Modelling the Spatial Invasive Range of *Heracleum mantegazzianum* in Europe

Wallentin, G.
Director of Studies, UNIGIS MSc Salzburg
Interfaculty Department of Geoinformatics – Z_GIS, University of Salzburg, Hellbrunnerstrasse 34,
5020 Salzburg, E-mail: gudrun.wallentin@sbg.ac.at

Abstract

The Western Caucasian plant Heracleum mantegazzianum (Giant Hogweed) is invading various parts of Europe at fast pace. This invasion is problematic, because the phytotoxic sap of Heracleum mantegazzianum can cause severe inflammations in human skin. To effectively take management actions, the potential spatial distribution of Giant Hogweed in Europe needs to be assessed. Usually, potential distribution is modelled based on parameters of the ecological niche as realised in a species native range. However, from the western Caucasus no observation data are available. This research investigates whether an environmental envelope model can reliably project the potential distribution of an invasive plant with observations from the invaded range only. Historic invaded-range observations in Europe were used to produce a time series of potential distribution maps. The results showed that the modelled area of potential distribution initially grew larger as new observations became available over time, until the area consolidated despite adding new observations. It is argued that this "point of consolidation" indicates the point in time, from when it is valid to project potential ranges of invasion with invaded-range observations only. For Giant Hogweed the potential range of invasion covers most parts of Europe and Central Russia.

1. Introduction

Heracleum mantegazzianum (Giant Hogweed) is a fast spreading invasive plant species in Europe. Its toxic furanocoumarins in the sap can cause severe skin inflammations (Phytophotodermatitis) when exposed to sunlight (Camm et al., 1976). Besides this health threat, conservation concerns have been expressed especially for riparian ecosystems (Page et al., 2006), where *Heracleum mantegazzianum* can displace native vegetation (Figure 1). Its further habitat preference for widely abundant ruderal places and abandoned grasslands is less critical from a conservation perspective (Thiele and Otte, 2006), but important for its successful spread. Giant Hogweed is a plant that grows to heights of 2-5m. It was introduced into England as an attraction in landscape gardens, and appeared on the seed list at London Kew Botanic Gardens in 1817. Ten years later the first population was recorded growing wild in Cambridgeshire (Pyšek et al., 2008). From there it invaded Central Europe (Nielsen et al., 2005 and Jahodová et al., 2007) and its range is still expanding (Nehrbass and Winkler, 2007).

Figure 1: A riparian stand of *Heracleum mantegazzianum* in Scotland

This brings up the question, which parts of Europe still are prone to invasion, and where the natural limits of distribution are. In Europe a few studies on distribution ranges of *Heracleum mantegazzianum* focused on the national level: Nielsen et al. (2008) modelled the current distribution in Denmark, Pyšek et al. (2008) studied the history of invasion in the Czech Republic on different scales and Catterall et al. (2012) used a Bayesian modelling approach for the last decades of invasion in the UK. The potential distribution on a regional scale for the entire continent of Europe has not yet been investigated. Models of potential distribution build on the concept of ecological niches, which represent the 'environmental envelope' of conditions required by a species. Using a statistical algorithm, observed occurrence locations and continuous environmental data layers are combined to project the potential distribution range. The condition implied in this approach is that the species is in an equilibrium state with the environment, i.e. that the realised niche equals the potential niche of environmental conditions. This presents a severe problem for modelling the potential distribution of invasive species. Therefore, Guisan and Thuiller (2005) suggest for invasive species to preferably use observation data from the native range. However, for some species native range observations are not available without costly field studies. The native range of *Heracleum mantegazzianum* for instance is in the subalpine zone of the Western Caucasus Mountains of Georgia, Azerbaijan, and southern Russia, where it is found in meadows, clearings, and forest edges between 1500 and 1850 m (Page et al., 2006 and Jahodová et al., 2007). Despite an increasing number of studies on Giant Hogweed in Europe, there are no observation data available and little is known about the biology of this species in its native range (Pergl et al., 2006). This research investigates how well an environmental envelope modelling approach can project the potential distribution of invasive species based on observation data from the invaded range only. Using *Heracleum mantegazzianum* as an example, the objectives of this research are (1) to analyse how the extent of the projected distribution area increases as new observation points are added over time, (2) to determine at what point the projected area consolidates despite new observations are added, and (3) to evaluate whether this consolidation point plausibly represents the potential distribution of the invasive range.

2. Methods

Historic observations of *Heracleum mantegazzianum* in Europe were used to analyse the projected distribution at several time steps to assess the modelled distribution in dependence with the available data at these points in time. The species distribution was modelled with the Open Modeller software (Muñoz et al., 2009) using the Bioclim approach (Nix, 1986) that implements the Bioclimatic Envelope Algorithm. This algorithm takes environmental information from a user-defined set of layers for each observation point and calculates the mean and the standard deviation to define the environmental envelope. Each grid cell is then classified as suitable, marginal or unsuitable. Whereas other algorithms that rely on more advanced statistical methods (e.g. PCA, or Bayesian approaches) perform better with small datasets or systematically sampled temporal snapshots (Hernandez et al., 2006, Catterall et al., 2012), the Bioclim algorithm represents the classical implementation of the environmental niche concept as discussed in this paper. For the environmental information, 21 data layers of the present climate with a resolution of 10' (New et al., 2002) were accessed through the Open Modeller library, including mean annual temperature, mean number of frost days, mean precipitation, etc. The observation data were accessed through the Global Biodiversity Information Facility (GBIF). In total, more than13,000 observed locations from 144 data providers worldwide were retrieved. About half of the data records also held temporal information. The first observations were recorded in 1913, the most current ones in 2012, where more than 60% of observations had been observed after the year 2000. The approximate native range of *Heracleum mantegazzianum* was digitised from Jahodová et al. (2007). The observation data was filtered to provide the information available in 1950 and each decade thereafter until 2010. Potential distribution maps were calculated for each of these time steps and additionally for the entire present day dataset including data without collection date. Finally, the resulting time series of distribution maps was analysed with respect to the available observation data to find the supposed 'point of consolidation' at which the full potential extent of the invaded range is projected, regardless of new observation points being added.

3. Results

In Figure 2 the time series of distribution maps shows the projected distribution range of *Heracleum*

mantegazzianum computed for the available data at each modelled time step. The observations document the invasion over time, as the plant increasingly occupies the potential ecological niche in Europe until the entire niche is realised and the area of the modelled potential distribution consolidates. Although the observation data is concentrated in Central Europe, the ecological niche model predicts a potential distribution over almost entire Europe, from northern Spain to Moscow and from the Balkan region to the north of Scandinavia. To quantify the point at which the projected distribution area consolidates, the number of observations is plotted against the area for each time step (Figure 3).

a) until 1950   b) until 1960   c) until 1970   d) until 1980

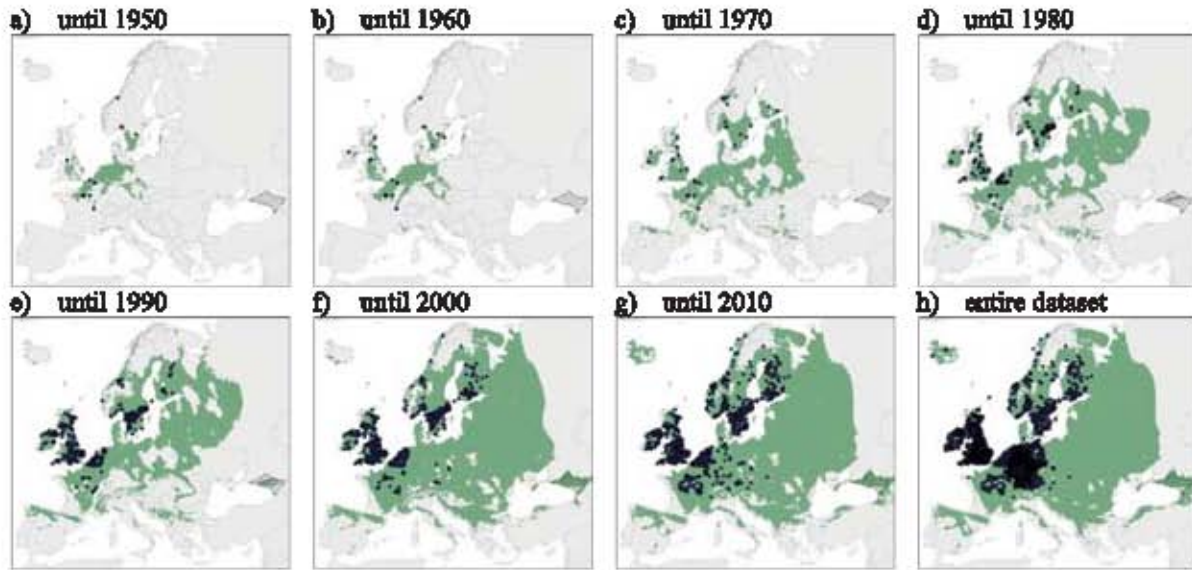e) until 1990   f) until 2000   g) until 2010   h) entire dataset

Figure 2: Potential distribution area based on historic observation data: the time series shows, how observation data (blue dots) affect the modelled extent of potential distribution of *Heracleum mantegazzianum* (green). The darker grey area is the native range in the western Caucasus region.

The extent of modelled habitat saturates with the number of observations
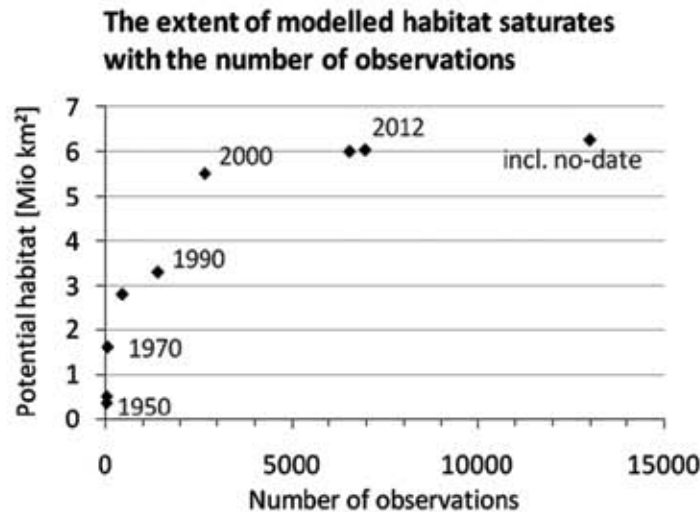
Figure 3: The number of observation points for each time step plotted against the area of the respective potential distribution shows a saturation curve: since the year 2000 the number of observations increased about five times without resulting in a significant change of potential distribution area.

The year 2000 is the turning point towards saturation. Although current observation data (including undated observations) outnumber the 2000 data five times, the projected distribution area only increases by 13%, mainly by filling holes within the potential range (Figure 2f-h). Since 2000, valid maps of the potential distribution can be computed with the available data. The reliability of the model since the year 2000 is confirmed by the good prediction of the native range based on observations from the invaded range in Europe only, although the two areas are geographically remote. In the 1980 projection, 17% of the native Caucasian habitat is identified as suitable, and since the year 2000 over 80% of the native range identified by Jahodová et al., (2007) is correctly classified as suitable habitat. As an alternative indicator to determine the consolidation point, the distance of observation points to the modelled distribution edge was analysed. When the environmental envelope from sparse sample points is projected to geographic space, a high fraction of these points are located near the outer distribution boundaries: the observation points encircle the projected distribution area. In 1950 only 27 observation points generate an area extending over app. 200 x 900km in northern Germany, the Benelux countries and parts of France (Figure 2a, b). The median distance of an observation point to the nearest distribution edge is only 2,5km. For the 1950 and 1960 data, more than 20% of observation points are closer than 1km to the modelled distribution edge (Figure 4).
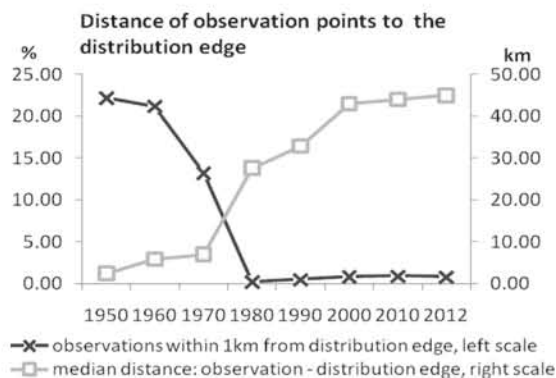

Figure 4: As new observations are added, the fraction of observations within 1km of the modelled distribution edge diminishes and the median distance of observation points to the outer boundary of the modelled distribution grow. These distance measures correlate with the number of observation points, but cannot be used as indicator of a 'consolidated' model result

In 1980 available observations points increased significantly, from 53 in 1970 to 442. Since then less than 1% of points are nearer than 1km to the edge. However, the modelled distribution still increased over the next two decades until 2000. So distance of observation points to the projected distribution edge is not an adequate predictor for a model to deliver consolidated results.

4. Discussion and Conclusions

Heracleum mantegazzianum potentially can invade most parts of Europe. For the eastern parts of the projected range in the Ukraine, Belarus and Russia no observation point data were accessible through GBIF. However, the literature confirms that Giant Hogweed is an invasive species in Central Russia (Pergl et al., 2012), which further validates the model. The ecological niche model of *Heracleum mantegazzianum* resulted in a plausible area of its potential geographical distribution, although the model was based on observation points in the invaded range only. This result does not oppose to the suggestion of preferably using native range observations to project the invaded range (Guisan and Thuiller, 2005) nor is it in contrast to the postulated better accuracy of distribution models based on systematically sampled presence-absence data (Brotons et al., 2004). However, this research shows that using unstructured samples of presence-only observation data of the invaded range can result in plausible distribution maps. In this sense, the finding is of high relevance for making use of the rapidly increasing accessibility of species data not only from environmental monitoring facilities, but also crowd-sourcing platforms (Jetz et al., 2012). What needs yet to be done is to find a computational approach to determine the 'point of consolidation', at which additional observations do not significantly influence the modelled distribution area. Plotting the modelled distribution area against observation points requires a large amount of observation data beyond the saturation point. For the projection itself these additional data are not needed. In this research, distance measures between observation points and distribution edges were analysed as a less data hungry alternative. Although the median distances grow as the model consolidates, distance measures proved not to be suitable for clearly identifying the consolidation point. An efficient way to determine the point of consolidation therefore remains to be a challenge for further research.

References

Brotons, L., Thuiller, W., Araújo, M. B. and Hirzel, A. H., 2004, Presence Absence Versus Presence only Modelling Methods for Predicting Bird Habitat Suitability. *Ecography, 27,* 437-448.

Camm, E., Buck, H. W. L. and Mitchell, J. C., 1976, Phytophotodermatitis from Heracleum mantegazzianum. *Contact Dermatitis,* 2, 68-72.

Catterall, S., Cook, A. R., Marion, G., Butler, A. and Hulme, P. E., 2012, Accounting for Uncertainty in Colonisation Times: A Novel Approach to Modelling the Spatio-Temporal Dynamics of Alien Invasions using Distribution Data. *Ecography,* 35, 901-911.

Guisan, A., and Thuiller, W., 2005, Predicting Species Distribution: Offering more than Simple Habitat Models. *Ecology Letters,* 8, 993-1009.

Hernandez, P. A., Graham, C. H., Master, L. L. and Albert, D. L., 2006, The Effect of Sample Size and Species Characteristics on Performance of Different Species Distribution Modeling Methods. *Ecography,* 29, 773-785.

Jahodová, Š., Fröberg, L., Pyšek, P., Geltman, D., Trybush, S. and Karp, A., 2007, Taxonomy, Identification, Genetic Relationships and Distribution of Large Heracleum Species in Europe. In Ecology and Management of Giant Hogweed (Heracleum mantegazzianum), edited by M. J. W. C. P.Pyšek, W. Nentwig and H. P. Ravn (Wallingford, UK: CAB International).

Jetz, W., Mcpherson, J. M. and Guralnick, R. P., 2012, Integrating Biodiversity Distribution Knowledge: Toward a Global Map of life. *Trends in Ecology and Evolution,* 27, 151-159.

Muñoz, M. E. S., Giovanni, R., Siqueira, M. F., Sutton, T., Brewer, P., Pereira, R. S., Canhos, D. a. L. and Canhos, V. P., 2009, Open Modeller: A Generic Approach to Species' Potential Distribution Modelling. *GeoInformatica.*

Nehrbass, N. and Winkler, E., 2007, Is the Giant Hogweed still a Threat? An Individual Based Modelling Approach for Local Invasion Dynamics of Heracleum mantegazzianum. *Ecological Modelling,* 201, 377-384.

New, M., Lister, D., Hulme, M. and Makin, I., 2002, A High Resolution Data set of Surface Climate over Global Land Areas. *Climate Research,* 21.

Nielsen, C., Hartvig, P. and Kollmann, J., 2008, Predicting the Distribution of the Invasive Alien Heracleum mantegazzianum at Two Different Spatial Scales. *Diversity and Distributions,* 14, 307-317.

Nix, H. A., 1986, A Biogeographic Analysis of Australian Elapid Snakes. In: R. Longmore (ed.) Atlas of Elapid Snakes of Australia. Canberra: Australian Government Publishing Service.

Page, N. A., Wall, R. E., Darbyshire, S. J., Mulligan, G. A., 2006, The Biology of Invasive Alien Plants in Canada. 4. Heracleum mantegazzianum Sommier & Levier. *Canadian Journal of Plant Science,* 86, 569–589.

Pergl, J., Perglová, I., Pyšek, P. and Dietz, H., 2006, Population Age Structure and Reproductive Behaviour of the Monocarpic Perennial Heracleum mantegazzianum (Apiaceae) in its Native and Invaded Distribution Ranges. *American Journal of Botany,* 93, 1018-1028.

Pergl, J., Pyšek, P., Perglová, I. and Jarošík, V., 2012, Low Persistence of a Monocarpic Invasive Plant in Historical Sites Biases Our Perception of its Actual Distribution. *Journal of Biogeography,* 39, 1293-1302.

Pyšek, P., Jarošík, V., Müllerová, J., Pergl, J. and Wild, J., 2008, Comparing the Rate of Invasion by Heracleum mantegazzianum at Continental, Regional, and Local Scales. *Diversity and Distributions,* 14, 355-363.

Thiele, J. and Otte, A., 2006, Analysis of Habitats and Communities Invaded by Heracleum mantegazzianum Somm. et Lev. (Giant Hogweed) in Germany. *Phytocoenologia,* 36, 281-320.