.*, 2009). As a consequence, the projected spread of the species in Fig. 5 will likely overpredict the region where ragweed will become a major problem in the future. This emphasizes the need to integrate interactions between demography, phenology and dispersal in the future development of process-based distribution models (Chuine & Beaubien, 2001; Dullinger *et al.*, 2009; Dormann *et al.*, 2012). In the context of this study, the phenology model seems to accurately predict the limits of ragweed reproduction, but the species has been introduced across Europe and the serious invasion seems to be limited by other factors not captured in the model.

The model was mainly parameterized with experimental data on populations near the northern edge of the native range (Table 1). A sensitivity analysis showed that range prediction was most sensitive to the two photoperiod response parameters (L_0 and α) and the minimum growth temperature (T_{min}). Uncertainty in these parameters will therefore lead to uncertainty in the position of the predicted range margin. This may be particularly important for T_{min} , since ragweed

emergence was only modestly well predicted (Fig. 1). Where confidence intervals or distributions can be placed on these parameters it would be possible to estimate this margin uncertainty through a sampling of parameter space. Such an exercise was beyond the scope of this study, but could prove useful for evaluating apparent mismatch between the observed and predicted range margins.

Further uncertainty may arise through geographical variation in the model parameters, not captured in the model and consistent with local adaptation (Chuine & Beaubien, 2001; Hodgins & Rieseberg, 2011). However, we do not consider this very important for this study as the model was parameterized near the northern edge of the native range where plants should be close to the limit of adaptation to cold and northerly conditions. Furthermore, the European populations are mainly derived from the northern part of the native range (Gaudeul *et al.*, 2011).

A further limitation was our use of long-term average temperature data. We would ideally have used annually varying daily meteorological data but these were not available at sufficiently high resolutions for both continents. Nevertheless, we were able to investigate this for a region where such data were available (UK). We found a sharp spatial transition between successful reproductions in nearly all years vs. very few years, centred on the climatologically predicted margin (Appendix S3). Furthermore, post-1990 warming, which is not captured in the average temperature database, had only a small impact on the prediction (Appendix S3). We suggest that this justifies our deterministic range prediction based on climatological data.

This study is among the first applications of a forward process-based model for predicting invasive species' distributions (Kearney *et al.*, 2008). While correlative models are useful tools for understanding the structure and dynamics of species' ranges (Thuiller *et al.*, 2005; Chapman *et al.*, 2008; Chapman & Purse, 2011; Petitpierre *et al.*, 2012) we believe that a wider adoption of forward process-based models would be a major advance. However, the development of such models is difficult and we suggest it requires the following steps: (i) collection of experimental data on how environmental drivers affect key biological processes; (ii) formulation of models to capture those effects; (iii) collection of independent data to test predictions of the processes (Fig. 1); (iv) model estimation of the region of potential persistence (Figs. 2–3) and (v) testing model predictions against distribution data (Fig. 4). Further work will establish the most important modelled processes, though these are likely to include phenology (this study, Chuine, 2010; Chuine & Beaubien, 2001), mortality (Morin *et al.*, 2007), energy/mass balances

(Kearney & Porter, 2009), dispersal (Kearney *et al.*, 2008; Smolik *et al.*, 2010; Bullock *et al.*, 2012), local adaptation (Morin *et al.*, 2007) and interspecific interactions (Bullock *et al.*, 2008). Capturing all of these within one model will always be challenging and require considerable empirical and theoretical effort. However, we suggest that integrating strongly climate-dependent biological processes such as phenology into distribution models will be very important for accurately predicting impacts of climate change on biodiversity and the progress of ongoing invasions.

Acknowledgements

This research was funded by the European Commission under ENV.B2/ETU/2010/0037. We thank the contributors of ragweed distribution data. Phenology data were provided by the USA National Phenology Network and the many participants who contribute to its Nature's Notebook program. We thank Carsten Dormann and three anonymous referees for commenting on an earlier draft.

References

- Allouche O, Tsoar A, Kadmon R (2006) Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, **43**, 1223–1232.
- Brandes D, Nitzsche J (2006) Biology, introduction, dispersal, and distribution of common ragweed (*Ambrosia artemisiifolia* L.) with special regard to Germany. *Nachrichtenblatt des Deutschen Pflanzenschutzdienstes*, **58**, 286–291.
- Bullock JM, Edwards RJ, Carey PD, Rose RJ (2008) Geographical separation of two *Ulex* species at three spatial scales: does competition limit species' ranges? *Ecography*, **23**, 257–271.
- Bullock JM, White SM, Prudhomme C, Tansey C, Perea R, Hooftman DAP (2012) Modelling spread of British wind-dispersed plants under future wind speeds in a changing climate. *Journal of Ecology*, **100**, 104–115.
- Chapman DS (2010) Weak climatic associations among British plant distributions. *Global Ecology and Biogeography*, **19**, 831–841.
- Chapman DS (2013) Greater phenological sensitivity to temperature on higher Scottish mountains: new insights from remote sensing. *Global Change Biology*, **19**, 3463–3471.
- Chapman DS, Purse BV (2011) Community versus single-species distribution models for British plants. *Journal of Biogeography*, **38**, 1524–1535.
- Chapman DS, Oxford GS, Dytham C (2008) Process from pattern in the distribution of an endangered leaf beetle. *Ecography*, **32**, 259–268.
- Chikoye D, Weise SF, Swanton CJ (1995) Influence of common ragweed (*Ambrosia artemisiifolia*) time of emergence and density on white bean (*Phaseolus vulgaris*). *Weed Science*, **43**, 375–380.
- Chuine I (2000) A unified model for budburst of trees. *Journal of Theoretical Biology*, **207**, 337–347.
- Chuine I (2010) Why does phenology drive species distribution? *Philosophical Transactions of the Royal Society B: Biological Sciences*, **365**, 3149–3160.
- Chuine I, Beaubien EG (2001) Phenology is a major determinant of tree species range. *Ecology Letters*, **4**, 500–510.
- Cunze S, Leiblein MC, Tackenberg O (2013) Range expansion of *Ambrosia artemisiifolia* in Europe is promoted by climate change. *ISRN Ecology*, **2013**, 9.
- Dahl Å, Strandhede S-O, Wihl J-Å (1999) Ragweed – An allergy risk in Sweden? *Aerobiologia*, **15**, 293–297.
- Déchamp C, Méon H, Reznik S (2009) *Ambrosia artemisiifolia* L. an invasive weed in Europe and adjacent countries: the geographical distribution (except France) before 2009. In: *Ambrosie: The First International Ragweed Review*. (eds Déchamp C, Méon H), pp. 26–46. Saint-Priest, France, AFEDA.
- Deen W, Hunt LA, Swanton CJ (1998a) Photothermal time describes common ragweed (*Ambrosia artemisiifolia* L.) phenological development and growth. *Weed Science*, **46**, 561–568.
- Deen W, Hunt T, Swanton CJ (1998b) Influence of temperature, photoperiod, and irradiance on the phenological development of common ragweed (*Ambrosia artemisiifolia*). *Weed Science*, **46**, 555–560.
- Deen W, Swanton CJ, Hunt LA (2001) A mechanistic growth and development model of common ragweed. *Weed Science*, **49**, 723–731.
- Dormann CF, Schymanski SJ, Cabral J *et al.* (2012) Correlation and process in species distribution models: bridging a dichotomy. *Journal of Biogeography*, **39**, 2119–2131.
- Dullinger S, Kleinbauer I, Peterseil J, Smolik M, Essl F (2009) Niche based distribution modelling of an invasive alien plant: effects of population status, propagule pressure and invasion history. *Biological Invasions*, **11**, 2401–2414.
- Essl F, Dullinger S, Kleinbauer I (2009) Changes in the spatio-temporal patterns and habitat preferences of *Ambrosia artemisiifolia* during its invasion of Austria. *Preslia*, **81**, 119–133.
- Fumanal B, Gaudot I, Bretagnolle F (2008) Seed-bank dynamics in the invasive plant, *Ambrosia artemisiifolia* L. *Seed Science Research*, **18**, 101–114.
- Gallien L, Münkemüller T, Albert CH, Boulangeat I, Thuiller W (2010) Predicting potential distributions of invasive species: where to go from here? *Diversity and Distributions*, **16**, 331–342.
- Gaudeul M, Giraud T, Kiss L, Shykoff JA (2011) Nuclear and chloroplast microsatellites show multiple introductions in the worldwide invasion history of common ragweed, *Ambrosia artemisiifolia*. *PLoS One*, **6**, e17658.
- Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.
- Hodgins KA, Rieseberg L (2011) Genetic differentiation in life-history traits of introduced and native common ragweed (*Ambrosia artemisiifolia*) populations. *Journal of Evolutionary Biology*, **24**, 2731–2749.
- Inouye DW (2008) Effects of climate change on phenology, frost damage, and floral abundance of montane wildflowers. *Ecology*, **89**, 353–362.
- Johns T, Gregory J, Ingram W *et al.* (2003) Anthropogenic climate change for 1860 to 2100 simulated with the HadCM3 model under updated emissions scenarios. *Climate Dynamics*, **20**, 583–612.
- Kearney M, Porter W (2009) Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. *Ecology Letters*, **12**, 334–350.
- Kearney M, Phillips BL, Tracy CR, Christian KA, Betts G, Porter WP (2008) Modelling species distributions without using species distributions: the cane toad in Australia under current and future climates. *Ecography*, **31**, 423–434.
- Kelly AE, Goulden ML (2008) Rapid shifts in plant distribution with recent climate change. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 11823–11826.
- Lavoie C, Jodoin Y, De Merlis AG (2007) How did common ragweed (*Ambrosia artemisiifolia* L.) spread in Quebec? A historical analysis using herbarium records. *Journal of Biogeography*, **34**, 1751–1761.
- Lenoir J, Gégout JC, Marquet PA, De Ruffray P, Brisse H (2008) A significant upward shift in plant species optimum elevation during the 20th century. *Science*, **320**, 1768–1771.
- Menzel A, Sparks TH, Estrella N *et al.* (2006) European phenological response to climate change matches the warming pattern. *Global Change Biology*, **12**, 1969–1976.
- Morin X, Augspurger C, Chuine I (2007) Process-based modeling of species' distributions: what limits temperate tree species' range boundaries? *Ecology*, **88**, 2280–2291.
- Oswalt M, Marshall G (2008) Ragweed as an example of worldwide allergen expansion. *Allergy, Asthma & Clinical Immunology*, **4**, 130–135.
- Petitpierre B, Kueffer C, Broennimann O, Randin C, Daehler C, Guisan A (2012) Climatic niche shifts are rare among terrestrial plant invaders. *Science*, **335**, 1344–1348.
- Rich TCG (1994) Ragweeds (*Ambrosia* L.) in Britain. *Grana*, **33**, 38–43.
- Saar M, Gudinskas Z, Plompuu T, Linno E, Minkien Z, Motiekaityt V (2000) Ragweed plants and airborne pollen in the Baltic states. *Aerobiologia*, **16**, 101–106.
- Sauliune I, Gudinskas Z, Veriankaite L, Malciute A, Lesciauskiene V (2011) Distribution of *Ambrosia* plants and airborne pollen in Lithuania. *Journal of Food Agriculture & Environment*, **9**, 547–550.
- Sherry RA, Zhou X, Gu S *et al.* (2007) Divergence of reproductive phenology under climate warming. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 198–202.
- Shrestha A, Erivelton SR, Thomas AG, Swanton CJ (1999) Modeling germination and shoot-radicle elongation of *Ambrosia artemisiifolia*. *Weed Science*, **47**, 557–562.

- Smolik MG, Dullinger S, Essl F *et al.* (2010) Integrating species distribution models and interacting particle systems to predict the spread of an invasive alien plant. *Journal of Biogeography*, **37**, 411–422.
- Thuiller W, Lavorel S, Araújo MB, Sykes MT, Prentice IC (2005) Climate change threats to plant diversity in Europe. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 8245–8250.
- USA National Phenology Network (2013) *Plant Phenology Data for the United States*. Tucson, Arizona, USA: USA-NPN. Available at <http://www.usanpn.org/results/data> (accessed 28 January 2013).
- Willemssen RW (1975) Effect of stratification temperature and germination temperature on germination and the induction of secondary dormancy in common ragweed seeds. *American Journal of Botany*, **62**, 1–5.
- Wood SN (2003) Thin plate regression splines. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **65**, 95–114.
- Wood S (2013) mgcv v1.7-26: Mixed GAM Computation Vehicle with GCV/AIC/REML smoothness estimation.
- Yin X, Kropff MJ (1996) The effect of temperature on leaf appearance in rice. *Annals of Botany*, **77**, 215–221.
- Ziska L, Knowlton K, Rogers C *et al.* (2011) Recent warming by latitude associated with increased length of ragweed pollen season in central North America. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 4248–4251.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1. Additional detail on parameterization of the ragweed phenology model.

Appendix S2. Information sources for the distribution of *Ambrosia artemisiifolia* in North America and Europe.

Appendix S3. Validity of using long-term average climate data in the model.