LUMA-GIS Thesis nr 26

Evaluating habitat suitability and spectral heterogeneity models to predict weed species presence

Christina Fotiou

2014
Department of
Physical Geography and Ecosystem Science
Centre for Geographical Information Systems
Lund University
Sölvegatan 12
S-223 62 Lund
Sweden

Christina Fotiou (2014). Evaluating habitat suitability and spectral heterogeneity models to predict weed species presence.
Master degree thesis, 30 credits in Master in Geographical Information Sciences
Department of Physical Geography and Ecosystems Science, Lund University

Evaluating habitat suitability and spectral heterogeneity models to predict weed species presence

---

Christina Fotiou

Master degree thesis in

Physical Geography and Ecosystem Science

Supervisor

Karin Hall

Department of Physical Geography and Ecosystem Science

Lund University, Sweden, 2014

Table of Contents

List of Figures

List of Tables

Abstract

The aim of this study is to analyze whether the presence of weed species resistant to herbicides is related to spatial-landscape heterogeneity. Additionally, Habitat Suitability (HS) models are tested to evaluate their presence. Weed species, *Chenopodium album*, *Conyza canadensis*, *Amaranthus retroflexus, Papaver rhoeas*, *Sorghum halepense*, *Echinochloa crus-galli* and *Lolium rigidum*, most of them common in two biogeographical regions of the European area, the Mediterranean, and the Continental, are used for the study. The corresponding countries selected, fitted in the differing bioclimatic conditions, are Greece and Germany. Given the difficulties of field-based data collection, and the availability of satellite-derived data, such as those derived from the Landsat, the use of remote sensing for estimating spatial heterogeneity is regarded as a powerful tool. Hence, the Normal Difference Vegetation Index (NDVI) is used as a data source for comparing species presence and spatial heterogeneity. In order to test relationships of species presence with NDVI, a number of regression models, such as Ordinary Least Square (OLS) and Logistic regression are tested. Further, Rank - abundance diagrams and Variograms of NDVI values are computed to detect spectral and hence spatial heterogeneity around species. Concerning HS models, since many available datasets do not provide reliable information about species absences, a model that is used to evaluate Habitat Suitability, by using presence only data is BIOMAPPER. Finally, so as to make predictions of species future distribution, based on bioclimatic data, OpenModeller software and GARP analysis are used. Most of the landscape area around species observation points is highly heterogeneous. The highest spectral variability is recorded around *P. rhoeas*, according to Variograms. Moreover, OLS regression results were not statistically significant. However, according to Logistic regression method, most of the species studied are related with low NDVI values, except from *P. rhoeas* in Germany related strongly with medium and dense vegetated areas. Studying species marginality through HS models, almost all weed species neither prefer to live in a wide range of conditions nor to be specialists in their environment, with the exception of *P. rhoeas* that tends to live in less marginal habitats. Additionally, the core factors affecting species habitat in Greece are mainly mean temperature, and secondarily annual precipitation, whereas the main factor is altitude for species in Germany followed by bioclimatic conditions. Concerning, species distribution in future climatic conditions, there is a general tendency of species moving to higher altitudes, minimizing their population size. However, in Germany, C_3 summer annual weeds, such as *C. album* and *C. canadensis* seem to be tolerant to climatic changes, and expand to new territories. Commonly in Greece, the distribution of winter annual C_3 weeds, such as *L. rigidum* and *P. rhoeas* is promoted by climatic changes. Concluding, the response of species range due to climatic changes, seems to be species-specific, but also related to

bioclimatic factors. Further study is needed to come up with prediction models of weed species future distribution.

Keywords: GIS; NDVI; Regression Analysis; Habitat Suitability-Biomapper; Spectral-Spatial Heterogeneity models; GARP models; Herbicide Resistant Weeds; Climate Change

1. Introduction

1.1 Background

The aim of the study is to exploit methods of relating species presence with landscape heterogeneity and environmental variables using remote sensing data. For this study, the Normalized Difference Vegetation Index (NDVI) was used. The novelty of this study is that this is the first attempt to relate NDVI with species presence.

A few attempts have been made in other studies to relate NDVI with species abundance or richness. Seto et al. (2004) reported the potential of predicting species richness using single-date Normalized Difference Vegetation Index (NDVI) derived from Landsat Thematic Mapper (TM). They used NDVI as an indicator of vegetation productivity, and examine the relationship of three measures of NDVI—mean, maximum, and standard deviation—with patterns of bird and butterfly species richness at various spatial scales. Results of this study, indicate a positive correlation, but with no definitive functional form, between species richness and productivity. Moreover, Oindo and Skidmore (2002), examined the relationship between interannual maximum NDVI parameters and species richness of vascular plants and mammals. Statistical analyses revealed that higher average NDVI, results in lower species richness. Further studies have also been done to relate NDVI with plant species richness, and the results were positive (Gillespie 2006; Levin et al. 2007; Parviainen et al. 2009). Finally, a number of researchers have used NDVI to model the interdependence of spectral heterogeneity and NDVI (Garrigues et al. 2006; Rocchini and Vannini 2010; Hall et al. 2012). Therefore, part of this study attempts to create spatial models relating spatial and spectral heterogeneity with species presence, using NDVI. For this reason, a number of different models, such as Habitat Suitability (HS) models, Rank - abundance diagrams, Spatial Variograms, Weight of Evidence (WofE), Ordinary Least Square models and Logistic Regression models were applied and evaluated.

The species used for this study were weeds resistant to herbicides. Evaluating the effect of environmental and bioclimatic conditions on these weed species is another novelty of the study. Their distribution was analyzed into two differing biogeographical areas Greece and Germany, characterized by Mediterranean and Continental climate, correspondingly. Despite the fact that these two regions have different bioclimatic conditions, most of the species studied are found in both countries.

A last attempt of this thesis was to asses weed risk expansion owing to changing climate. Therefore, projections of the climatic conditions of 2020 and 2050 were applied and the prediction of species distribution in future conditions was also studied by applying Genetic Algorithm for Rule-Set Production (GARP) models.

1.1.1 Weeds resistant to herbicides

The species studied are weeds resistant to herbicides. The evolution of weed populations resistant to herbicides is an increasing problem in many countries. Herbicide resistance can be defined as “the inherent ability of a weed to survive a rate of herbicide which would normally result in effective control” (Moss 2002). Resistance is an evolutionary process, whereby a population changes from being susceptible to being resistant. The process does not happen to individuals, but to populations, when the proportion of resistant individuals within the population increases over time (Moss 2002). The main factors influencing the resistant are agronomics. More specifically, herbicide resistance of weeds have occurred in cases where the same herbicides (or herbicides with the same mode of action) have been used repeatedly over a period of years, usually associated with intense agricultural or horticultural systems involving crop monoculture and minimum tillage systems, in which herbicides have been relied upon to achieve high levels of weed control. Herbicide resistance has been evolved within crops. The most severe and widespread problems have occurred in maize, cereal and rice crops. Additionally, herbicide resistance has also evolved in non-cropping situations, such as roadsides and railways, following intensive use of herbicides (Moss 2002).

1.1.2 C_3 and C_4 photosynthetic pathways and plants

In all plants CO_2 is fixed by the enzyme Rubisco. It catalyzes the carboxylation of ribulose-1,5-bisphosphate, leading to two molecules of 3-phosphoglycerate. Instead of CO_2, Rubisco can also add oxygen to ribulose-1,5-bisphosphate, resulting in one molecule each of 3-phosphoglycerate and 2-phosphoglycolate. Phosphoglycolate has no known metabolic purpose and in higher concentrations it is toxic for the plant (Anderson 1971). It therefore has to be processed in a metabolic pathway called photorespiration. Photorespiration is not only energy demanding, but furthermore leads to a net loss of CO_2. Thus the efficiency of photosynthesis can be decreased by 40% under unfavorable conditions including high temperatures and dryness (Ehleringer et al. 1991). Plants developed different ways to cope with this problem. Perhaps the most successful solution was C_4 photosynthesis (Gowik and Westhoff 2011).

The establishment of C_4 photosynthesis includes several biochemical and anatomical modifications that allow plants with this photosynthetic pathway to concentrate CO_2 at the site of Rubisco. Thereby its oxygenase reaction and the following photorespiratory pathway are largely repressed in C_4 plants. In most C_4 plants the CO_2 concentration mechanism is achieved by a division of labor between two distinct, specialized leaf cell types, the mesophyll and the bundle sheath cells, although in some species C_4 photosynthesis functions within individual cells (Edwards et al, 2004). Since Rubisco can operate under high CO_2 concentrations in the bundle sheath cells, it works more efficiently than in C_3 plants (Oaks 1994). Additionally, C_4 plants exhibit better water-use efficiency than C_3 plants. Because of the CO_2 concentration mechanism they can acquire enough CO_2 even when keeping their stomata more closed. Thus, water loss by transpiration is reduced (Long 1999). Therefore, the C_4 photosynthetic carbon cycle is an elaborated addition to the C_3 photosynthetic pathway. It evolved as an adaptation to high light intensities, high temperatures, and dryness (Edwards et al. 2010).

1.2 Problem statement

Weed species resistant to herbicides threaten agricultures, ecosystem and human health. Resistance to herbicides worries agriculturers, since it has important economic and environmental consequences to agriculture. Weed herbicide resistance could narrow agriculture production, since weeds are still growing at the expense of cultivated crops. Higher volumes of herbicides are dropped on the agricultures, causing nutritional problems, and also turning more weed species resistant. Effective herbicide weed management requires an assessment of their distribution, related with the landscape features, environmental and climatic conditions.

Given the difficulties of field-based data collection, the use of remote sensing for relating weed species distribution with landscape and environmental conditions could be a powerful tool for weed risk expansion assessment, regarding the availability of satellite-derived data. An index that is extensively used to study these relationships is the Normal Difference Vegetation Index.

Measurements of environmental variables that are key determinants of species distributions, such as land use, and soil conditions, are also important for defining species habitats or species spatial range. Moreover, the global climate is changing, driving also changes in the flora. Many weed species with high phenotypic plasticity and tolerant to changes, could possibly adapt to climatic pressures, and expand their range, whereas others could not acclimate and, get extinct. Hence, the assessment of weed risk expansion owing to changing climate is another task to be examined in this study.

1.3 Objectives

In this study the questions of the work are described as follows:

- Is landscape heterogeneity related to species presence? Does the pattern differ between the studied regions?
- What about species, environmental and bioclimatic conditions? Is there any relationship?
- Which main factors are related with species presence and distribution?
- Can we make assumptions of future species distribution?

Finally, comparisons of the Eco-geographical models applied in this study are made.

2. Background

2.1 Study area and climatic conditions

Two study areas were selected laying in different biogeographic regions, Greece (39.00° N, 22.00° E), a Mediterranean country, and Germany (51.00° N, 10.00° E) a Continental country. The Mediterranean climate is characterized by warm and dry summers, mild and wet winters (fig. 1). Greece is located in South Europe, on the southern end of the Balkan Peninsula, consisting of a large mainly mountainous area and a large number of islands (Worldatlas 2014a).

Germany is a west-central European country, stretches from the Alps to the North and Baltic sea. The climate is temperate (cool and wet winters and summers) at the greater part of Germany and marine/oceanic in the North and Northwest part with relatively mild winters, cool summers and rainfall all year round. In the East the climate is Continental, warm summers and mild cloudy winters (Worldatlas 2014b).

2.2 Climate change in Europe

It is generally accepted that the earth is getting warmer with significant ecological consequences (Walther et al. 2002; Thuiller et al. 2005). Recently, this century, Central and Northern Europe have received more rain than in the past. In contrast, Southern and Southeastern Europe have become drier. There have also been changes in the flora. During the last thirty years the number of plant species has decreased in parts of Europe, which is attributed to loss of habitat and climate change (Mannetje 2006).

Mannetje (2006) refers that extreme cold winters will no longer occur in the Northern Hemisphere towards the end of the present century, whilst extreme warm summers and heavy showers will occur more frequently. Agriculture in Europe will profit with increased yields as a result of rises in temperature and atmospheric CO_2 and by expansion in a northerly direction and in central Europe. However, at the same time in parts of Southern Europe harvests will become more insecure because of water shortages due to reduced summer rainfall by about 20%. The intensity of rain showers is likely to increase and may become more tropical in nature. The higher temperatures will lead to increased evaporation that together with reduced precipitation will lead to more frequent droughts (Mannetje 2006). The mean surface air temperature is predicted to increase by 1.5 to 5.8 °C by the end of the present century. The temperature in the northern latitudes is expected to rise more than elsewhere (IPPC 2001).

Concerning the climate of the Mediterranean region, it would become hotter, drier and more variable. Annual mean temperature would increase by 1-2 °C over present conditions, but at inland locations, such as in Turkey, northern Italy, at a distance from the sea, maximum temperatures could rise by up to 5 °C. Annual precipitation would likely decrease by up to 20% over the southern Mediterranean, while reduction in summer rainfall over the northern Mediterranean could exceed 30%. Drought periods would be expected to shift in time and last longer. The hotter and drier climate would lead to lower agricultural yields, particularly in summer crops that are not irrigated. A drier climate, accompanied by reduced precipitation and surface runoff, and increasing

demand from the agricultural sector, would exacerbate the already high level of water stress in the region. Latest studies show that global warming of more than 2 °C could lead to a loss of over 50% of plant species in the Northern Mediterranean region (Mannetje 2006).

2.3 Remote sensing

2.3.1 What is Remote sensing?

Remote sensing is a technique of observing/recording objects from a a distance, by using sensors based on ships, aircraft, satellites, or other spacecraft, i.e. airborne or spaceborne platforms. Optical Remote Sensors detect solar radiation, in differing wavelengths, usually extending from the visible and near infrared (NIR) to the short-wave infrared (SWIR), scattered or reflected from surfaces from the earth. The radiation collected by the sensors is used to form images depicting the area or object from which it was reflected. Different objects, such as water, vegetation, soil, buildings, etc. reflect visible and NIR light in different ways, resulting in different spectral reflectance signatures (Sanderson 2012). The graph in figure 1 shows the typical reflectance spectral of water, bare soil and two types of vegetation.

Moreover, There are two kinds of remote sensors, passive sensors that collect radiation reflected or emitted from an object, and active sensors that transmit a signal and collect its reflectance (CRISP 2013). The quality of remote sensing data consists of its spatial, spectral, and radiometric resolutions (Janssen and Huurneman 2001).

Spatial resolution

The size of a pixel that is recorded in a raster image.

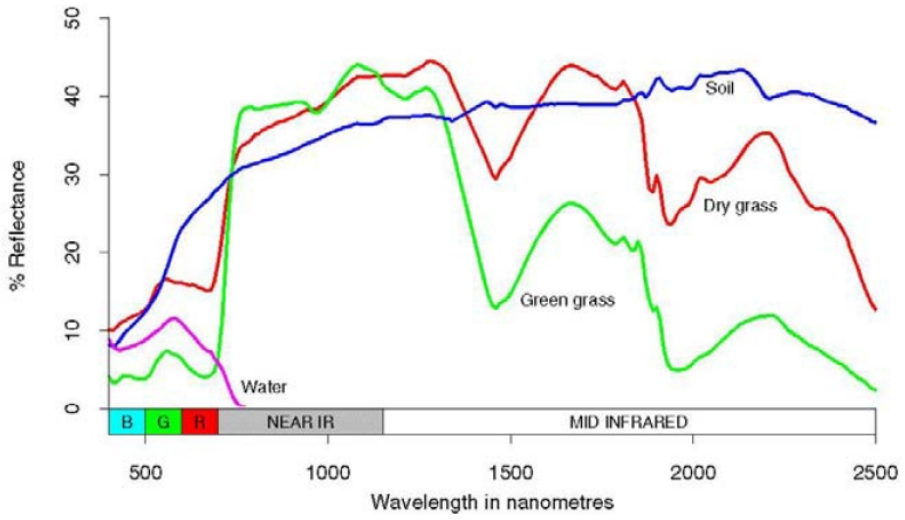


Figure 1. Typical spectral reflectance of water, bare soil and two types of vegetation.

Spectral resolution

The recorded frequency bands' wavelength width, usually related to the number of frequency bands recorded by the platform.

Radiometric resolution

The number of different radiation intensities the sensor is able to distinguish. Typically, this ranges from 8 to 14 bits, corresponding to 256 levels of the gray scale and up to 16,384 intensities or "shades" of color, in each band. It also depends on the noise of the instrument.

Radiometric correction

Gives a scale to the pixel values, e. g. the monochromatic scale of 0 to 255 will be transformed to actual radiance values.

Topographic correction

It eliminates the variations of the illumination depending on the exposure of a pixel. It recovers true reflectivity or radiance of objects in horizontal conditions.

Atmospheric correction

It eliminates atmospheric haze, rescaling each frequency band so that its minimum value (usually realized in water bodies) corresponds to a zero pixel value.

Finally, remote sensing map images have to be geo-referenced, since raster image has no particular size, to a geographic location, by the process of scaling, rotating, translating and de-skewing the image.

2.3.2 Landsat ETM+

Thematic mappers (TM) and multispectral scanners (MSS) are passive scanning systems that collect raster data in several bandwidths simultaneously between visible light and thermal bandwidths (0.4–8.0 μm).

Landsat represents the world's longest continuously acquired collection of space-based moderate-resolution land remote sensing data. It is a joint initiative between the U.S. Geological Survey (USGS) and NASA. The Landsat Project and the data it collects support government, commercial, industrial, civilian, military, and educational communities throughout the United States and worldwide. Landsat Enhanced Thematic Mapper Plus (ETM+) images consist of eight spectral bands with a spatial resolution of 30 meters for Bands 1 to 7 (table 1). The resolution for Band 8 (panchromatic) is 15 meters (USGS 2013a).

Finally, the Landsat Enhanced Thematic Mapper Plus (ETM+) sensor onboard the Landsat 7 satellite has acquired images of the Earth nearly continuously since July 1999, with a 16-day repeat cycle. Landsat 7 images are referenced to the Worldwide Reference System-2 (USGS 2013b).

Table 1. Spectral bands, wavelength and resolution of ETM +.

Enhanced Thematic Mapper Plus (ETM+)	Landsat 7	Wavelength (micrometers)	Resolution (meters)
	Band 1	0.45-0.52	30
	Band 2	0.52-0.60	30
	Band 3	0.63-0.69	30
	Band 4	0.77-0.90	30
	Band 5	1.55-1.75	30
	Band 6	10.40-12.50	60 * (30)
	Band 7	2.09-2.35	30
	Band 8	.52-.90	15

* ETM+ Band 6 is acquired at 60-meter resolution. Products processed after February 25, 2010 are resampled to 30-meter pixels.

2.3.3 Normalized Difference Vegetation Index (NDVI)

An Index that is widely used in Remote Sensing is the Normalized Difference Vegetation Index (NDVI) that is based on how plants reflect different wavelengths of the spectrum. Therefore, this Index is used to determine the density of green on a patch of land. Reflectance is the ratio of energy that is reflected from an object. Spectral reflectance of objects differs in the range of wavelength of the electromagnetic spectrum. Plant photosynthetic activity is determined by chlorophyll content. Chlorophyll in plants absorbs blue and red portion of the spectrum, i.e. visible light (from 0.4 to 0.7 μm), and reflects green. Near infrared radiant energy (NIR, from 0.7 to 1.1 μm) is strongly reflected from the plant surface (Kumar and Silva 1973). The contrast between vegetation and soil is at a maximum in the red and near infrared region. Therefore, spectral reflectance data can be used to compute a variety of vegetative indices. The Normal Difference Vegetation Index (NDVI) could be used as a successful predictor of photosynthetic activity, since it includes both NIR and red light wavelength.

The NDVI is calculated from reflectance measurements in the red and near infrared (NIR) portion of the spectrum (Eqn 1):

$$NDVI= (R_{NIR} –R)/(R_{NIR} +R) \quad \textit{(Eqn 1)},$$

where R_{NIR} is the reflectance of NIR radiation and R is the reflectance of visible red radiation.

Calculations of NDVI for a given pixel always result in a number that ranges from -1 to

+1. Negative values correspond to sea, lakes, rivers and oceans. Very low NDVI values (0 - 0.1) correspond to barren rock, sand, or snow. Sparse vegetation such as shrub and grasslands result in moderate NDVI values (0.2 - 0.5), whereas high NDVI values (0.6 - 0.9) correspond to dense vegetation. Generally, if the NIR in higher than the red radiation, then the vegetation is likely to be dense. On the contrary, the vegetation is sparse, when the difference in the reflected visible and NIR wavelength is minimized (Rouse et al. 1973).

2.4 Ecological & Geographical Modeling of Species Distribution in the Landscape

2.4.1 Ecological Niche

Ecological niche relates a set of environmental variables to species fitness, and according to Hutchinson (1957), is the volume in the environmental space that permits positive growth. The presence of a species in a location depends on three parameters (Soberon and Peterson 2005; Soberon 2007): (i) the local environment allows the population to grow (Grinnellian niche), (ii) the interactions with other local species (predation, competition, mutualism, etc.) allow the species to persist (Eltonian niche), and (iii) the location is actually accessible, given the dispersal abilities of the species. These constraints determine the geographical distribution of the species (Hirzel and Gwenaëlle 2008). Moreover, Hutchinson (1957) defined the Realized niche as a subset of the Fundamental niche a species was constrained to occupy because of interactions with other species. However, the difference between realized and fundamental niche are becoming lower, when spatial heterogeneity and dispersal limitations are considered (Pulliam 2000).

2.4.2 Habitat Suitability (HS) models

When Fallacious absences (FA) are suspected, Habitat Suitability models that use only presence data for the analysis should be preferred (Helfer and Métral 2001). Fallacious absences stem from at least five causes, namely: (i) limited dispersal: geographical barriers or slow dispersal prevent a species from occupying some parts of its potential distribution; (ii) local extinction: environmental or demographic stochasticity has momentarily driven a local population to extinction; (iii) patch size: the area of suitable habitat is too small to harbour a viable population; (iv) alternative habitats: a generalist species may use several types of habitat, possibly at different periods; (v) biotic interactions (e.g. succession stage, competition, predation).

Additionally, Habitat Suitability (HS) models relate ecogeographical variables (EGV's) to the likelihood of species occurrence, related on Hutchinson ecological niche theory. HS model analysis is based on ENFA, Environmental Factor Analysis, comparing in a multidimensional space of ecological variables the species distribution, proposing habitat suitability maps.

Crucial factors related to species distribution and their suitable habitats is global marginality, tolerance, and specialization (Hirtzel et al. 2002). Marginality refers to the impact of ecogeographical variables in the probability of species and population occurrence, especially derived from unfavorable conditions and resource scarcity

(Shreeve et al. 1996). The global marginality takes into account all the EGVs and provides a summary of how much the species habitat differs from the available conditions. A low value (close to 0) indicates that the species tends to live in average conditions, whereas a high value (close to 1), in extreme habitats throughout the study area. Moreover, a low value of the global tolerance (close to 0) indicates a "specialist" species ending to live in a very narrow range of conditions, whereas a high value (close to 1) indicates a generalist species. The global specialization is the inverse of global tolerance, but as it varies between 1 and infinity, it is less easy to interpret (Hirtzel et al. 2002).

2.4.3 Rank - Abundance diagrams & Landscape Heterogeneity

Rank-abundance diagrams could be used in remote sensing for evaluating both spectral richness and equitability. Former methods, used the richness of digital numbers (DNs), the Shannon entropy/Boltzmann (Shannon and Weaver 1962) or Pielou evenness (Pielou 1969) indices for measuring spatial heterogeneity. Shannon entropy is a diversity index indicating the uncertainty of the prediction of a community, whereas Pielou eveness is a measure that quantifies the equitability of the species community. Rank–abundance diagrams were first introduced by MacArthur (1957), who developed the theory with examples related to bird communities by plotting species abundance against ranks. Further, Whittaker (1965) straightforwardly reworked the method with a graphical representation of all possible theoretical situations of the community diversity of a species. Rank-abundance diagrams are showing the diversity of an array of values. Once the relative abundance of a species is ordered, a rank is assigned to each species as a function of its abundance (Rocchini and Neteler 2012).

3. Material and Methods

3.1 Flowchart

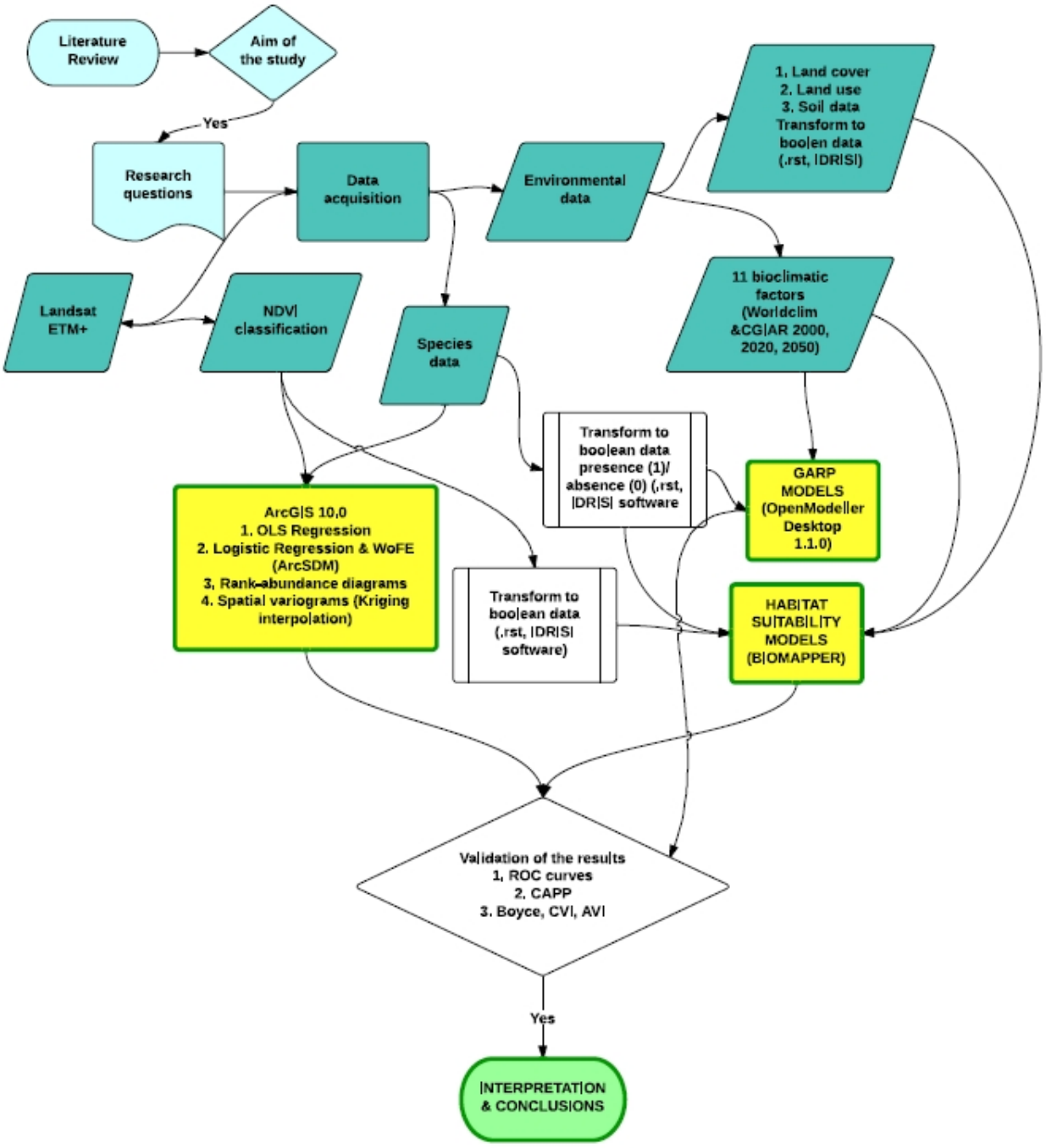


Figure 2. Master thesis research method flowchart.

According to the above flowchart (fig. 2), the first step for this study was to review the literature related to resistant to herbicide weed species and their response to differing environmental conditions. Moreover, I searched through the literature of models that can be applied to study weed species distribution in a variety of environmental and climatic conditions. A number of ecological models were studied and a few of them were selected. The next step was to set the aim of the study and the research questions to be answered. Further, a number of data that can be used to answer the questions setted, were acquired through internet, and downloaded free of charge. The data that were used, were grouped to species data, Landsat ETM+ data, environmental data, such as land cover, land use, soil data (transformed to Boolean data), and 11 bioclimatic factors (Worldclim and CGIAR for 2000, 2020 and 2050). The relationship of species presence with NDVI was studied by applying Ordinary Least Square (OLS) regression, Logistic regression and WofE (Weight of Evidence) using ArcGIS 10 and ArcSDM software. Moreover, this relationship was also examined by applying the Rank-abundance diagrams and spatial variograms spatial models. So as to study Habitat Suitability for each weed species, NDVI classified data and species data were turned to presence/absence Boolean data using IDRISI software, and were combined with the environmental and the 11 bioclimatic factors in the biomapper software. The projection of species distribution in 2000, 2020 and 2050 was performed by applying GARP models (openModeller Desktop 1.1.0 software), matching species Boolean data with the 11 bioclimatic factors.

3.2 Data

3.2.1 Species

An available reliable free species database is the Global Biodiversity Information Facility (GBIF 2013). GBIF was established by governments in 2001 to encourage free and open access to biodiversity data, via the Internet to promote scientific research, conservation and sustainable development. Most of the species data used for this study are common weeds found in both studied countries and they all share a common characteristic that is herbicide resistance (International Survey of Herbicide Resistant Weeds 2013). Most of the species collections dates range between 2000 and 2010. The data files were in excel format and inserted into ArcGIS 10 as point shapefiles. The weed species used for this study are shown in table 2.

3.2.2 Administrative boundaries

The data for the administrative boundaries of Greece and Germany were downloaded from GADM database of Global Administrative Areas (GADM 2013). GADM is a spatial database of the worlds's administrative areas for GIS and similar software use.

3.2.3 Bioclimatic data and altitude

Climate layers for both countries were downloaded from WorldClim (WorldClim 2013), a set of global climate layers (climate grids) and altitude grids with a spatial resolution of about 1 square kilometer. The data layers were derived from interpolation of average

monthly temperature and rainfall value data on a 30 arc-second resolution grid. For this study eleven bioclimatic factors were used coded as follows:
BIO1 = Annual Mean Temperature
BIO2 = Mean Diurnal Range (Mean of monthly (max temp - min temp))
BIO3 = Isothermality (BIO2/BIO7) (* 100)
BIO4 = Temperature Seasonality (standard deviation *100)
BIO5 = Max Temperature of Warmest Month
BIO6 = Min Temperature of Coldest Month
BIO7 = Temperature Annual Range (BIO5-BIO6)
BIO12 = Annual Precipitation
BIO13 = Precipitation of Wettest Month
BIO14 = Precipitation of Driest Month
BIO15 = Precipitation Seasonal

Moreover, altitude data were downloaded from WorldClim for both Greece and Germany. In order to make future predictions for species distributions the same bioclimatic factors were downloaded for 2020 and 2050 from GCM data portal (GCM 2013), data provided by the CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS). The parameters selected were delta method, SRES_a1B scenario, CSIRO_MK3_0 model with a resolution of 2.5 arc minutes.
Delta Method is a statistical downscaling method based on thin plate spline spatial interpolation of anomalies (deltas) of original Global Climate Model (GCM) outputs. Anomalies are interpolated between GCM cell centroids and are then applied to a baseline climate given by a high resolution surface. The method makes the following gross assumptions:
1. Changes in climates vary only over large distances (i.e. as large as GCM side cell size).
2. Relationships between variables in the baseline ('current climates') are likely to be maintained towards the future (CCAFS 2014).
Concerning CSIRO Mk3.0 model, a 12-month average temperature was calculated for 1961-1990 based on the simulated transient historical data. Result is a 12 month simulated climatology at the scale of the CSIRO Mk3.0 grid. For each forecast month a difference anomaly was calculated (ex. January_2021 – Mean_historical_January). Result is a 100 year monthly set of temperature anomalies at the scale of the CSIRO Mk3.0 grid (DATA BASIN 2014).
In this study, according to the IPCC Special Report on Emission Scenarios (SRES), the A1 storyline and Emissions Scenario family was used that describes a future world of very rapid economic growth, global population that peaks in mid-century and declines thereafter, and the rapid introduction of new and more efficient technologies. Major underlying themes are convergence among regions, capacity building and increased cultural and social interactions, with a substantial reduction in regional differences in per capita income. The A1 scenario family group that was applied in this study was the A1B or a balance across all sources group, where balanced is defined as not relying too heavily on one particular energy source, on the assumption that similar improvement rates apply to all energy supply and end-use technologies (IPCC 2014).

3.2.4 Land cover

Dominant landcover types for 2000 (classified of the CORILIS layers) were downloaded from the European Environmental Agency database having a resolution of 10 km. A land cover type is dominant in a point when its density value in that point is bigger than a threshold value. The dominant land cover types came up according to CORILIS methodology. Data were generalized by crossing them with a regular grid of 3x3 km resolution and then smoothed using a search radius of 20 km. The smoothing algorithm uses the Gaussian function to weight data in the neighborhood. Then the smoothed features were reclassified into Dominant Landscape types. The criterion Vn > mean + standard deviation was used to assign dominant character to smoothed features. Vn is the smoothed value of class n in a given cell of the map. The landcover types used were agriculture, artificial, forest, dispersed urban, composite vegetation, open natural grasslands, rural and pasture areas (EEA 2013).

3.2.5 Soil data

Soil data for both countries were downloaded from the European Soil Portal - Soil Data and Information Systems (European Soil Portal 2013). The data are raster or grid files in the ETRS89 Lambert Azimuthal Equal Area (ETRS_LAEA) co-ordinate system, having a cell size of 1kmX1km. The soil database was constructed in 1985. The current distribution is called version 1.0. It corresponds to the stage of development of the European Soil Data Base on 1999.

The soil characteristics that were used in this study were the following:

- FAO85LEVEL 1 (Soil major group code of the STU from the 1974 (modified CEC, 1985) FAO-UNESCO Soil Legend), reclassified into classes b (cambisol), g (gleysol), j (fluvisol), i (lithosol), l (luvisol)
- 3 classes of Bs-top (Base saturation of the topsoil) (Medium, Low, and High)
- Moor, arable, pasture and grassland areas

Table 2. Common herbicide resistant weeds found in Greece and Germany, included in the study.

HERBICIDE RESISTANT WEEDS					
Species	**Common Name**	**Description**	**Photosynthetic pathway**	**Greece (present)**	**Germany (present)**
Amaranthus retroflexus	Redroot Pigweed	Summer annual dicot broadleaf	C_4	+	+
Chenopodium album	Lambsquarters	Summer annual dicot	C_3	+	+
Conyza canadensis	Horseweed	Summer annual dicot	C_3	+	+
Echinochloa crus-galli	Barnyardgrass	Summer annual monocot grass	C_4	+	+
Papaver rhoeas	Com poppy	Winter annual dicot	C_3	+	+
Lolium rigidum	Rigid ryegrass	Winter annual grass monocot	C_3	+	-
Sorgum halepense	Johnsongrass	Summer perennial monocot grass	C_4	+	-

3.3 Atmospheric Correction of Landsat ETM+ orthorectified data - Dark object subtraction (DOS)

Remotely sensed radiance includes unwanted path radiance caused by atmospheric interactions. The removal of this radiance is called atmospheric correction. In this study, in order to eliminate the unwanted radiance a Dark Object Subtraction (DOS) method was followed, proposed by Chavez 1988. The DOS method includes finding the darkest object (usually water bodies) in the image, assuming that its spectral reflectance should be zero, and subtracting the path radiance from each pixel of the image. For this reason ENVI software was used. In ENVI, the DOS is called Dark Subtract, click Basic Tool -> Preprocessing -> General Purpose Utilities -> Dark Subtract -> select the image as the input file, click OK, select the Band Minimum, which means that the minimum value of each band will be automatically selected, and then this value will be subtracted from all pixels in this band.

3.4 Calculating Normalized Difference Vegetation Index (NDVI)

Spectral value images represented by digital number (DN) contain substantial noises. A significant proportion of such noises can be normalized by converting the DN to at-satellite reflectance value. The use of a radiometric correction is to convert the DN values to absolute radiance and at-satellite reflectance values. Absolute radiance is required when utilizing temporal data that may come from different sensors, variating in sun elevation angle, acquisition date, illumination geometry etc (Huang et al. 2001).

Therefore, the process to calculate the NDVI index using the Landsat ETM+2000 (Landsat org. 2013) ortho-rectified and atmospheric corrected data, is described by the following numbered steps, beginning with converting DN values to radiance, and further to reflectance values:

1. Reclassify the Landsat data so that all zero values are mapped to "NoData". Reflectance is not calculated for the sections where data is missing.

2. *Convert DN data to radiance data. Before converting to reflectance data, DN have to be converted to radiance*. This is done using the following expression (Eqn 2):

$$L_\lambda = (gain_\lambda * DN) + bias_\lambda \text{ (Eqn 2)},$$

where L_λ is the calculated radiance [in Watts / (sq. meter * μm * ster)] , DN is the Landsat ETM+ DN data, and the gain and bias are band-specific numbers. The latest gain and bias numbers for the Landsat ETM+ sensor are given in Chander et al. (2009) and are shown in table 3.

Table 3. Gain and bias for Landsat ETM+.

Band	Gain	Bias
1	0.778740	-6.98
2	0.798819	-7.20
3	0.621654	-5.62
4	0.639764	-5.74
5	0.126220	-1.13
7	0.043898	-0.39

3. *Convert radiance data to reflectance data*

Reflectance is a quantity measured by the Landsat sensors. To make better comparisons between the different scenes, a conversion of radiance to reflectance is needed, by removing differences caused by the position of the sun and the differing amounts of energy output by the sun in each band. The reflectance can be thought of as a "planetary albedo," or fraction of the sun's energy that is reflected by the surface. It can be calculated using the following expression (Eqn 3):

$$R_\lambda = \frac{\pi * L_\lambda * d^2}{E_{sun,\lambda} * \sin(\theta_{SE})} \quad \textit{(Eqn 3),}$$

where R_λ is the reflectance (unitless ratio), L_λ is the radiance, *d* is the earth-sun distance (in astronomical units), $Esun_{,\lambda}$ is the band-specific radiance emitted by the sun, and θ_{SE} is the solar elevation angle. These values are given in Chander et al. (2009) and are repeated in table 4.

Table 4. Band specific radiance emitted by the sun.

Band	$E_{sun,\lambda}$ [Watts/(sq.meter* μ m)]
1	1997
2	1812
3	1533
4	1039
5	230.8
7	84.9

The second and third pieces of information are *d*, the earth-sun distance, and θ_{SE}, the solar elevation angle. These two values are dependent on the individual scene, specifically the day of the year and the time of day when the scene was captured. The solar elevation angle and the day of year are listed in the header file for each scene.

4. *Enforce positive reflectances*

corrected_reflectance = CON([reflectance] < 0.0, 0.0, [reflectance]) *(Eqn 4)*

5. *Calculate NDVI*

NDVI = (band 4 – band 3) / (band 4 + band 3) *(Eqn 5)*

The methodology for assisting to calculate the NDVI was downloaded from the following site:
(Colorado State University 2013)

6. *Mosaic NDVI maps*

To create a single NDVI index raster, multiple raster NDVI index datasets were merged into a mosaic for each country. All rasters used for each country had the same spatial reference (WGS_84), the same pixel size, and the same number of bands. In many cases, there was some overlap of the raster dataset edges that were mosaicked. These overlapping areas were handled by taking the mean of the overlapping cell values. The

rule used in raster calculator was:
Mosaic and Con(IsNull([ndvi184033_c]), focalmean([ndvi184033_c], rectangle, 2, 2), [ndvi184033_c]) no data areas for mosaic Ndvi

3.5 Genetic Algorithm for Rule-Set Production (GARP Models)

One of the models that can be used to study species distribution in relation with bioclimatic conditions is the Genetic Algorithm for Rule-Set Production (GARP) model. This is a method that can be used to model the potential distribution of species based on species occurrence and not absence data.
GARP relates ecological characteristics of known occurrence points to those of points randomly sampled from the rest of the study region, developing series of decision rules that best summarizes those factors associated with the species presence (Stockwell & Peters 1999).
The software used for the analysis was openModeller desktop 1.1.0 (de Souza Muñoz et al. 2011). The environmental layers that represent abiotic factors used to determine species distribution for the years 2000, 2020 and 2020, were the 11 bioclimatic factors described in section 3.13 (Bioclimatic data and altitude). The results of the analysis were validated by using the Receiver Operating Characteristic (ROC) curves. The ROC plot is obtained by plotting sensitivity as a function of the falsely-predicted positive fraction, or commission error (1-specificity), for all possible thresholds of a probabilistic prediction of occurrence. The resulting area under the ROC curve provides a single measure of overall model accuracy, which is independent of a particular threshold. AUC values range from 0 to 1, with a value of 0.5 indicating model accuracy not better than random, and a value of 1.0 indicating perfect model fit (Fielding and Bell 1997).

3.6 Habitat Suitability Models (HS) using Biomapper

Another multivariate approach that can be used to study species distribution using environmental variables and species presence only data is Habitat Suitability Models (HS Models) using Biomapper software by Hirtzel et al. (2002). These Models are designed to build habitat suitability maps using not only bioclimatic variables, but a number of other environmental variables, such as, soil conditions, and land cover. Moreover, HS models cannot make predictions of species future distribution in relation to GARP models.
The first step in this analysis was that all variables used were converted into .rst file format using IDRISI software. The data were exported form ArcGIS 10.0 as .img files and imported to IDRISI. Concerning species data, Boolean maps of 0 and 1 values indicating absence and species presence were created for each species using IDRISI software.
Moreover, the predictor variables named as EGV's, i.e. Ecogeographical variables that were used in this study, were of two types, quantitative data, such as the 11 bioclimatic factors and altitude data, and the qualitative or Boolean data, such as all landcover and soil data downloaded for this study and referred above. The methodology to transform the data used for the analysis was downloaded by the Biomapper's wikispace (Biomapper 2013).
Quantitative data are ready to be used by the model. On the contrary, qualitative data need to be transformed into several Boolean maps, each describing a relevant category.

Concerning NDVI index, it was classified into 5 classes (table 5).

Table 5. NDVI values classification in Greece and Germany.

	Classes				
	1	**2**	**3**	**4**	**5**
Greece	-1 to 0	0 to 0.25	0.25 to 0.5	0.5 to 0.75	0.75 to 1
Germany	-1 to 0	0 to 0.4	0.4 to 0.6	0.6 to 0.8	0.8 to 1

The core part of HS models is the Ecological-Niche Factor Analysis (ENFA) that computes suitability functions by comparing the species distribution in the EGV space with that of the whole set of cells. The first factor extracted maximizes the marginality of the species, whereas the next factors maximize the specialization.
The result of ENFA is a score matrix and a set of eigenvalues.The first column of this matrix is the marginality factor. The score matrix (eigenvectors) provides information of how the factors are correlated with the variables. The other columns are the V-1 specialization factors, (V is the number of variables).
Additionally, the eigenvalues indicate how much variance is explained by the factors. The larger they are, the more information each factor is conveying (if your species was distributed randomly throughout the study area, the eigenvalues would be all close to 1, marginality would be close to 0 and tolerance would be close to 1). Finally, eigenvectors and eigenvalues are readily interpreted and can be used to build habitat-suitability maps (Hirzel et al. 2002).

3.6.1 Validating Habitat Suitability Models

3.6.1.1 Boyce Index

In order to validate the Habitat Suitability Models, Boyce et al. (2002) proposed a method that consists of partitioning the habitat suitability range into b classes (or bins), instead of only two. For each class i, it calculates two frequencies: Pi, the predicted frequency of evaluation points, and the expected frequency (Fi) of evaluation points, i.e. the frequency expected from a random distribution across the study area. For each habitat suitability map the prediction the predicted-to-expected (P/E) ratio is given. A low suitability class should contain fewer evaluation presences than expected by chance, resulting in a ratio < 1. Conversely, high suitability classes should have ratio increasingly higher than 1. Boyce et al. (2002) measure this monotonic increase by the Spearman rank correlation coefficient between Fi and i. This “Boyce Index” Bb varies from −1 to 1. Positive values indicate a model whose predictions are consistent with the presences distribution in the evaluation dataset, values close to zero mean that the model is not different from a chance model, negative values indicate an incorrect model, which predicts poor quality areas where presences are more frequent (Hirzel at al. 2006).

3.6.1.2 Absolute and Contrast Validation Indices

The Absolute Validation Index (AVI) and the Contrast Validation Index (CVI) were developed before Boyce Indices and cross-validation processes. The AVI is the proportion of validation cells that have a HS>50. However, an AVI of 1 (best value) could be obtained by a model predicting presence everywhere. In other words, the AVI cannot

tell if a model is better than chance. The CVI does this: If we call AVI0 the proportion of all cells that have a HS>50, then the CVI = AVI – AVI0. A model predicting presence everywhere would thus get a CVI of 0. A good model should have CVI>0.3. All presence-only evaluation Indices are always somewhat sensitive to the study area. Indeed, the larger the study area (or the more specialized the focal species), the higher the CVI (Hirtzel and Arlettax 2003).

3.7 Rank - Abundance diagrams and Spatial Heterogeneity

In order to test the potential of rank–abundance diagrams in discriminating differences of the heterogeneity of the landscape around species (Rocchini and Neteler 2012), 3 km buffer zones were created around species points, DN value's of the corresponding NDVI and their abundances were extracted, and further analyzed. For this reason NDVI values were converted into an 8-bit band splitting values into 256 equal intervals, hereafter simply referred to as DN values.

For each species, rank–abundance diagrams were derived as follows: (i) the abundance of each DN NDVI value in each species buffer zone was calculated, (ii) DNs were ranked, with the most abundant DN ranked first and (iii) relative abundance was plotted against ranks. Relative abundances were used in order to be consistent with previous studies using rank–abundance diagrams (Magurran 1988). In other words, the most abundant - common species will have a higher number of individuals and will be ranked 'first'. Then, relative abundances are plotted against ranks (rank–abundance diagram), and species are ranked from commonest to rarest along the abscissa and their abundances are plotted along the ordinate. The application of rank–abundance diagrams to remotely sensed imagery is simple. The individuals correspond to individual pixels and the species to reflectance values (Rocchini and Neteler 2012). Digital Number abundance was plotted even on a log scale in order to decrease the outlier effect, that is, the smoothing of the curve due to hyper-dominant DN values.

Additional indices were calculated, such as species richness, Shannon diversity or entropy index (Shannon and Weaver 1962) and Pielou eveness index (Pielou 1969). Species richness is referred as the number of different DN values around each species point. Moreover, Shannon index, H, indicating the relative proportion of each DN value, is calculated using the following equation (Eqn 6):

$$H = -\sum_{DN=1}^{N} P \ln (P), 0 \leq H \leq \ln(N) \qquad \textit{(Eqn 6)}$$

When the Shannon index takes into account the equitability of the system, then the result is the Pielou index J *(Eqn 7)* that indicates the maximum possible diversity with the same number of DNs N.

$$J = \frac{H}{\ln(N)}, 0 \leq J \leq 1 \qquad \textit{(Eqn 7)}$$

3.8 Spatial variability: Variograms via Kriging interpolation

In order to explain spatial variation of species points related to NDVI, Kriging multistep process was used, since it includes statistical analysis, variogram modelling, and surface variance exploration. The main application of Kriging is to interpolate values at un-sampled locations. Parallel, a semivariogram is constructed providing information on the spatial autocorrelation of the dataset. To fit a model to the empirical semivariogram, a spherical function was applied that rises and levels off for larger distances beyond a certain range. Semivariogram modeling is a key step between spatial description and spatial prediction. The sill of a semivariogram is the variance, and range is the distance at which the variogram reaches the sill.

In this study the Kriging tool was applied on mean NDVI on an area of 3km around species points, using the spherical function. Moreover, the sill and square root range of the resulted semivariogram models were used in order to construct spatial variability variograms, a methodology to characterize and quantify the spatial heterogeneity of the landscape (Gariggues et al. 2006). The range square root value summarizes the mean length scale (spatial scale) of the data, on the x -axis. The theoretical dispersion variance from variograms quantifies the spatial heterogeneity, on the y-axis (Gariggues et al. 2006).

3.9 Regression Models

3.9.1 Ordinary Least Square analysis & autocorrelation

Ordinary Least Squares (OLS) is a linear regression model for studying relationships. It is the starting point of spatial regression techniques, and it fits one equation to represent the dataset. The equation (Eqn 8) that describes OLS is the following:

$$y = \beta_0 + \beta_1 * \chi_1 + ... + \beta_n * \chi_n + \varepsilon$$

(Eqn 8),

where

y is the process we are trying to explain,

β is the coefficients of the model

χ is used to predict the dependent variable

The model is constructed by minimizing the sum of squared vertical distances between the observed and the predicted values. The difference between these actual and the predicted Y value is called the residual of the model, and it provides an indication of how well the model predicts each data point. So as to check for spatial autocorrelation, the Global Morans' I tool on the residuals of OLS analysis was used. Morans' I index evaluates whether the pattern expresses is clustered, dispersed, or random, by calculating a z-score and p-value. In this study, OLS was performed using the mean and standard deviation values of NDVI index on an area of 3km around species points. Finally, Morans' I autocorrelation was performed on residuals to check for spatial autoregression.

3.9.2 Weight of Evidence and Logistic Regression (SDMArcGIS)

Species' distribution models (SDMs) attempt to predict the potential distribution of species by interpolating identified relationships between species' presence/absence, or presence-only data on one hand, and environmental predictors on the other hand, to a geographical area of interest (Raes & ter Steege 2007).

Due to non-linearity of the data, logistic regression was the next step. Spatial Data Modeller (SDM) for ArcGIS 10 was used to perform the weight of evidence and logistic regression analysis. Species presence was the dependent variable, whereas NDVI reclassified into 5 classes (table 5) was the independent variable. NDVI in Germany had very few values between 0-0.25, therefore classified differently. Environmental settings were specified in Arctool box, weight of evidence (unique type) was calculated and finally logistic regression was performed. The calculation of weights requires one training site per unit area. The unit area selected was 3 square kilometers. Therefore, the result was one training point per 3 kilometers buffer zone.

Weight is a measure of an evidential-theme class (feature) as a predictor of training points. A weight is calculated for each theme class. For binary themes, these are often labeled W+ and W- (Raines and Bonham-Carter 2006). A positive weight indicates areas where training points are likely to occur, whereas a negative weight indicates an area where training points are not likely to occur. While calculating weights of evidence (WofE) for NDVI index grid classified, a contrast value is calculated for each class, by combining the positive and negative weights. Moreover, contrast is a measure of evidence layer's significance in predicting the location on training points (Bonham-Carter 1994).

A positive contrast that is significant, suggests that the evidential theme is a good predictor. The confidence of the evidential theme is the contrast divided by the standard deviation. The second type of confidence can be calculated for each theme by dividing the posterior probability by its standard deviation (Raines 1999). Confidence values and its corresponding level of significance are presented in table 6.

Table 6. Test values calculated in WofE and their respective studentized T values expressed as level of significance in percentages.

Test value	Studentized T value (confidence expressed as level of significance)
2.576	99.5 %
2.326	99 %
1.96	97.5 %
1.645	95 %
1.282	90 %
0.842	80 %
0.674	75 %
0.542	70 %
0.253	60 %

The normalized contrast provides a useful measure of significance of the contrast because of the uncertainties of the weights and missing data (Raines 1999).

3.9.3 Logistic & WofE Validation

3.9.3.1 CAPP curve method (Cumulative area posterior probability)

So as to validate logistic regression results, a CAPP curve is created by plotting posterior probability on the Y axis and cumulative area (from low to high posterior probability) on the X axis. Posterior Probability is the probability that a unit cell contains a training point, given states of information from the evidential themes. This measurement changes from location to location depending on the values of the evidence. (Raines and Bonham- Carter 2006). After creating the curves, a model is fitted, and the Area Under Curve (AUC) is calculated, using a widget (Wolframalpha 2013).

4. Results

4.1 Climatic changes

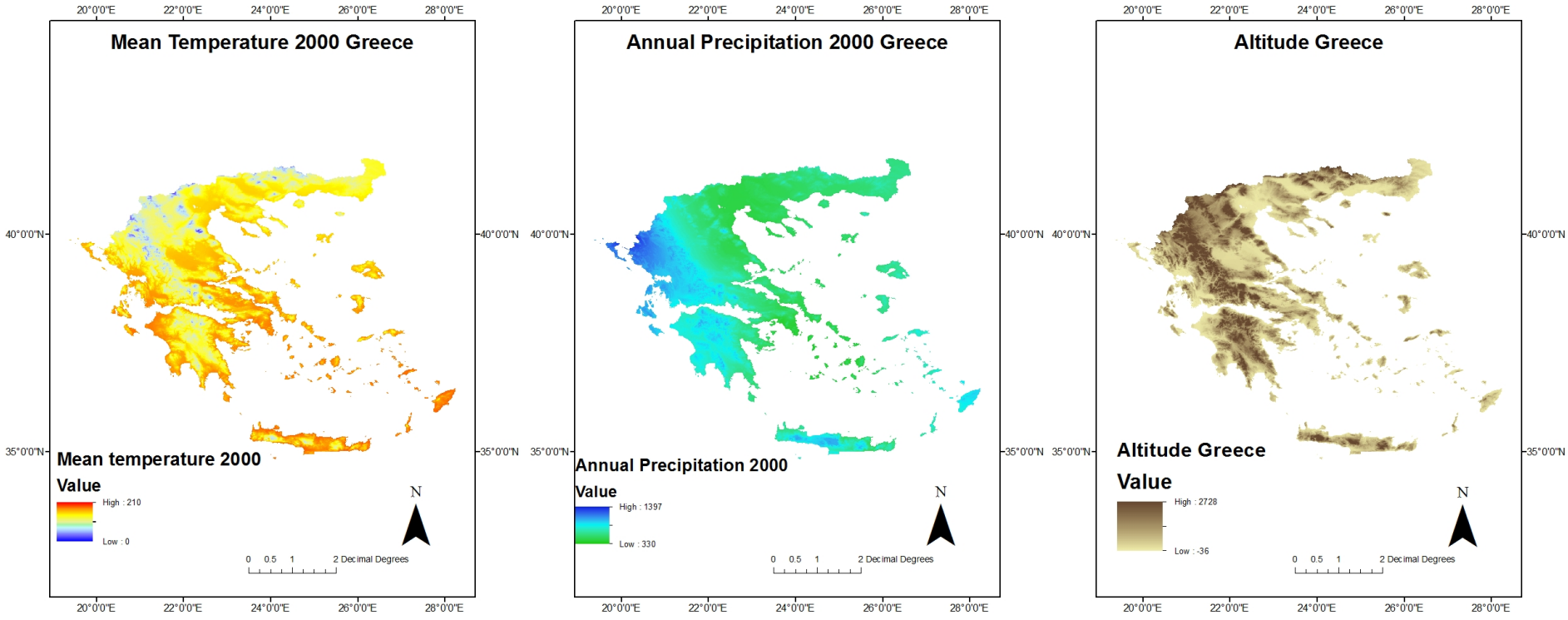


Figure 3. Mean temperature (0.1 °C), annual precipitation (mm) for 2000, and altitude (m) in Greece.

The main environmental factors affecting species distribution are found to be mean temperature (annual), annual precipitation and altitude. The mean temperature, annual precipitation and altitude of Greece in 2000 is presented in fig. 3. The central part of Greece is crossed by a high mountainous area (Pindos) creating two different patterns of precipitation. A wetter one at the western part of the country and a drier one at the eastern part (fig. 3). The mean temperature in Greece is higher at the Southern part of the country, at low land areas and also at the Greek islands.

Concerning Germany, the bioclimatic characteristics, and also the altitude values, of Greece are given at figure 4. Based on figure 4, mean temperature is higher at the western part of Germany, gets average values in the north eastern part, and low values in the South mountainous part of Germany. Moreover, at higher altitudes, south, and at a lower value central Germany, the precipitation is higher.

Concerning climatic changes, in Greece the temperature and precipitation pattern does not alter (figs 3 and 5). However, the minimum, low values as depicted on figs 3 and 5, mean temperature and annual precipitation tend to decrease by 2050. The high values of the same maps (figs 3 and 5) remain the same.
Comparing figs 4 and 6, the whole country of Germany becomes warmer, especially the Southern part, by 2050. The increase is recorded between the years 2020 and 2050 (fig. 6). However, according to figs 4 and 6, the temperature does not alter in the time period between 2000 and 2020. The pattern of annual precipitation does not seem to alter through the studied years. However, a small decrease is recorded for the maximum values of annual precipitation by 2050. Hence, Germany is becoming warmer by 2050.

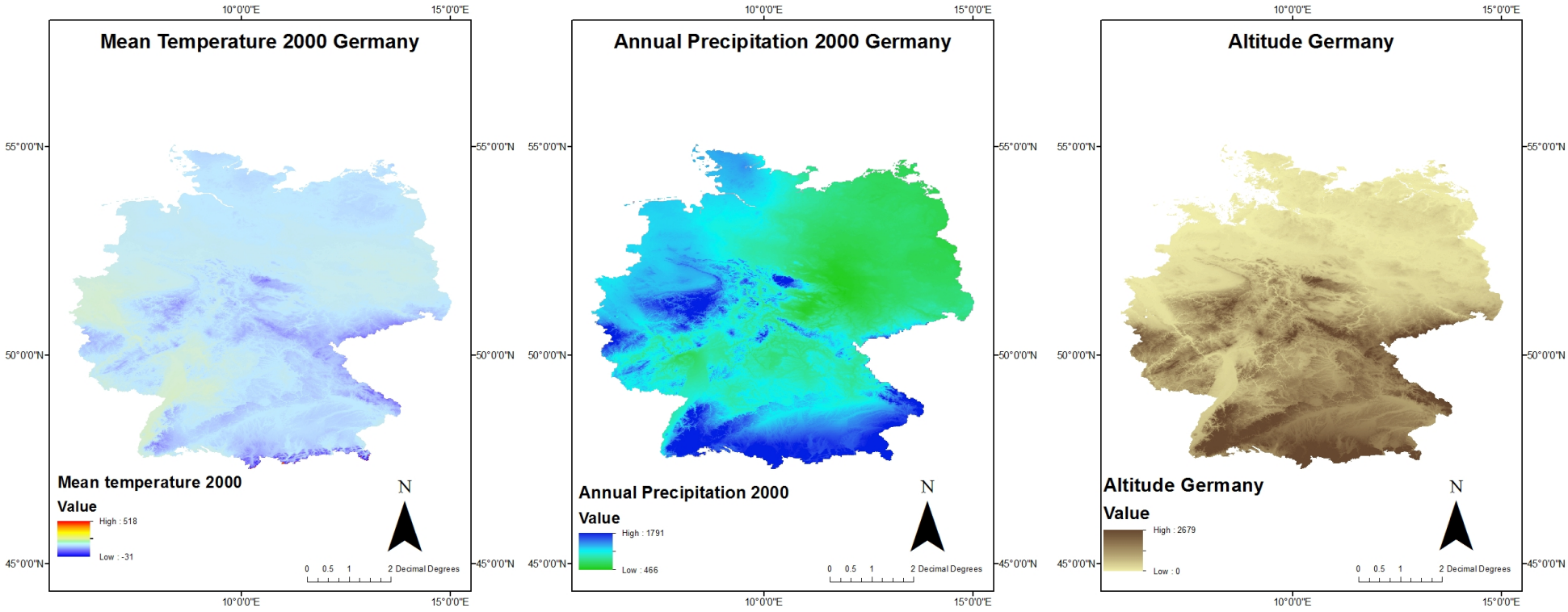


Figure 4. Mean temperature (0.1 °C), annual precipitation (mm) for 2000, and altitude (m) in Germany.

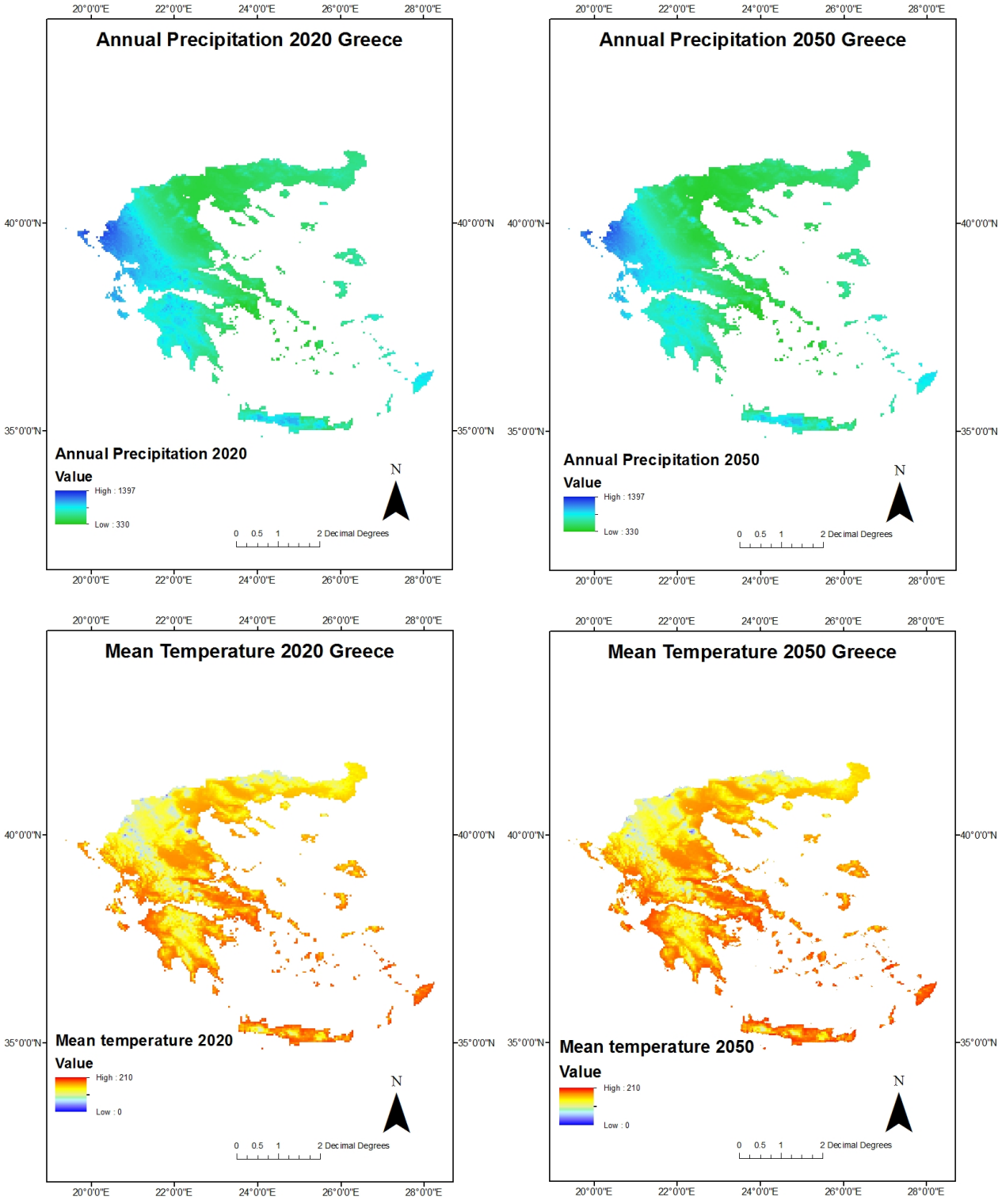


Figure 5. Mean temperature (0.1 °C) and annual precipitation (mm) for 2020 and 2050 in Greece.

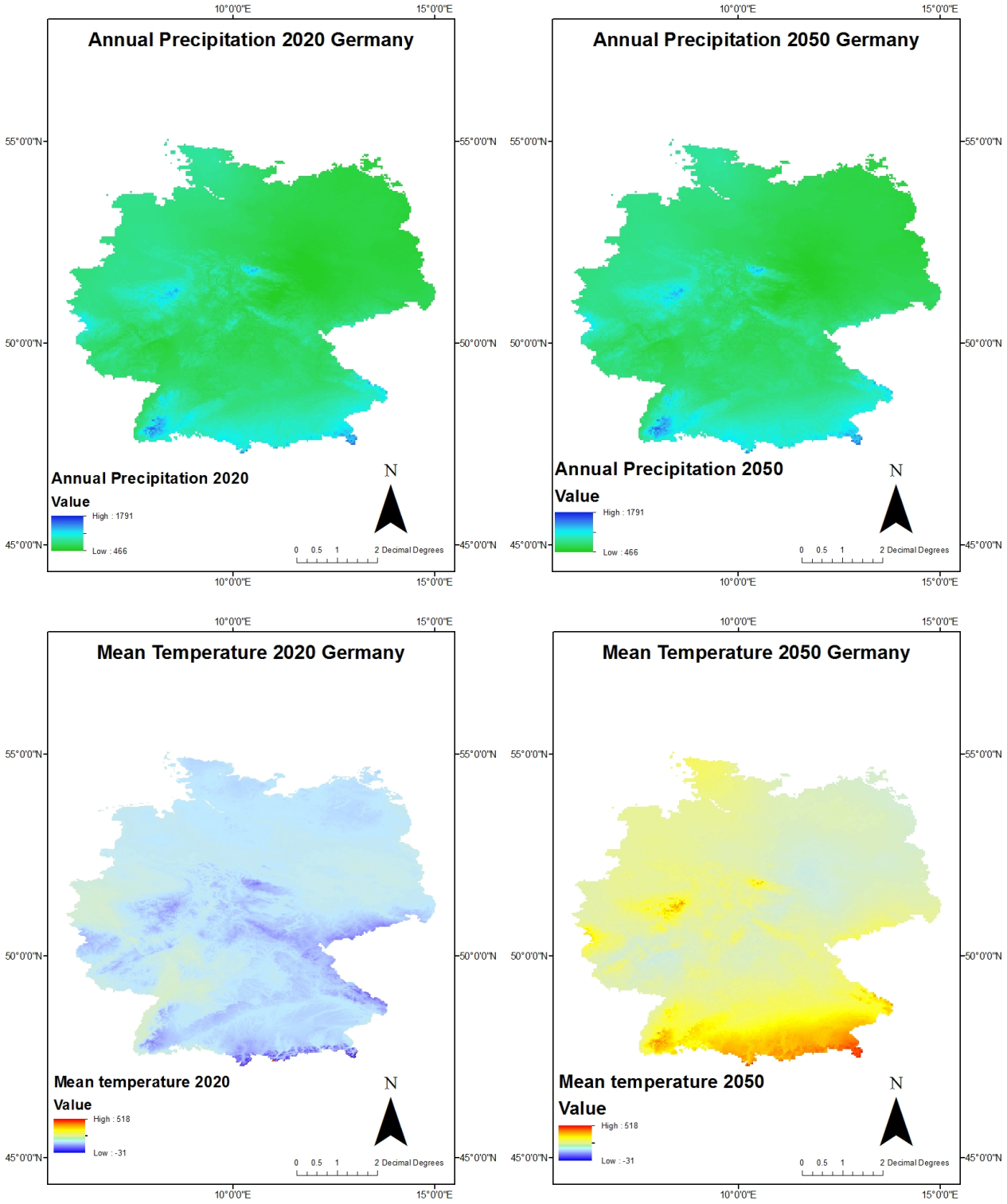


Figure 6. Mean temperature (0.1 °C) and annual precipitation (mm) for 2020 and 2050 in Germany.

4.2 Habitat Suitability Maps (HS maps)

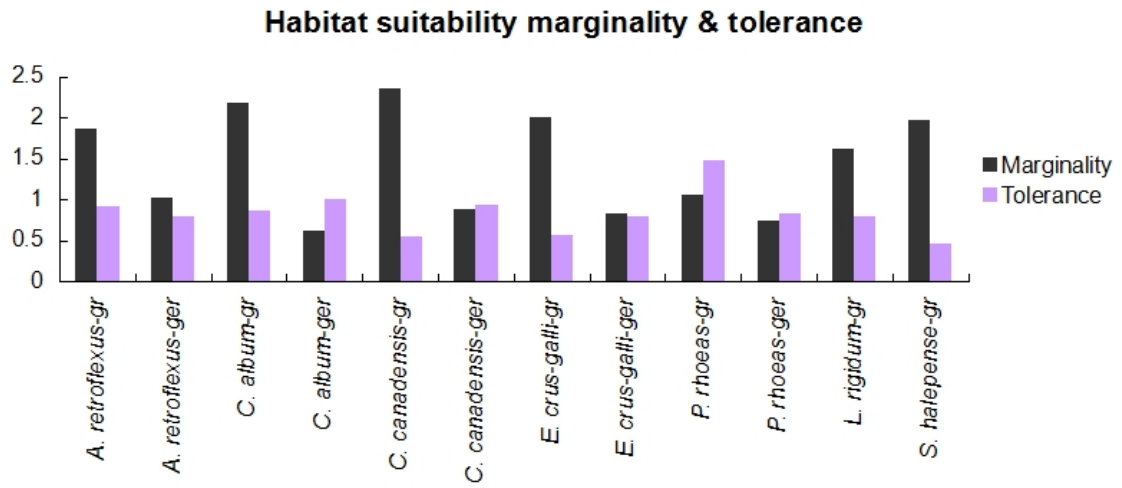


Figure 7. Marginality and tolerance of studied species, in Greece and in Germany, derived from Habitat Suitability models by Biomapper.

Marginality seems to be higher for all species in Greece, related to the same species in Germany, where marginality seems to be very low (fig. 7). The lowest marginality for species in Greece is recorded for *P. rhoeas*, whereas the highest for *C. canadensis*, *C. album*, and *E. crus-galli*. Concerning tolerance, all species both in Greece and Germany do not get a very high value (fig. 7).

In order to assess which factors affecting mostly species distribution (eigenvector) and also which factor is highly related with species marginality (eigenvalue), scatter-plots were created depicting the eigenvector and eigenvalue of factors used in Environmental Factor Analysis (ENFA) and Habitat Suitability models.

Concerning *A. retroflexus,* the most important factor affecting species distribution (eigenvector) is the variable annual mean temperature (bio_1) (fig. 8a). The same holds for the species both in Greece and in Germany, having though a higher impact on the species in Greece. The next factors affecting *A. retroflexus* distribution in Greece is annual precipitation (bio_12), precipitation of the wettest month (bio_13), and base saturation of the topsoil, high value (BS_top_H). Moreover, the main factors affecting *A. retroflexus* distribution in Germany are mean temperature (bio_1) and annual precipitation (bio_12) values, in accordance with the eigenvector values of figure 4. Additionally, the highest eigenvalue is getting the mean temperature (bio_1) for *A. retroflexus* in Greece, whereas altitude for the same species in Germany.

Concerning *C. album* in Greece, highest eigenvector is recorded for bio_1, followed by bio_12 and bio_13, whereas in Germany bio_1, bio_15 (seasonal precipitation), ndvi004, and arable land are the highest eigenvector variables (fig. 7b). Moreover, the factor that is related with the species marginality is bio_1 (mean temperature) in Greece, and altitude in Germany.

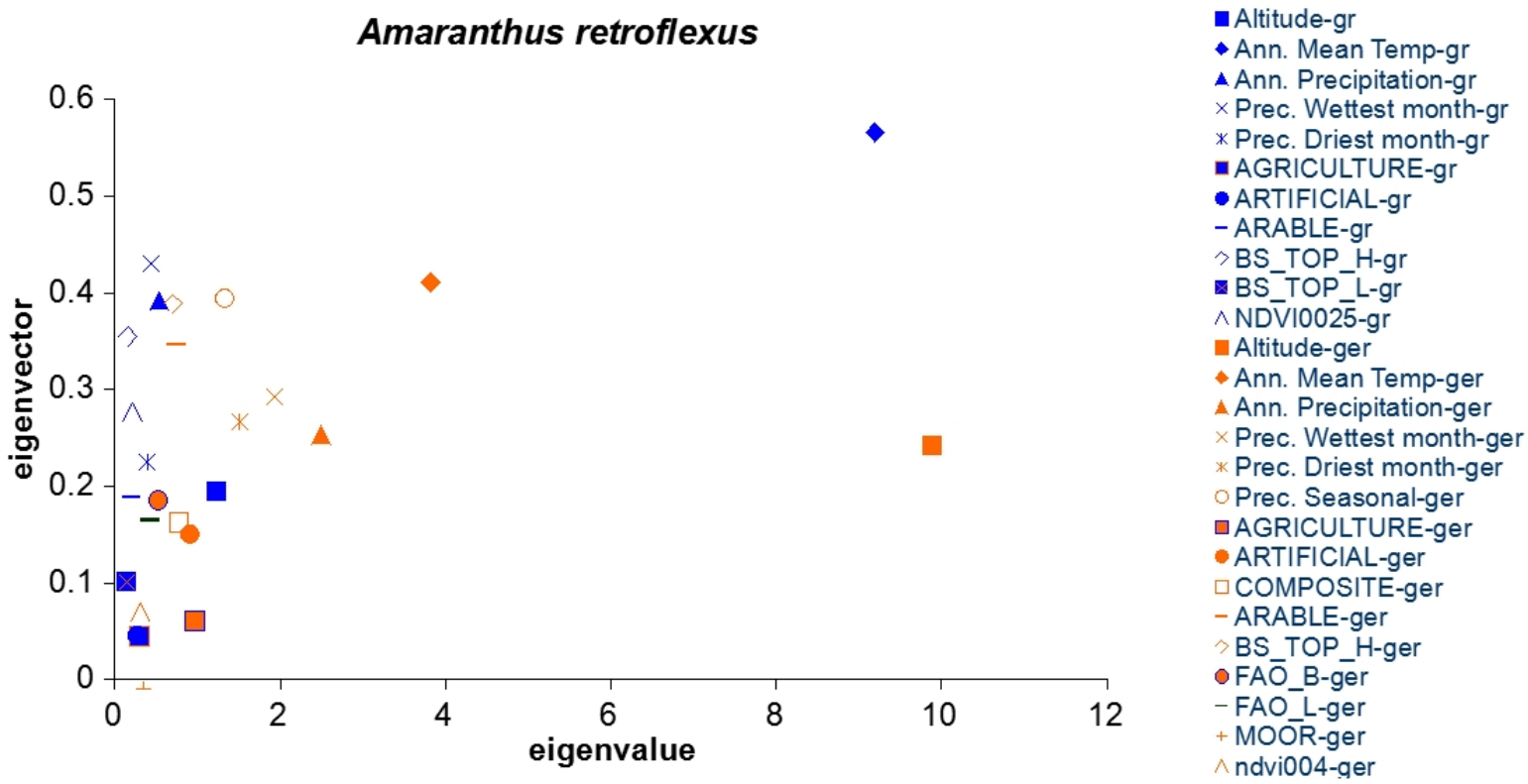


Figure 8a. Eigenvector (y-axis) versus eigenvalue (x-value) scatterplot derived from Environmental Factor Analysis for A. retroflexus, in Greece and Germany.

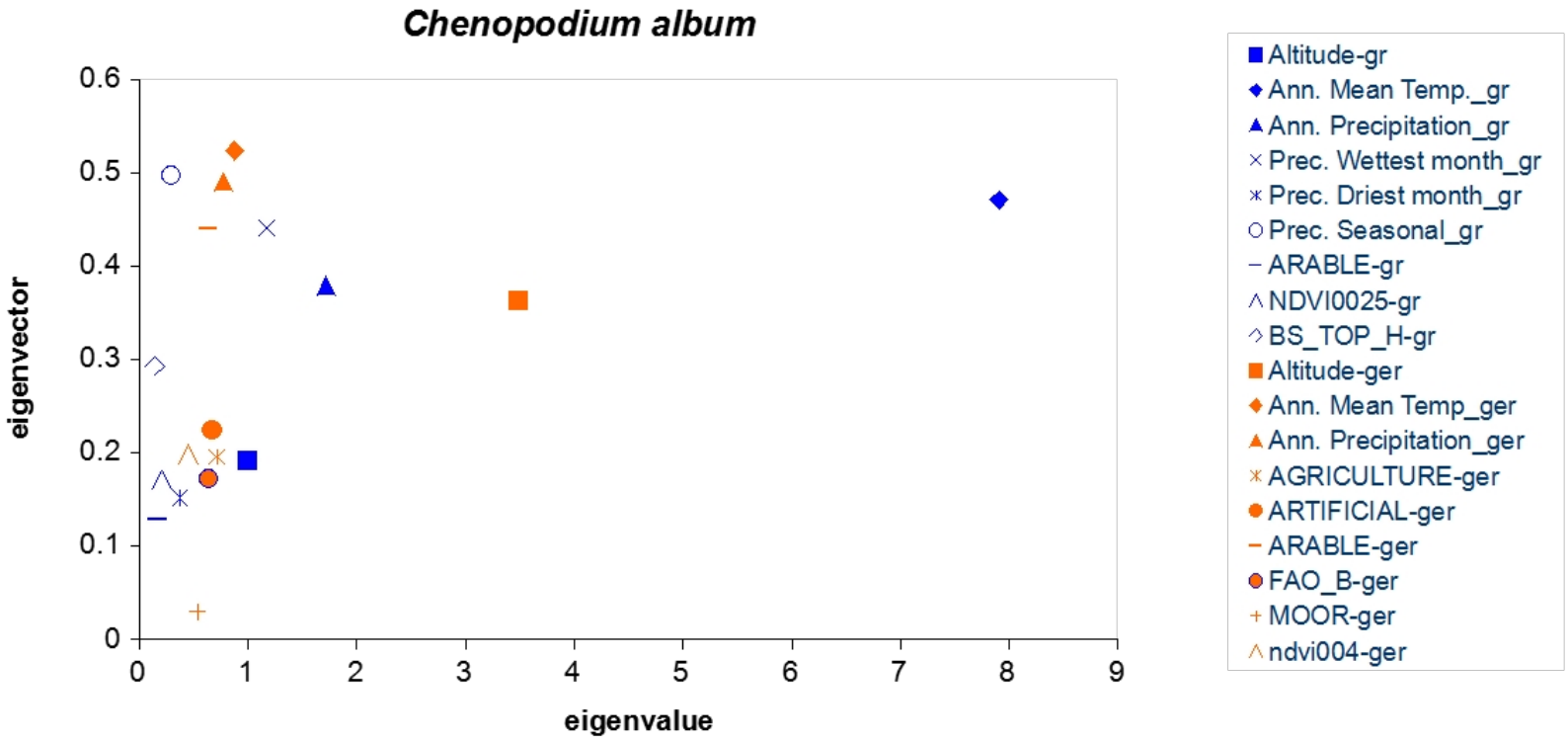


Figure 8b. Eigenvector (y-axis) versus eigenvalue (x-value) scatterplot derived from Environmental Factor Analysis for C. album, in Greece and Germany.

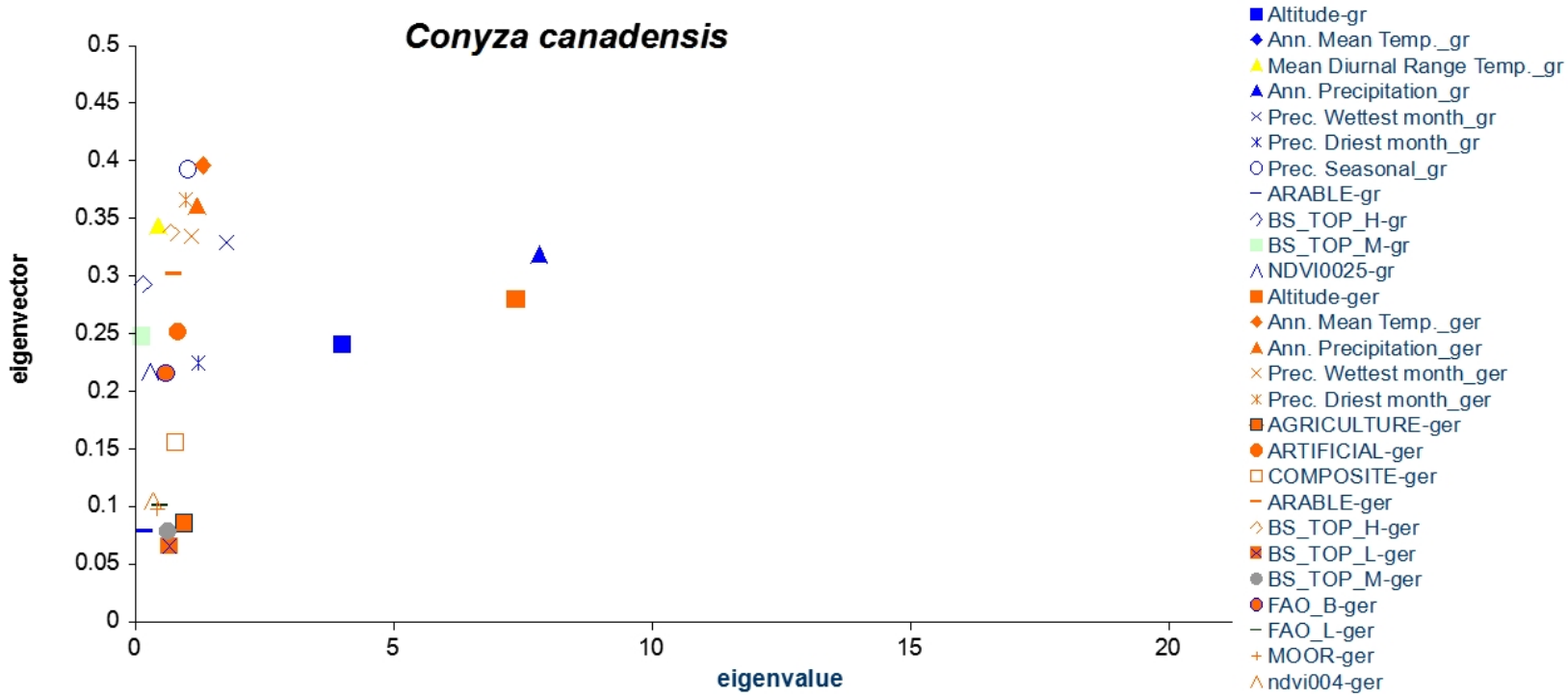


Figure 9a. Eigenvector (y-axis) versus eigenvalue (x-value) scatterplot derived from Environmental Factor Analysis for C. canadensis, in Greece and Germany.

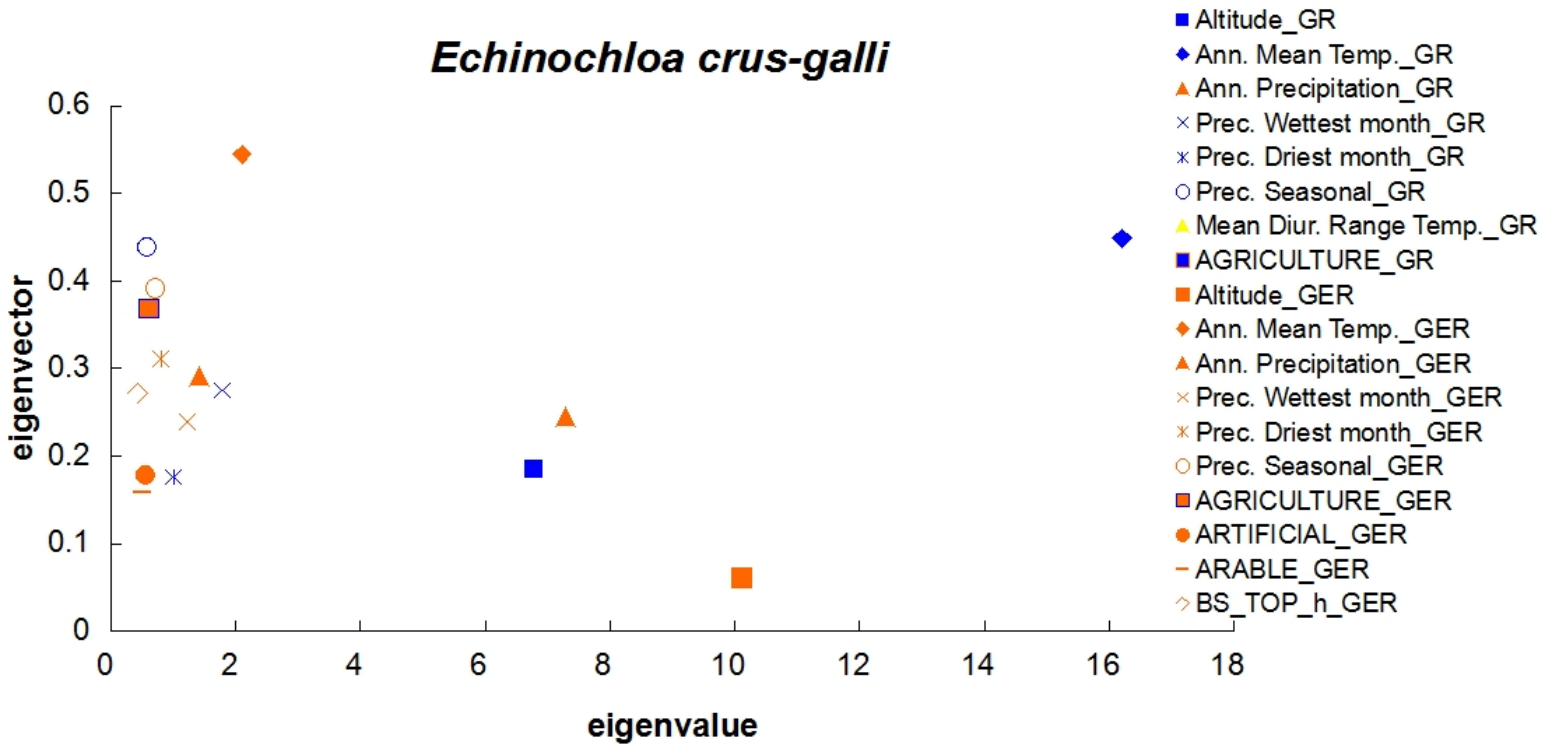


Figure 9b. Eigenvector (y-axis) versus eigenvalue (x-value) scatterplot derived from Environmental Factor Analysis for E. crus-galli, in Greece and Germany.

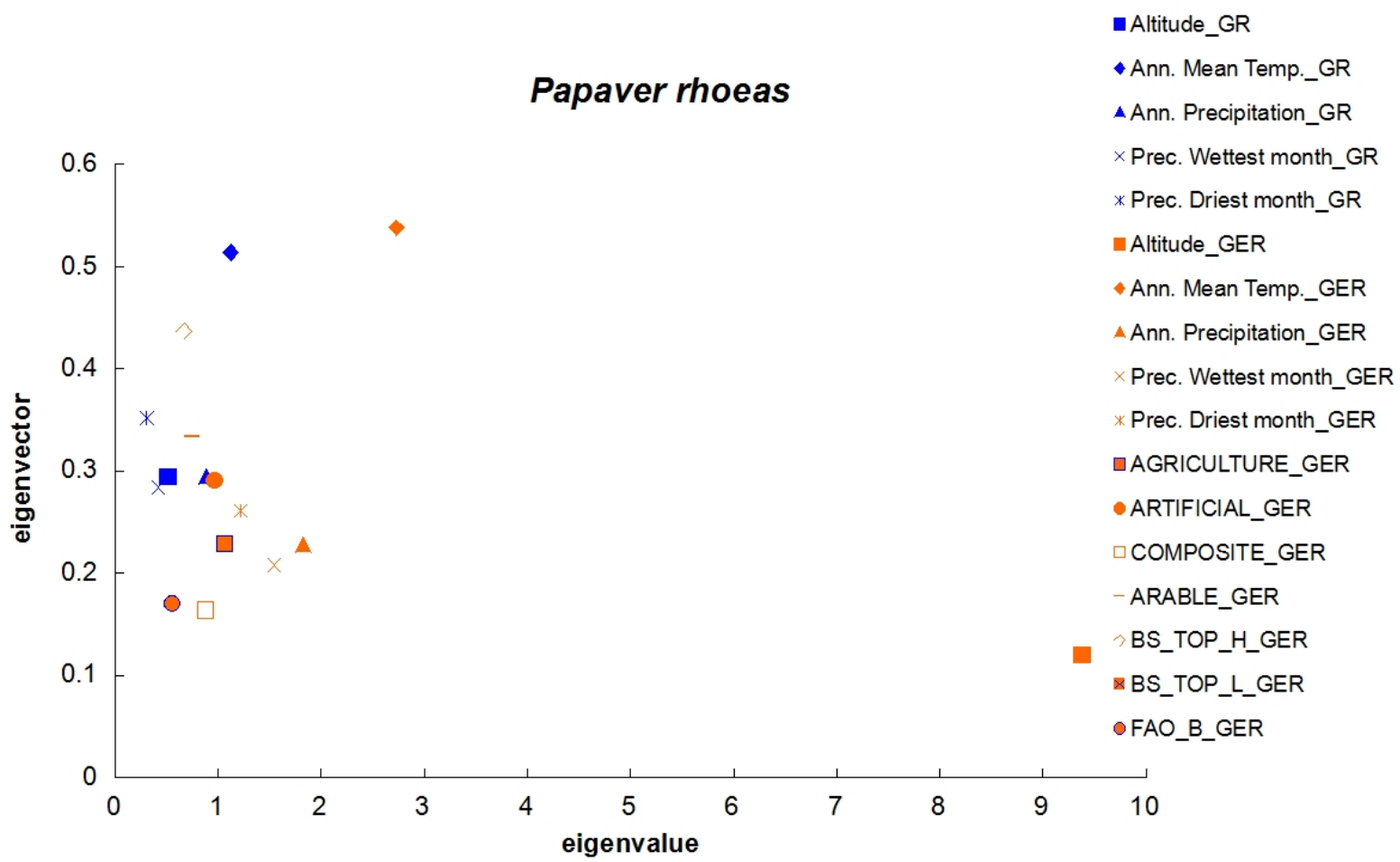


Figure 10a. Eigenvector (y-axis) versus eigenvalue (x-value) scatterplot derived from Environmental Factor Analysis for P. rhoeas, in Greece and Germany.

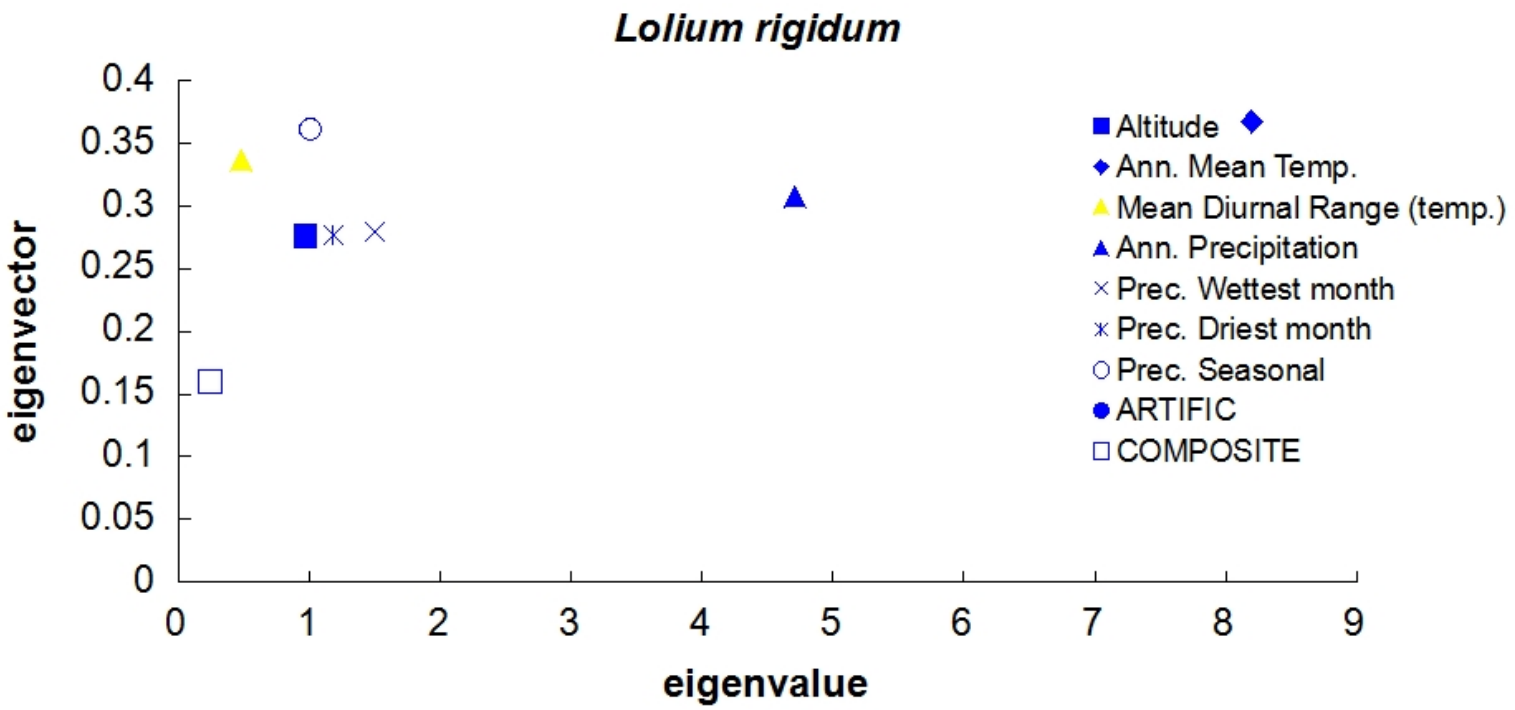


Figure 10b. Eigenvector (y-axis) versus eigenvalue (x-value) scatterplot derived from Environmental Factor Analysis for L. rigidum, in Greece.

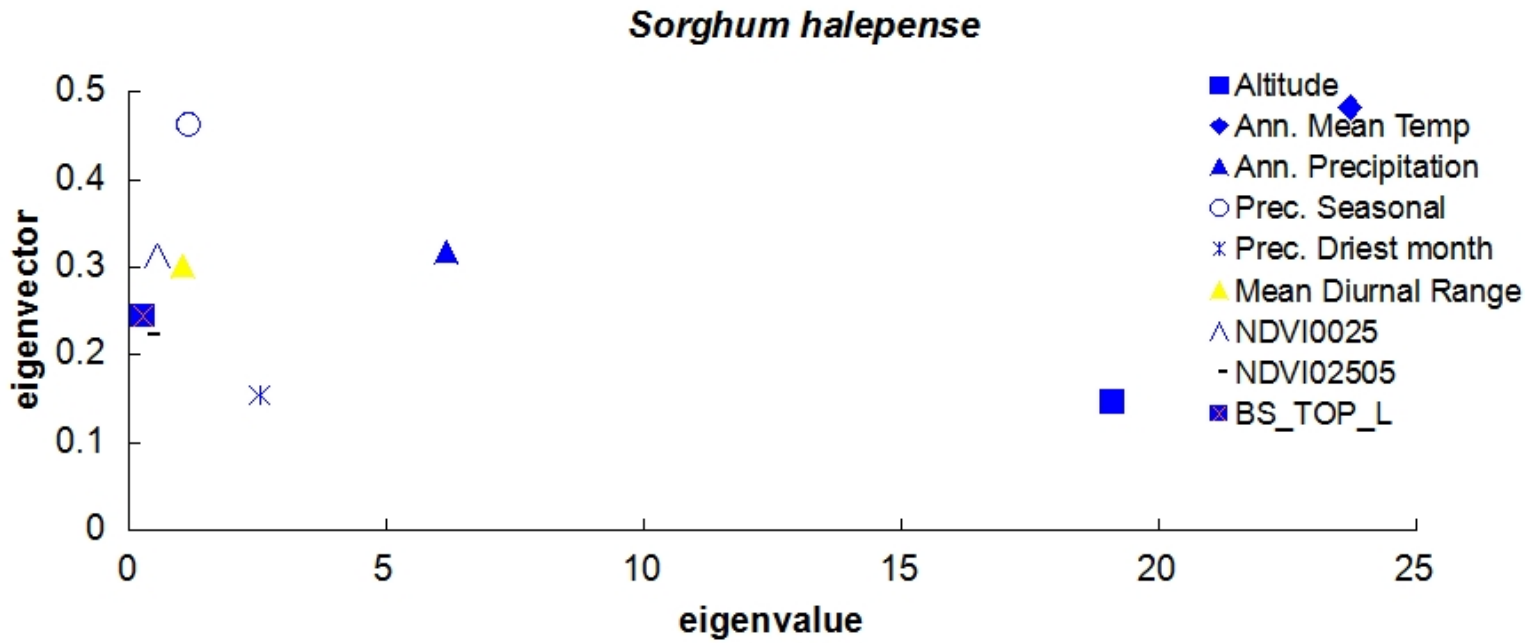


Figure 10c. Eigenvector (y-axis) versus eigenvalue (x-value) scatterplot derived from Environmental Factor Analysis for S. halepense, in Greece.

Higher marginality is recorded for *C. album* found in the Greek landscape, according to the eigenvalues of figure 8b.

C. canadensis distribution is related mainly with variables bio_15 (seasonal precipitation), bio_2 (mean temperature diurnal change), bio_14 (precipitation of driest month), and finally, bio_12 (annual precipitation) in Greece, whereas bio_1, bio_12, bio-13 (precipitation od wettest month), bio_14 (precipitation of driest month) in Germany (fig. 9a). The species marginality (eigenvalue) depends mostly on annual precipitation for the species in Greece, and altitude for the species in Germany (fig. 9a).

Moreover, bio_1 and bio_15 variables are the main factors affecting *E. crus-galli* distribution in Greece, whereas bio_1 and bio_15 and agriculture land factors in Germany (fig. 9b). Higher marginality is recorded for species in Greece than in Germany. In Greece, marginality for the species is related with mean temperature (bio_1), and secondarily with altitude, whereas in Germany with altitude and annual precipitation (bio_12) (fig. 9b).

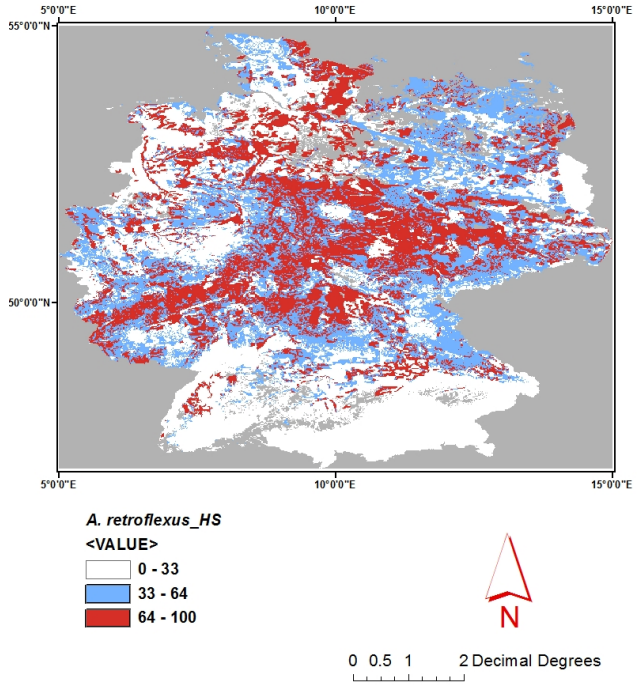
Concerning *P. rhoeas*, the main factor affecting species distribution in both countries is mean temperature (bio_1), according to the maximum eigenvector value. Moreover, the species in Greece presents very low marginality. On the contrary, high marginality is recorded for the species in Germany related to the altitude variable (fig. 10a).

Distribution of *L. rigidum* in Greece seems to be related to almost all bioclimatic factors studied, but also to the artificial landcover type and altitude (fig. 10b). Moreover, the marginality of the species is mostly related with mean temperature (bio_1), and annual precipitation (bio_12) (fig. 10b). Comparing to *L. rigidum*, marginality values are much higher for *S. halepense.* Hence, marginality (eigenvalue) of *S. halepense* in Greece, is related mostly with mean temperature (bio_1) and altitude (fig. 10c). Moreover, species distribution is mostly affected by mean temperature (bio_1), annual precipitation (bio_12), and seasonal precipitation (bio_15) (fig. 10c).

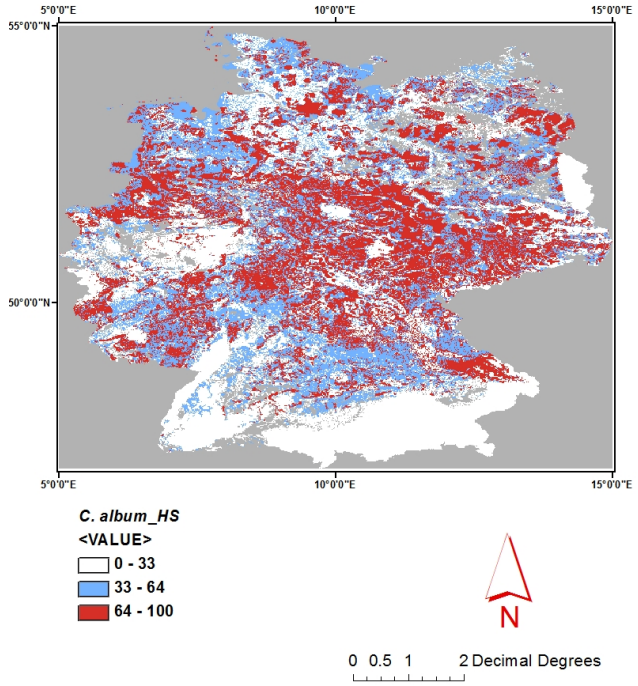
A. retroflexus, *C. album* and *E. crus-galli* are found mostly at the central part of Germany (fig. 11a), where the annual precipitation rate, the altitude and the mean temperature get average to low values, according to figure 4. On the contrary *P. rhoeas* is found at the north eastern part of Germany, where the elevation and the precipitation rate is low; however, the temperatures is a little higher than average (fig. 11b). Finally, *C. canadensis* is expanded at the central and west part of Germany, where the temperature and precipitation are higher related to other areas of Germany (figs 11b and 4).

As far as it concerns Habitat Suitability for the studied weed species in Greece, *A. retroflexus*, *C. album, C. canadensis, and S. halepense* are distributed in the central part of Greece, at lowland areas, with low altitude values, high temperatures especially around the capital of Greece, Attiki, and in non mountainous areas of Peloponnese, according to figures 12a and 3. Among these species, *C. album*, and secondarily *S. halepense* are the more expanded. Additionally, *E. crus-gall*i seems to prefer the mountainous area of Pindos, i.e. an area with low temperatures, and average precipitation (fig. 12a and fig. 3). However, this species is also found at drier and warmer areas of Sterea Ellada, and lowlands of Thessalia (central part of Greece) (figs 12a and fig. 3). Finally, the presence of species, *P. rhoeas* and *L. rigidum,* according to HS maps, is scarce and dispersed all over the Greek landscape (fig. 12b).

Amaranthus retroflexus HS Germany



Chenopodium album HS Germany



Echinochloa crus-galli HS Germany

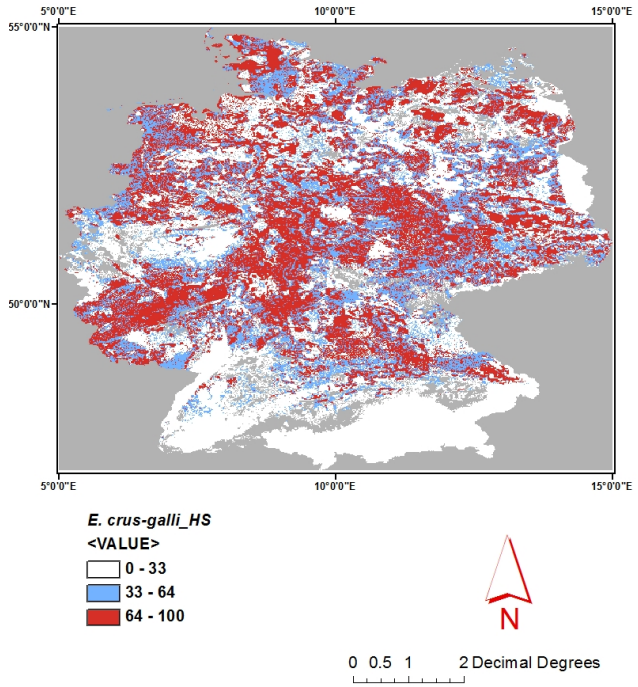


Figure 11a. Habitat Suitability maps for A. retroflexus, C. album and E. crus-galli weed species for current conditions in Germany. The Habitat becomes more suitable from lighter to darker colored map areas.

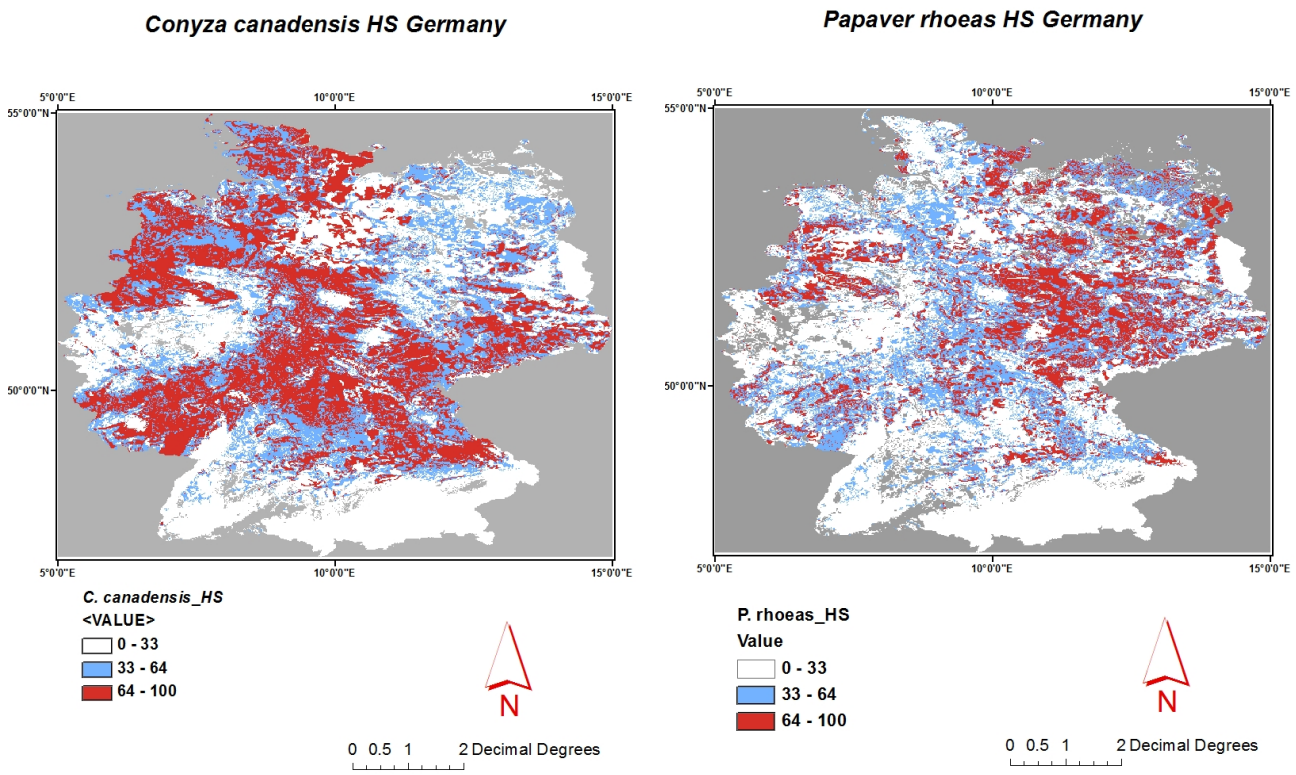


Figure 11b. Habitat Suitability maps for C. canadensis and P. rhoeas weed species for current conditions in Germany. The Habitat becomes more suitable from darker to lighter colored map areas.

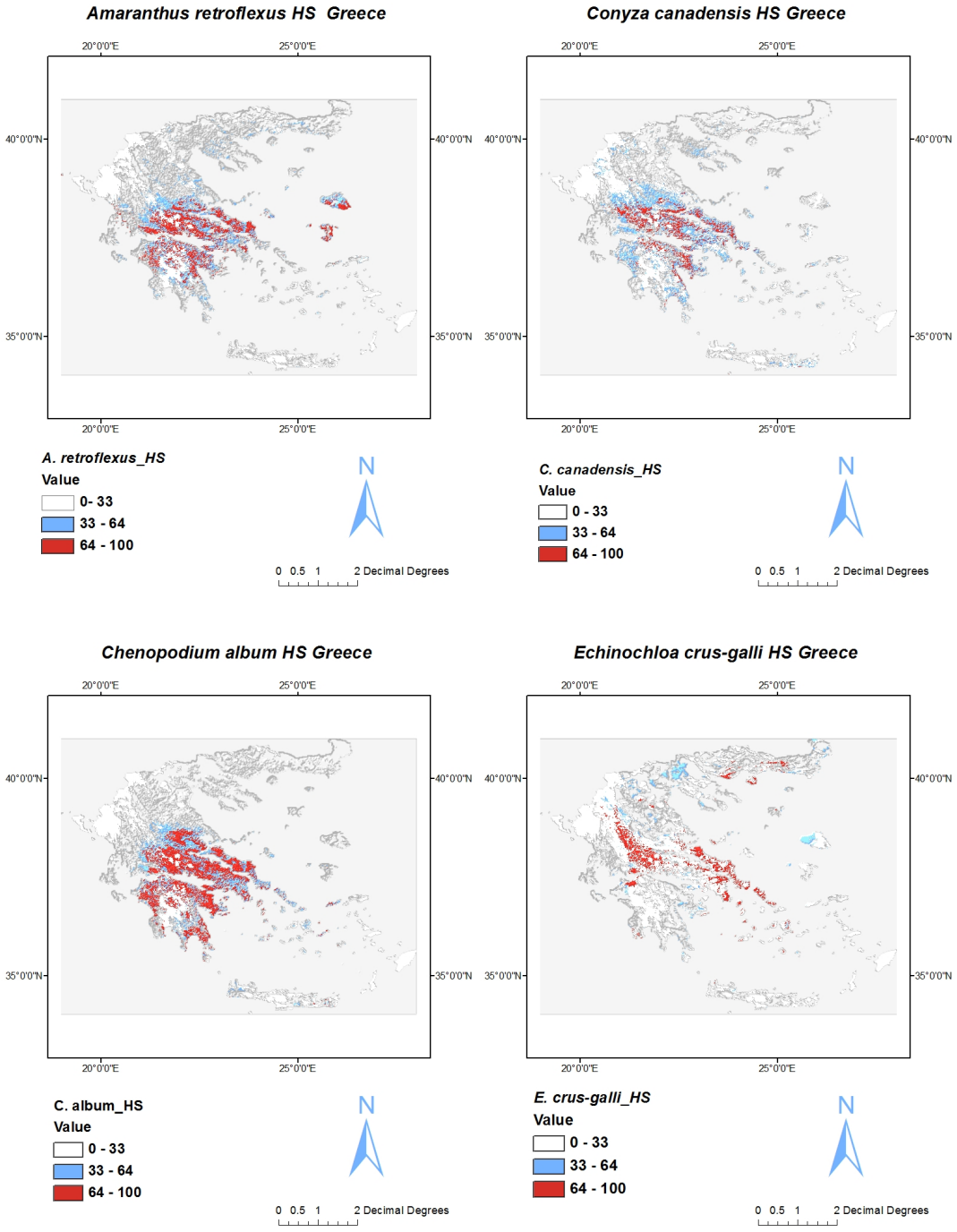


Figure 12a. Habitat Suitability maps for A. retroflexus, C. album, C. canadensis and E. crus-galli weed species for current conditions in Greece. The Habitat becomes more suitable from lighter to darker colored map are

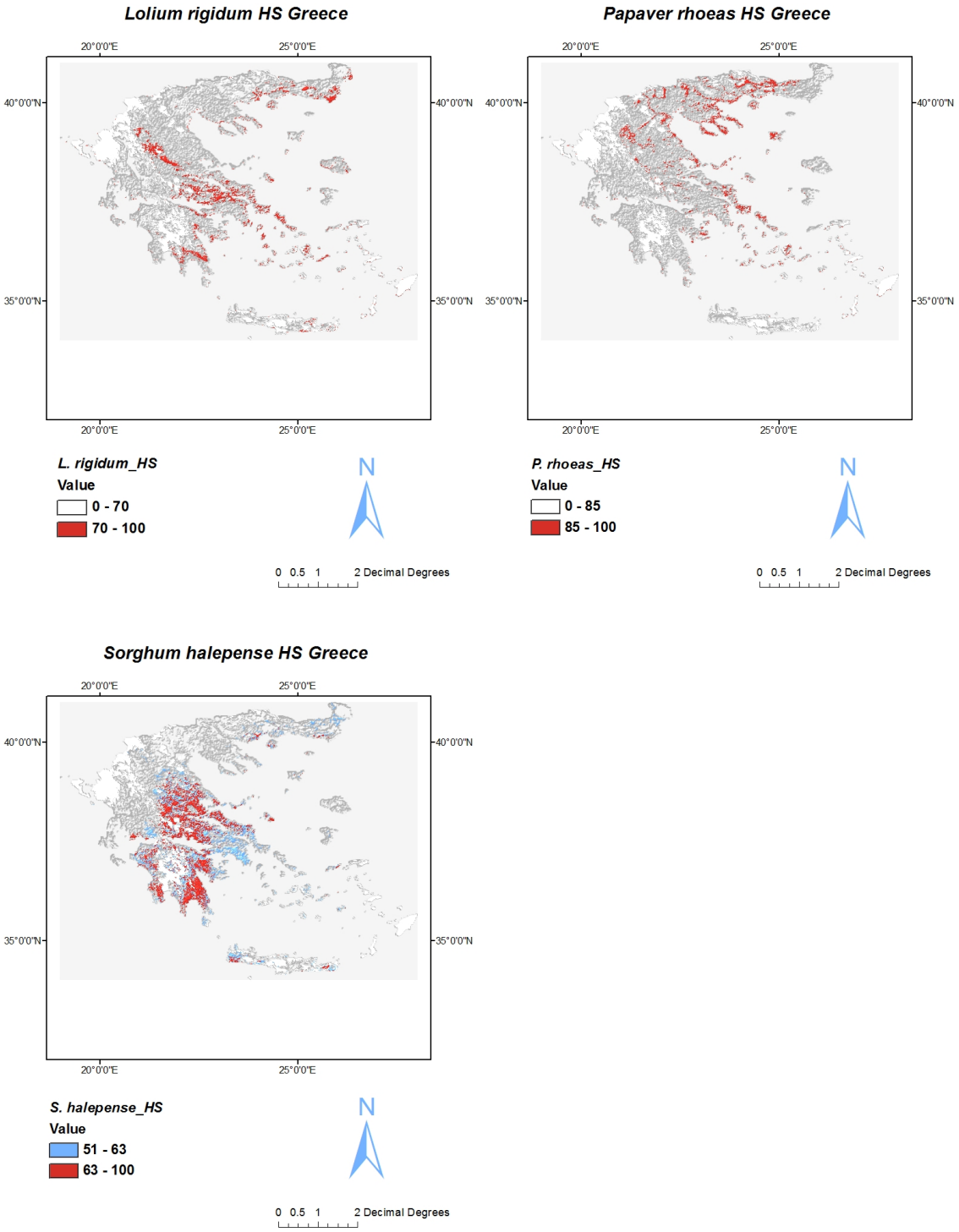


Figure 12b. Habitat Suitability maps for P. rhoeas, L. rigidum and S. halepense weed species for current conditions in Greece. The Habitat becomes more suitable from lighter to darker colored map areas.

According to the cross-validation displayed in table 7, the Boyce index is above 0.5, for almost all studied weed species in Greece, with the exception of *C. canadensis*, *, L. rigidum* and at a lesser extent of *E. crus-galli*, with Boyce index values close to zero, indicating that the model is not different from a chance model. Concerning *P. rhoeas* in Greece*,* the Boyce index is negative that might indicate an incorrect model. However, the cross-validation CVI index is over 0.3, for all species in Greece. *L. rigidum* and *P. rhoeas* are excepted, since the corresponding CVI value is much lower.

All species in Germany have low Boyce (near to zero) and also CVI (0.2 to 0.3) values, with the exception of *A. retroflexus* and *P. rhoeas* have higher index values.

Table 7. Cross-validation of Habitat Suitability Models, for all species in both countries. Boyce and Contrast Validation Index (CVI) are given, accompanied by their standard deviation (s.d.) values.

	Boyce index	**Boyce std**	**CVI**	**CVI std**
A. retroflexus-gr	0.692	0.078	0.482	0.086
A. retroflexus-ger	0.380	0.493	0.214	0.162
C. album-gr	0.601	0.189	0.463	0.126
C. album-ger	0.209	0.245	0.152	0.089
C. canadensis-gr	0.019	0.549	0.485	0.409
C. canadensis-ger	0.083	0.545	0.17	0.19
E. crus-galli-gr	0.317	0.472	0.48	0.297
E. crus-galli-ger	0.134	0.507	0.188	0.168
P. rhoeas-gr	-0.355	0.462	-0.264	0.275
P. rhoeas-ger	0.503	0.424	0.273	0.155
L. rigidum-gr	-0.093	0.399	-0.21	0.291
S. halepense-gr	0.480	0.347	0.473	0.16

4.3 GARP Models

Based on the GARP maps of figure 13a, in 2000 *C. album* occurs at the Southern part of Greece, such as Peloponnisos, and at areas around Attiki (Sterea Ellada - central Greece). Populations of the same species, are also found at East Macedonia and Thrace (fig. 13a). The scale of of all GARP maps indicates that the species distribution ranges between low (0 percent) and high (100 percent) probability values. The species is mostly distributed at dry areas with high temperatures (fig. 3). Concerning, *C. canadensis*, and *A. retroflexus,* they are found in smaller populations, almost at the same areas with *C. album*. Moreover, *E. crus-galli* is distributed with smaller population

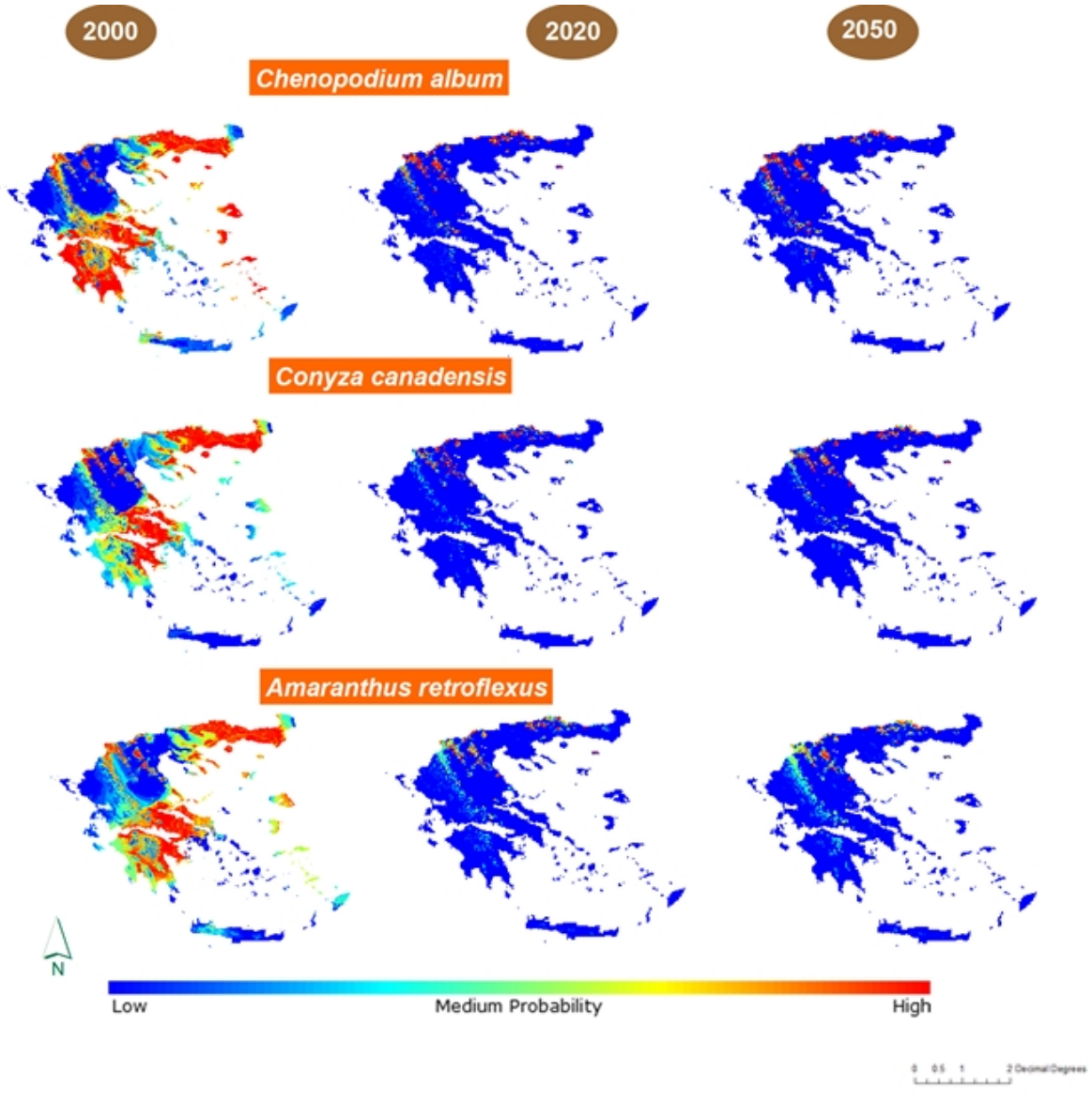


Figure 13a. GARP maps for C. album, C. canadensis, A. retroflexus in Greece, using annual mean temperature and annual precipitation, as environmental factors in 2000, 2020, and 2050.

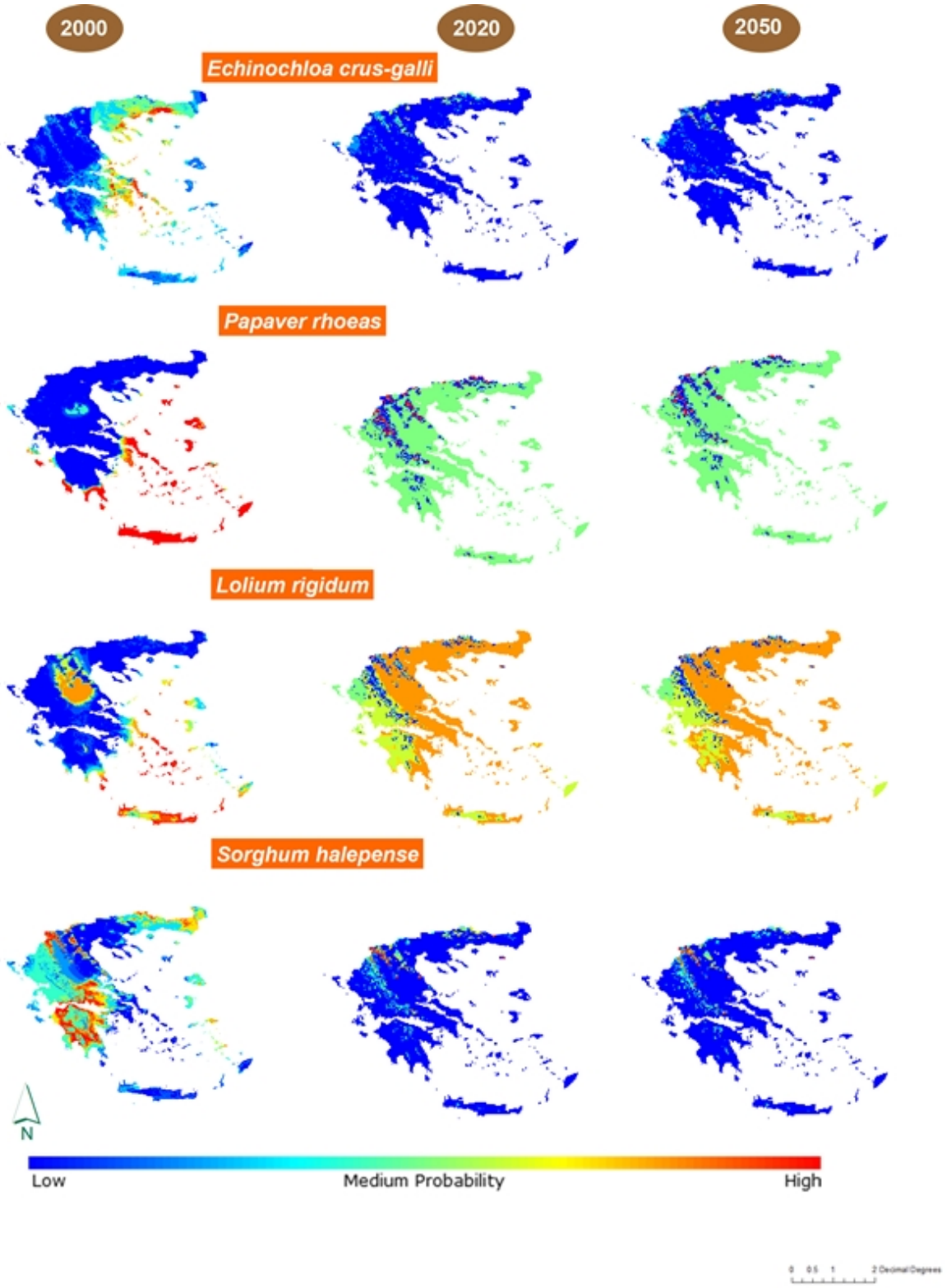


Figure 13b. GARP maps for E. crus-galli, P. rhoeas, L. rigidum, S. halepense in Greece, using annual mean temperature and annual precipitation, as environmental factors in 2000, 2020, and 2050.

sizes around Attiki, Central - East Macedonia and Thrace (fig. 13b), mainly at dry - hot low land areas (fig. 3). Additionally, *P. rhoeas* is found mostly around the Aegean islands, and artificial areas, such as Athens, the capital of Greece(fig. 13b). *L. rigidum* seems to expand at the same areas with *P. rhoeas*, included a low land agricultural area, Thessaly, but in smaller populations. Finally, *S. halepense* seems to be distributed at low land areas of Peloponnese and also, however, with lower probability, at the western wetter part of Greece (figs 13b and 3). Therefore, most of the studied species in Greece, prefer to loccur at dry - hot low mand areas, with the exception of *S. halepense that* seems to prefer wet high temperature areas of low altitude (fig. 3).

Projecting species distribution in 2020 and 2050, all species in Greece, with the exception of *P. rhoeas* and *L. rigidum*, tend to move to higher altitudes (figs 13b and 5), having extremely low population sizes. On the contrary, *P. rhoeas* tends to expand with a medium probability all over Greece by 2020 and 2050 (fig. 13b). The same holds for *L. rigidum*, having though a higher probability of expanding to the drier and more hot east part of Greece (figs 13b and 5). Therefore, in future climatic conditions in Greece, characterized by higher minimum temperature compared to present conditions (figs. 3 and 5) most of the studied species, are limited, except from *L. rigidum* and *P. rhoeas*. Based on table 8, almost all GARP models for species distribution in Greece, in 2000, are 71 to 90 % accurate, The Area Under Curve (AUC) for all models is over 0.75 for the species in Greece.

Table 8. Cross - validation of GARP analysis for all species studied in Greece and Germany. Area Under Curve (AUC) for the whole model (species in 2000, 2020 and 2050) and Accuracy of observed values (percent of all values) for the species in 2000 is recorded.

Species	Greece		Germany	
	AUC	Accuracy (%)	AUC	Accuracy (%)
Amaranthus retroflexus	0.76	86.96	0.64	92.11
Chenopodium album	0.75	93	0.61	88.22
Conyza canadensis	0.8	89.47	0.61	94.85
Echinochloa crus-galli	0.93	71.43	0.67	89.10
Papaver rhoeas	0.8	63.27	0.67	81.57
Lolium rigidum	0.84	84.2		
Sorghum halepense	0.82	81.25		

In Germany, current distribution of *A. retroflexus* is with higher probability at lowland areas of the Northern half of Germany (figs 4 and 14a). However, there is also probability, lower though, to find the species in other areas of the country, since the species seems to have a wide distribution across the country (fig. 14a). Projecting the distribution of *A. retroflexus* in Germany in 2020, the species are mostly found at the southern part of Germany, where the altitude and the precipitation is higher, but the temperature lower, compared to northern Germany (figs 14a, 4 and 6). Finally, the same species in 2050, tends becomes limited (fig. 14a). Concerning, *C. album* and *C.*

canadensis, they are found with higher probability at the western part of Germany (fig. 14a). The probability of *C. album* and *C. canadensis* distribution (fig. 14a), is very similar to the current mean temperature pattern of Germany (fig. 4). Both species in 2020, seem to land up almost the whole country of Germany (fig. 14a). However, in 2050, *C. album* seems to occur across the country, but the probability is quite low. On the contrary, there is a higher probability for *C. canadensis* to become expanded across Germany, except from the western part of the country (fig. 14a). In 2050, the whole Germany, becomes warmer, especially the Southern areas (fig. 6).

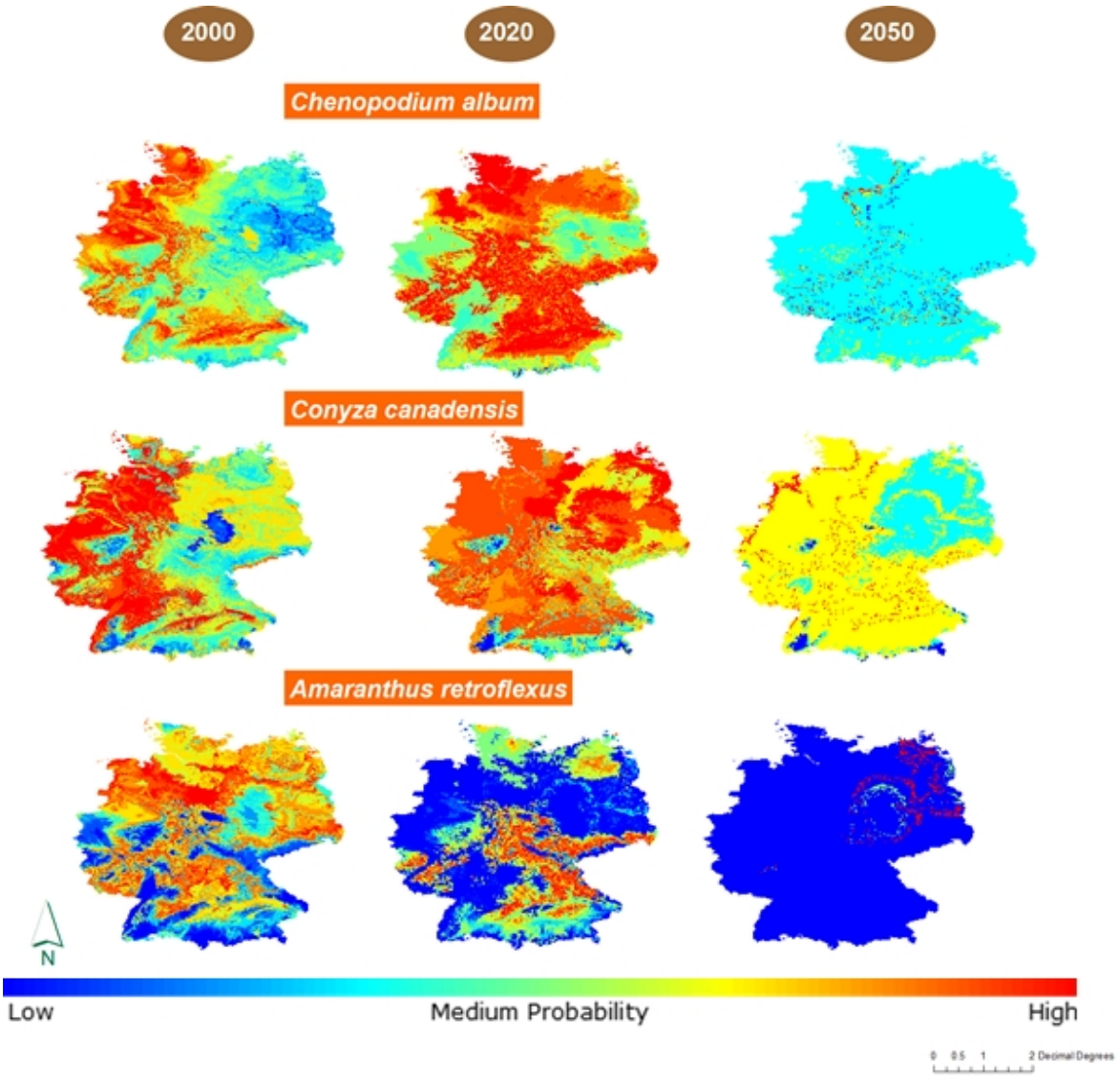


Figure 14a. GARP maps for C. album, C. canadensis, A. retroflexus in Germany, using annual mean temperature and annual precipitation, as environmental factors in 2000, 2020, and 2050.

As far as it concerns *E. crus-galli* in Germany (fig. 14b), it is found only at the northern part of Germany in 2000. Projecting its' distribution in 2020, the species occurs mainly at the southern part of Germany, where the altitude is higher and the temperature lower (fig. 6). In 2050, in Germany, the species distribution is very limited and it is almost vanished, becoming locally extinct. The main factor contributing to this change could be the warming of the country by 2050 (fig. 6). Following the distribution of *P. rhoeas* in 2000, the species has a medium probability to be found almost across Germany. Additionally, there is a higher probability to find the species at the western cooler and wetter part of Germany (figs 14b and 4). In 2020, the pattern of distribution is similar, but the higher probabilities of distribution at the western part are decreasing (fig. 14b). Finally, in 2050, the probability of finding the species across the country becomes lower (fig. 14b).

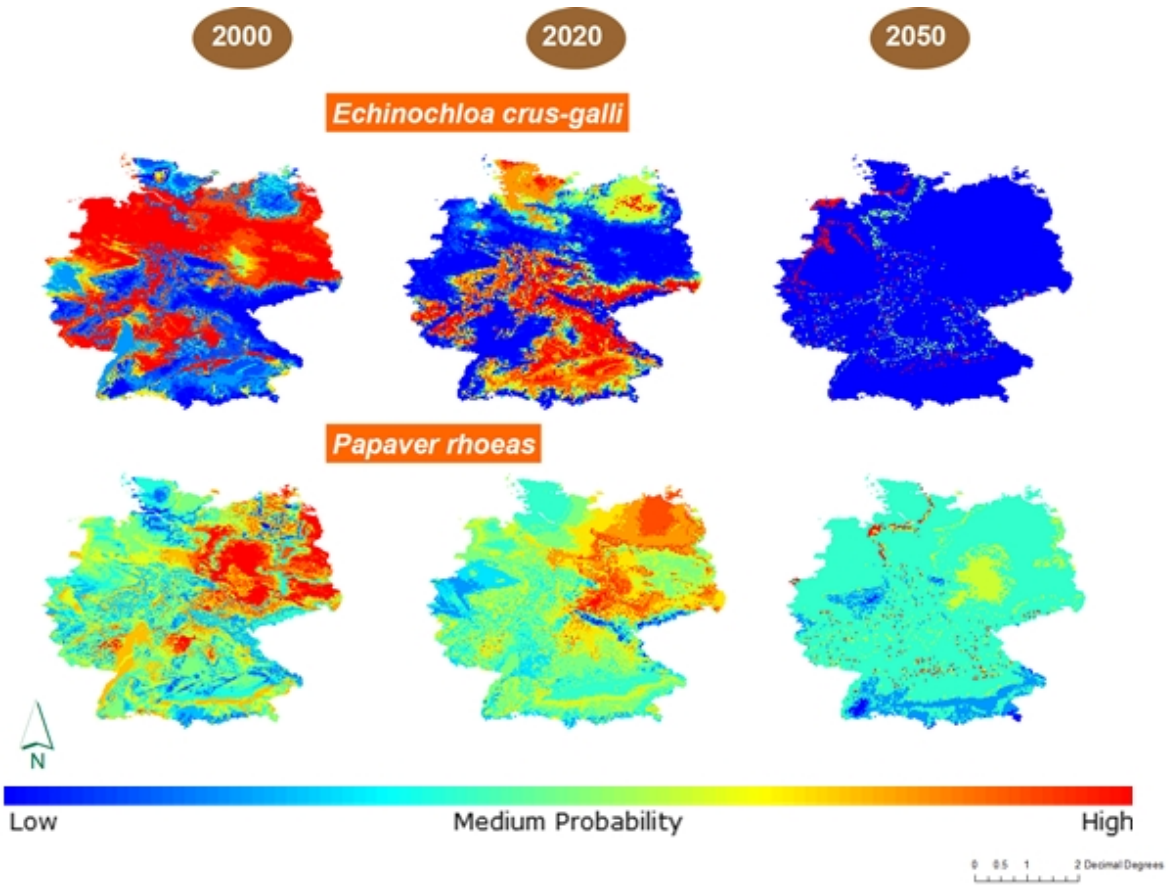


Figure 14b. GARP maps for E. crus-galli and P. rhoeas in Germany, using annual mean temperature and annual precipitation, as environmental factors in 2000, 2020, and 2050.

Based on table 8, almost all GARP models for species distribution in Germany, in 2000, are 82 to 95. However, the Area Under Curve (AUC) for all models is between 0.6 and 0.67 for the species in Germany, indicating models better than random.

4.4 Rank - abundance diagrams

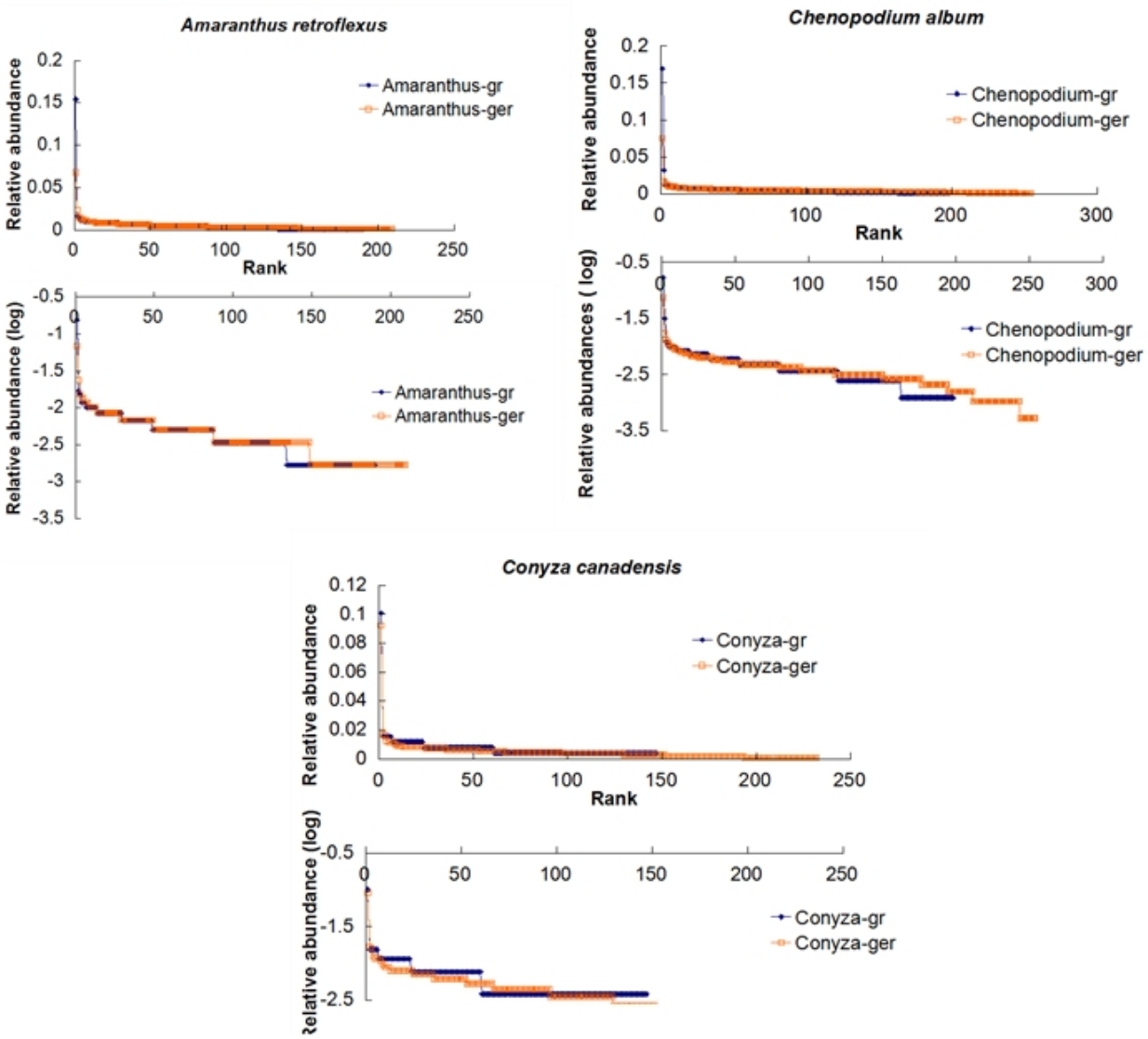


Figure 15a. Rank - abundance diagrams of mean NDVI value of 3 km buffer zone around species points landing both in Greece and in Germany. Log values of relative abundance are also given. The species presented are A. retroflexus, C. album and C. canadensis.

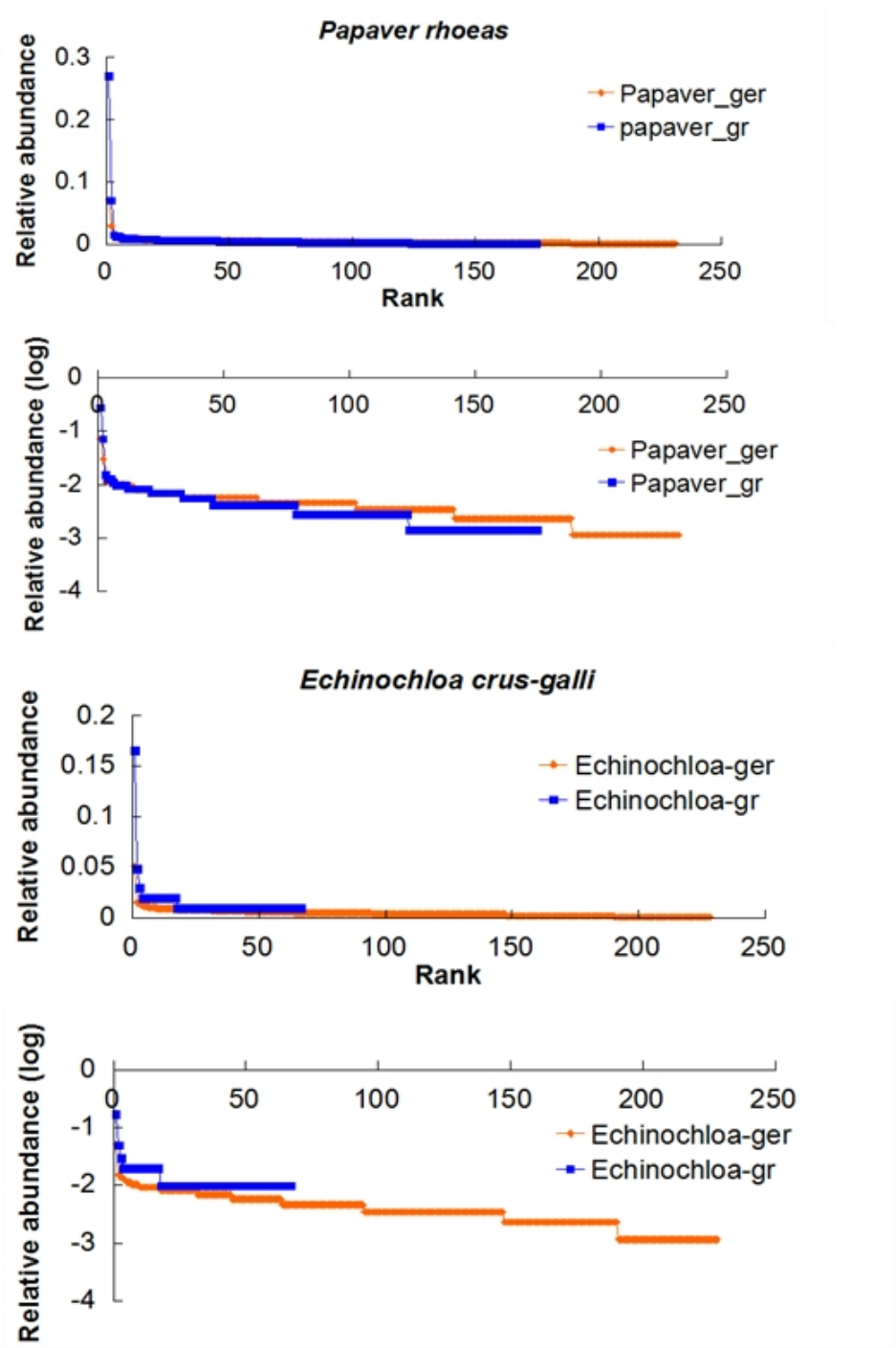


Figure 15b. Rank - abundance diagrams of mean NDVI value of 3 km buffer zone around species points landing both in Greece and in Germany. Log values of relative abundance are also given. The species presented are E. crus-galli and P. rhoeas.

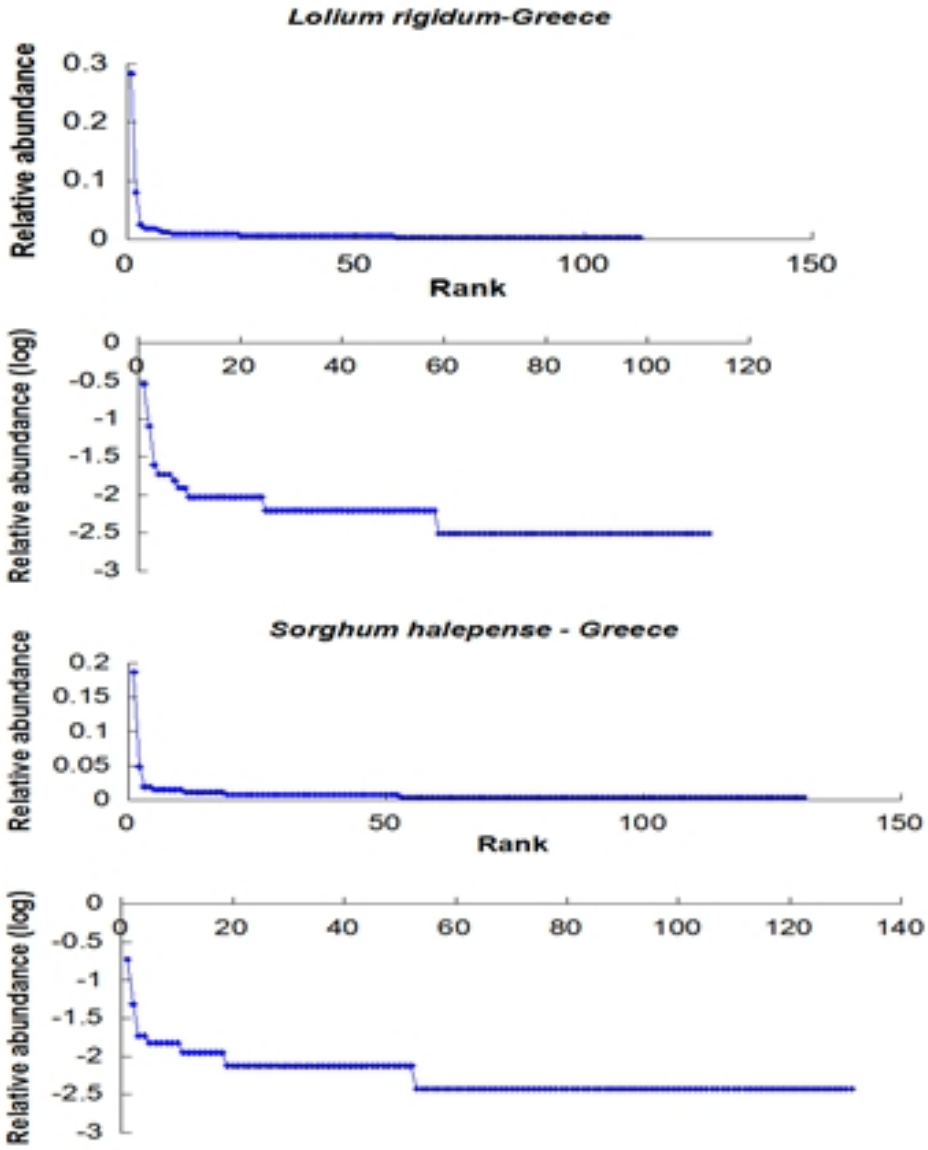


Figure 15c. Rank - abundance diagrams of mean NDVI value of 3 km buffer zone around species points landing in Greece. Log values of relative abundance are also given. The species presented are L. rigidum and S. halepense.

The relative abundance of different pixels (DNvalues) around species points is ranked. Moreover, DN abundance was plotted even on a log scale in order to decrease the outlier effect, that is, the smoothing of the curve due to hyper-dominant DNvalues. The tails of the curve of rank-abundance diagrams, provide information of the abundance of the dominant and rare values. The left-side of the tail indicates dominant values, whereas the right-side the rare values. Concerning the log curves, the more sigmoid the curve, the more heterogenous the landscape is.

All curves are strongly right-sided, indicating equitability of the distribution of the abundances of DN values (figs 15a, 15b and 15c). Another common characteristic of the above rank - abundance diagrams, is the one or two dominant values in every species

distribution. Moreover, dominant DN values correspond to very low NDVI values (figs 15a, 15b and 15c).
The relative abundance, and also log curves for *A. retroflexus* in both countries follow the same pattern (fig. 15a).On the contrary, the log curves of *C. album*, *C. canadensis* and *P. rhoeas* landing in Germany seem to be longer and more sigmoid related to the Greek ones (figs 15a and 15b). However, concerning *E. crus-galli*, the relative abundance curve and log-curve of the species in Germany is much longer, and more sigmoid compared to the same species in Greece (fig. 15b). This could indicate, a smaller DN diversity of NDVI values around the species in Greece (table 9), and therefore a lower spectral heterogeneity.
Finally, the studied weed species in Germany seem to land in more diverse landscape (table 9). Among all species, *E. crus-galli*, *L. rigidum*, and *S. halepense* in Greece, hold the lowest spectral diversity, Shannon and Pielou index (table 9). On the contrary, all species in Germany, and also *C. canadensis*, and *E. crus-galli* show the highest *H* value (table 9).

Table 9. Diversity, Shannon and Pielou index, using the NDVI Digital Number (DN) values around species points in a buffer zone of 3 km.

Country	*Species*	*Diversity*	*H*	*J*
Greece	***A. retroflexus***	190	4.71	0.89
	C. album	198	4.67	0.88
	C. canadensis	147	4.7	0.94
	E. crus-galli	67	3.87	0.92
	P. rhoeas	175	4.08	0.79
	L. rigidum	112	3.73	0.79
	S. halepense	131	4.28	0.88
Germany	***A. retroflexus***	209	5.04	0.94
	C. album	254	5.2	0.94
	C. canadensis	232	5.04	0.93
	E. crus-galli	228	5.17	0.95
	P. rhoeas	231	5.12	0.94

4.5 Ordinary Least Square (OLS) regression

According to table 10, most of the species distribution does not seem to be related with NDVI index, since the Rsquare is near to zero, and the models are mostly not statistically significant (f-probability, non significant). The Rsquare for *E. crus-galli* is 0.33, but the test is not significant. Moreover, the Morans' I analysis applied on the residuals of the OLS regression, indicated that either there is spatial autocorrelation inside the species or more predictor variables are needed to explain the species distribution. Additionally, there is a high probability that the relationships are not linear. Normal distribution of the residuals is recorded in Germany for *A. retroflexus* and *P. rhoeas*, and in Greece for *C. canadensis*, *E. crus-galli* and *P. rhoeas*. Statistically significant models, according to f-prob, are recorded for *A. retroflexus* in Greece, with an Rsquare 0.16, though the residuals are clustered. In Germany, *C. album* distribution seems to be statistically significant (f-prob and t-prob) and positively related with a low Rsquare of 0.10 to NDVI index. However, the distribution of the residuals in this case is not normal.

Finally, significant is the t-test probability, in Greece for *S. halepense* (clustered residuals), *L. rigidum* (clustered res.), *C. canadensis* (normal res.), *P. rhoeas* (normal res.) both in Greece and Germany, and *E. crus-galli* (clustered res.), *C. album* (clustered res.), and A. retroflexus (*normal res.*) in Germany. All the relationships resulted in very low Rsquare values. Moran's I autocorrelation evaluates whether the pattern of the relationship expressed is clustered, dispersed or random. Normal distribution refers to random distribution of the residuals of the analysis. For the Moran's I statistic, the null hypothesis states that the attribute being analyzed is randomly distributed among the features in the study area, i.e. the spatial processes promoting the observed pattern of values is random chance. The values of the dataset tend to cluster spatially, when high values cluster (group) near other high values or/and low values cluster near other low values.

Table 10. Ordinary Least Square regression model of species distribution in Greece, and in Germany, using as predictor, mean and standard deviation NDVI values, in an area of 3km around species points. R^2, coeff., t-stat, t-prob, f-stat, f-prob, Morans' I z, Morans'I p and the residuals are also presented for checking spatial autocorrelation.

SPECIES	COUN-TRY	R^2	Coeff.	t-stat	t-prob	f-stat	f-prob	Morans'I z	Morans' I p	Residuals
Amaranthus retroflexus	GR	0.16	10.70	0.64	ns	4.35	0.02	8.32	0.00	clustered
	GER	0.14	140.30	3.00	0.01	2.68	ns	1.38	ns	normal
Chenopodium album	GR	0.07	28.00	1.96	ns	2.52	ns	7.62	0.00	clustered
	GER	0.10	103.10	6.40	0	6.73	0.00	0.28	0.00	clustered
Conyza canadensis	GR	0.26	41.8	2.82	0.01	2.75	ns	1.51	ns	normal
	GER	0.02	28.00	1.07	ns	0.84	ns	2.48	0.01	clustered
Echinochloa crus-galli	GR	0.33	64.40	1.62	ns	1.21	ns	0.36	ns	normal
	GER	0.00	52.20	2.04	0.05	0.05	ns	2.64	0.01	clustered
Papaver rhoeas	GR	0.02	36.30	2.86	0.01	0.51	ns	1.64	ns	normal
	GER	0.05	72.70	3.74	0.00	1.33	ns	0.48	ns	normal
Lolium rigidum	GR	0.01	54.00	2.36	0.03	0.09	ns	0.32	0.00	clustered
Sorghum halepense	GR	0.02	80.20	2.66	0.02	0.18	ns	2.37	0.02	clustered

4.6 Variograms

The highest spatial variability around species points is recorded for *P. rhoeas* in Greece, but also in Germany; however, with a lower sill. The spatial scale for this species is low in both areas, indicating that a smaller number of cells is needed to record the maximum spatial variability. Low mean NDVI length scale indicates that high spectral variability is recorded in a small buffer zone or spatial scale around species points. High NDVI image variability (sill) indicates high spectral heterogeneity around species points. According to the figure above, all the other species in both countries follow the pattern of a much lower spatial variability, and a higher mean NDVI length scale. Moreover, both length scale and sill for *A. retroflexus* are almost identical for the species in the two studied countries. The highest length scale in recorded for *S. halepense,* and the lowest spatial variation for *C. album* in Greece (fig. 16).

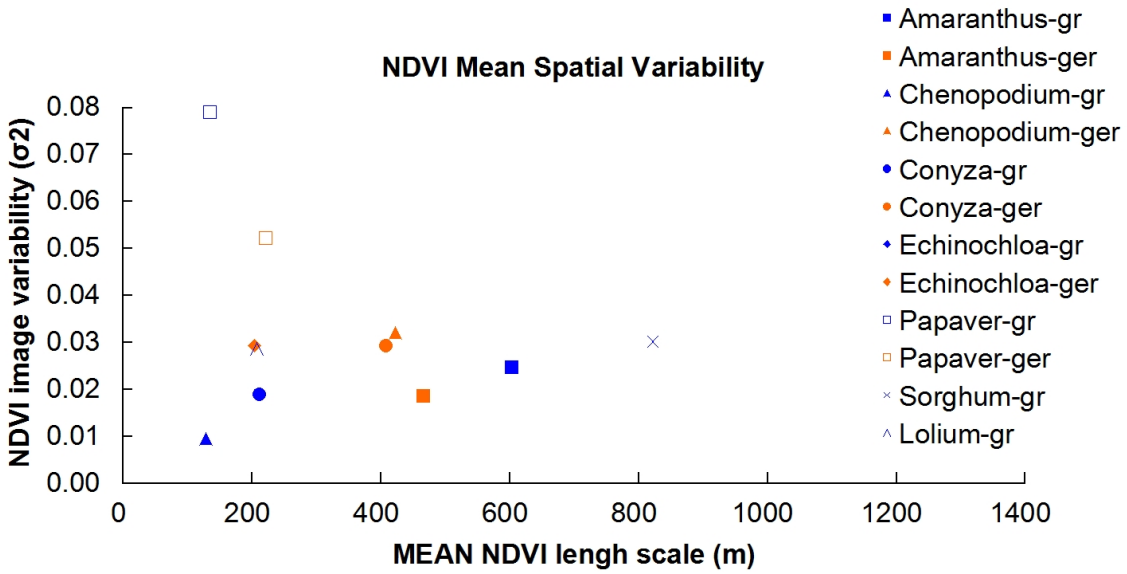


Figure 16. NDVI spatial variability (σ^2, variogram sill) versus the NDVI mean length scale (square root of the variogram integral range).

4.7 Weight of Evidence and Logistic regression

The relationship of NDVI with species presence is presented in table 11, for species in Greece and on table 12, for species in Germany. The relationship varies between the species. According to table 10, *E. crus-galli*, *L. rigidum* and *S. talepense* do not seem to have any relationship with NDVI. On the contrary, *C. canadensis*, and *P. rhoeas* seems to be related strongly (95 - 97.5 %) with the first class on the NDVI that corresponds to values 0 to 0.25. For both species,the contrast value is over 0.9 that could indicate that the class 0 to 0.25 of the NDVI could be a good predictor of species presence. Moreover, the logistic repression's p-value for the same class and species is also statistically significant (table 11). The second NDVI class, i.e. 0.25 to 0.5 (class 2) is also well related to the presence of *C. album* in Greece, with a level of significance between 99.5 and 100 percent (table 11). Moreover, the p-value of the logistic regression also indicates a statistically significant relationship (table 11). Finally, *A. retroflexus* in the Greek landscape is related with the third class of NDVI that gets values between 0.5 and 0.75, with a level of significance 90 to 95 percent; however, the relationship is not statistically significant according to the corresponding p-value of the logistic regression (table 11).

Table 11. NDVI relationships with species presence studied in Greece. WofE contrast values, their significance, Logistic Regression Post Probability values and their corresponding two-tailed p-values.

Species	Classes*	LRPost Prob	p-value (two-tailed)	CONTRAST	STUD_CNT	Level of sig. (%)
A. retroflexus	1	0.01	0.02	0.44	0.98	85-90
	2	0.01	0.02	-0.52	-0.70	<70
	3	0.01	0.10	**0.64**	1.46	**90-95**
	4	0.01	0.00	0.28	0.60	70-75
C. album	1	0.00	0.16	0.02	0.07	<70
	2	0.01	**0.01**	**0.98**	3.72	99.5-100
	3	0.02	0.00	-0.78	-2.10	<70
	4	0.01	0.00	-1.58	-2.20	<70
C. canadensis	1	0.01	**0.01**	**0.91**	1.91	95-97.5
	2	0.01	0.03	-0.95	-1.50	<70
	3	0.01	0.10	0.33	0.67	70-75
	4	0.00	0.33	-1.00	-0.97	<70
E. crus-galli	1	0.00	0.39	-0.45	-0.40	<70
	2	0.00	0.39	0.27	0.29	<70
	3	0.00	0.25	-0.31	-0.28	<70
	4	0.00	0.39	0.69	0.62	70-75
P. rhoeas	1	0.01	**0.01**	**0.94**	3.28	99.5-100
	2	0.01	0.00	-0.35	-1.01	<70
	3	0.02	0.00	-0.44	-1.14	<70
	4	0.01	0.09	-0.82	-1.37	<70
L. rigidum	1	0.01	0.02	0.14	0.28	<70
	2	0.01	0.04	0.00	0.00	<70
	3	0.01	0.02	0.18	0.34	<70
	4	0.01	0.17	-0.48	-0.64	<70
S. halepense	1	0.00	0.07	0.47	0.87	80-85
	2	0.01	0.11	-0.31	-0.52	<70
	3	0.01	0.03	-0.09	-0.14	<70
	4	0.00	0.34	-0.28	-0.27	<70

** 5 classes for NDVI index, 0 to 0.25 (1), 0.25 to 0.5 (2), 0.5 to 0.75 (3) and 0.75 to 1 (4)*

Table 12. NDVI relationships with species presence in Germany. WofE contrast values, their significance level, Logistic Regression Post Probability values and their corresponding two-tailed p-values.

Species	Classes*	LRPostProb	p-value	CONTRAST	STUD_CNT	Level of sig. (%)
P. rhoeas	1	0.00	**0.00**	0.03	0.08	<70
	2	0.00	**0.00**	**0.52**	1.30	**90-95**
	3	0.00	**0.00**	**0.72**	2.12	**97.5-99**
	4	0.00	**0.00**	-0.20	-0.44	< 70
C. canadensis	1	0.00	**0.00**	0.34	1.31	90-95
	2	0.00	**0.00**	0.18	0.60	70-75
	3	0.00	**0.00**	**0.46**	1.93	**95-97.5**
	4	0.00	**0.00**	0.21	0.77	75-80
E.crus-galli	1			**0.58**	1.79	**95-97.5**
	2			0.28	0.73	75-80
	3			0.33	0.99	80-90
	4			-0.03	-0.08	<70
C. album	1			**0.45**	1.95	**95-97.5**
	2			0.30	1.12	80-90
	3			0.25	1.06	80-90
	4			0.24	0.95	80-90
A. retroflexus	1			**0.62**	1.51	**90-95**
	2			-0.31	-0.50	<70
	3			0.47	1.15	80-90
	4			0.19	0.40	60-70

** 5 classes for NDVI index, 0 to 0.4 (1), 0.4 to 0.6 (2), 0.6 to 0.8 (3) and 0.8 to 1 (4)*

Proceeding with species in Germany, the highest relationship is recorded between *P. rhoeas* and NDVI class 3 (0.6 to 0.8 NDVI value), getting a contrast value of 0.7 with a 97.5 - 99 percent level of significance, and class 2 (0.4 to 0.6 NDVI value), getting a contrast value of 0.52 with a 90-95 percent level of significance (table 12). The logistic regression also results to statistically significant relationships (table 12).

Moreover, *E. crus-galli*, *C. album*, and *A. retroflexus* presence are related with the 1st class of NDVI (0 to 0.4), with a medium contrast value (0.45 - 0.62), and level of significance 95 to 97.5 percent for *E. crus-galli*, and *C. album*, and 90 to 95 percent for *A. retroflexus* (table 12). Finally, *C. canadensis* is well related with the 3rd class of NDVI (0.6 to 0.8), with 95 to 97.5 level of significance, and also statistically significant p-value derived from the logistic regression model (table 12).

Validating the logistic regression models (table 13), *P. rhoeas* in Germany. *C. canadensis* both in Germany and Greece, and *L. rigidum* and *S. halepense* in Greece, results into statistically significant models, since the Area Under Curve (AUC) is over 60 percent.

Table 13. CAPP curves equations fitted, Rsquare and AUC values for weed species in Greece, and also Germany.

Species	Country	Rsquare	Equation (y)	AUC (%)
P. rhoeas	***Greece***	0.87	0.0123x^2+ 0.062x+0.0023	48.82
	Germany	0.87	-0.0059x^2+0.0087x-0.0014	**76.3**
C. canadensis	***Greece***	*0.98*	0.0056x^2+0.0064x+0.005	58.92
	Germany	0.92	-0.0063x^2+0.0101x-0.0017	**69.04**
E. crus-galli	***Greece***	0.90	0.0005x^2+0.0022x+0.0004	53.76
A. retroflexus	***Greece***	*0.99*	-0.0026x^2+0.0192x-0.0037	44.65
C. album	***Greece***	*0.99*	0.0168x^2+0.0008x+0.0003	35.2
L. rigidum	***Greece***	0.99	0.0021ln(x)+0.0097	**78.77**
S. halepense	***Greece***	0.89	0.0011ln(x)+0.0057	**80.92**

5. Discussion

5.1 Landscape heterogeneity & species presence

The landscape is more heterogenous (spectral heterogeneity) for the species found in Germany than in Greece, especially for *P. rhoeas*. A number of different models can be used to study landscape heterogeneity in relation to species presence. The results from all models used were not contradictive, with the exception of OLS regression and HS models that did not indicate significant relationships. Among the models used, WofE and Variograms seem to be the most reliable spatial models that can be used to study NDVI, landscape heterogeneity and species presence relationships.

5.1.1 Rank-abundance diagrams & Diversity Indices

Rank - abundance diagrams were created so as to speculate the heterogeneity of the landscape around species points. The main pattern of the curves for almost all species in both countries is the one or two dominant values and the evident right-sided tails. The right-sided tails combined with the sigmoid shape of the log curves is an indication of a large number of unique rare values, related to heterogenous landscapes or landscapes with high spectral heterogeneity.

More specifically, for *A. retroflexus*, the spectral heterogeneity of the landscape around species seems to be similar in both countries. On the contrary, the landscape around *C. album*, *C. canadensis, P. rhoeas* and *E. crus-galli* seems to be more heterogenous in Germany, since the curves are longer and slightly more sigmoid. Additionally, all species in Germany have high landscape diversity (NDVI Shannon and Piellou index values). Lower spectral diversity, according to diversity indices, was recorded for *E. crus-galli*, *L. rigidum*, and *S. halepense* in Greece. The low spectral diversity for these species comes in accordance with the regression (OLS and Logistic regressiion) and WofE results, where the species presence were not related with NDVI. An explanation for the low spectral heterogeneity around these three species in Greece could be the fact that they are mostly grown in monoculture rice and sorghum fields, and not in more diverse landscapes.

Moreover, based on Rank-abundance diagrams, the studied weed species in Germany seem to land in more diverse landscape. Additionally, all species in Germany have high landscape diversity (NDVI) values (Shannon and Piellou indices), and most of them generally are related with more dense areas compared to Greece, according to Logistic regression and WofE models. On the contrary, according to variograms, all species studied, except from *P. rhoeas,* have a similar NDVI variability pattern. *P. rhoeas* has the strongest relationships with NDVI and the highest spectral variability around it (variograms).

5.1.2 Variograms & heterogeneity

According to the variograms, the highest spectral variability is recorded around *P. rhoeas* both in Greece and Germany. The spatial scale for this species is low in both areas, indicating that a smaller number of cells is needed to record the maximum spatial variability. *P. rhoeas* is a species with a wide distribution across countries. As a species, it prefers calcareous areas and soils (Emorsgateseeds 2014), therefore it can be

expanded easily into the Greek landscape. The lowest variability of NDVI index, is recorded for *C. album* and *C. canadensis* in Greece. The same holds for the species in Germany, with slightly a few variations. Therefore, a clear difference of NDVI index variability is recorded only for *P. rhoeas.* All other species studied have a similar NDVI index variability pattern.

5.2 NDVI & species presence

Most of the studied weed species in both countries are related with NDVI. However, not all models applied indicated significant relationships. Habitat Suitability and Ordinary Least Square Regression Models resulted to non significant relationships, compared to Logistic Regression and WofE Models. The relationships were more obvious for the weed species in Germany, especially for *P. rhoeas* that was related with high NDVI values, i.e. dense vegetated areas.

5.2.1 Habitat Suitability Models

According to Habitat Suitability models there is not an important relationship of NDVI with species presence. However, this could be as a consequence of using a large number of variables to describe species presence. In these models, the bioclimatic factors, especially mean temperature and annual precipitation are the main factors affecting species distribution.

5.2.2 OLS regression models

The distribution of most of the studied species in both countries does not seem to be related to NDVI index, since the Rsquare is near to zero, and the models are mostly not statistically significant (f-probability, non significant). Statistically significant relationships, but of low strength, have been found for *A. retroflexus* in Greece (R^2=0.16), and *C. album* in Germany (R^2=0.10); however, the residuals were not normal distributed. Further, in most cases the residuals distribution of the OLS regression is clustered, indicating that either there is spatial autocorrelation or more predictor variables are needed to explain the species distribution. Finally, there is a high probability, the relationship between species presence and NDVI not to be linear.

5.2.3 Logistic regression & WofE

On the contrary Logistic regression and WofE models resulted to statistically significant relationships of weed species with NDVI values, comparing to OLS models. However, significant relationships were not recorded for all species. More specifically, *E. crus-galli*, *L. rigidum* and *S. halepense* in Greece do not seem to have any relationship with NDVI. The species with the strongest relationships with NDVI is *P. rhoeas* landing in Germany, found in dense vegetated areas. The same species in Greece is found in very low vegetated areas. Generally, most of the species studied are related with low NDVI values. The weed species studied in Greece, are mostly found in low vegetated areas, whereas the same species in Germany are related with more dense vegetation. These differences could be the consequence of the differing bioclimatic conditions, resulting to differing vegetation types between the two countries; dry and hot, sparse, but more diverse vegetation in Greece, more temperate and wet, more dense vegetation in Germany. Finally, the logistic regression has a low predictive performance in Greece,

according to validation results. This can be contributed to the primarily clustered distribution of the species in the Greek landscape. Most of the species point data used for the study are mainly found at the Central and Southern part of the country. However, the analysis was performed for the whole country area.

5.3 Environmental Factors & weed species distribution

In Greece, all weed species studied tend to live in extreme habitats throughout the study area (high marginality value). The lowest marginality for species in Greece is recorded for *P. rhoeas*, whereas the highest for *C. canadensis*, *C. album*, and *E. crus-galli*. *P. rhoeas* is a species that prefers to live in calcareous, dry and hot areas, a condition that is very common in the Greek landscape.

Additionally, *C. canadensis* and *C. album* are summer annual C_3 plants, whereas *E. crus-galli* is a summer annual C_4 grass that prefers to live in dry hot, and mainly artificial areas. On the contrary, marginality for all studied species landing in Germany gets a low value that indicates that the same species in Germany, tend to live in average conditions. Moreover, the tolerance value for all species both in Greece and in Germany does not get a very high value. A low value of the global tolerance (close to 0) indicates a "specialist" species ending to live in a very narrow range of conditions, whereas a high value (close to 1) indicates a species that is not too picky on its living environment. This could indicate that all weed species studied neither prefer to live in a narrow range of conditions nor to be specialists in their environment. This condition is more evident for species in Germany, since the environmental conditions are less extreme than Greece, where the climate is relatively more dry and hot, especially during the summer, consisting a stressful environment for the species. This could be an explanation for the higher species marginality in Greece.

Additionally, the main factor explaining species marginality in Greece, according to the eigenvalues resulted from models applied, is mean temperature. Secondary factors in Greece, are altitude for *E. crus-galli* (C_4 summer annual monocot grass) and *S. halepense* (C_4 summer perennial monocot grass) and annual precipitation for *L. rigidum* (C_3 winter annual grass) and *C. canadensis (*C_3 summer annual dicot).

Therefore, in Greece C_4 plants marginality is not related to water availability, whereas C_3 plants marginality is. These results could come in accordance with Ozturk et al. (1981), found that the productivity of the C_4 species declined at the highest level of water availability, while the C_3 species were least productive under dry conditions. However, the study of Ozturk et al. (1981), was done for species in Peru, where the climate is very diverse, with a large variety of climates and microclimates.

Concerning Germany, the main factor affecting species lower though recorded marginality, is altitude, a second factor only for the species *E. crus-galli* is annual precipitation. This result comes in accordance to Patterson (1995a), described *E. crus-galli* as more competitive under moist conditions.

Hence, the climatic conditions are the main factors related to species marginality in Greece, whereas altitude in Germany. The contribution of climate to species distribution is 70 % for *A. retroflexus* and *P. rhoeas*, according to Hyvönen et al. (2012). Concerning *Chenopodium vulvaria* and *Chenopodium hybridum* the climate contribution to dispersal is 90 % and 55 %, correspondingly, according to the same author.

In addition, the abundance of C_4 (including Amaranthaceae) plants of total C_4 dicots,

total C_4 monocots and C_4 Poaceae, was related to the climatic variables of annual mean daily temperature, annual precipitation and DeMartonne's aridity index, according to Pyankov et al. (2010). Moreover, the abundance of total C_4 plants monocots, C_4 Poaceae decreases with increasing temperature but does not with precipitation (Pyankov et al. 2010). This comes in accordance with our results.

C_4 photosynthesis is a biochemical mechanism of a special mode of photosynthesis for concentrating CO_2 internally. Therefore, C_4 plants can still operate well with partially closed stomata. They generally have high water use efficiency (Black, 1973). It is intriguing that monocots respond much more to the temperature factor and, conversely, dicots to the precipitation factor of the aridity index. Grasses often possess morphological and anatomical traits, such as narrow rolled leaves and epidermal cavities protecting stomata under stress of low water availability and thick cuticles. These contribute to controlling transpiratory water loss when water is limiting, whereas they do not help to avoid temperature stress. In contrast, lower transpiration reduces transpirational cooling and this increases temperature stress. Thus, the stress adaptive C_4 model should prove to be less in demand in relation to water stress and more important in relation to temperature stress, which may explain why the abundance of C_4 monocots increases more in response to temperature than precipitation. Conversely, broad-leaved dicots may suffer more immediate stress as a result of transpiration when precipitation is low and the abundance of C_4 dicots responds more directly to transpiration (Pyankov et al. 2010).

5.4 Projecting future weed species distribution

Temperature is a major factor governing the seasonal growth of weeds and their geographic distribution (Woodward 1988). Davis et al. (2005) pointed out that there are three biotic responses to climate change, with all three involving evolution: (i) persistence if species tolerance limits are still within the changed climate parameters, (ii) range shifts (migration) to allow organisms to continue to thrive within their tolerance limits or (iii) extinction.

The first step in the risk assessment of weed range expansion is predicting the probability of successful establishment of weed species by matching climate data (Hyvönen et al. 2012). As far as it concerns climatic changes, the temperature increases are predicted to be greater in winter than in summer and greater in the Northern Hemisphere. The range expansion of many weeds into higher latitudes may accelerate with global warming (Patterson 1995b). Projections for the range shifts of European plant species due to climate change have indicated that the greatest species loss will occur in Mediterranean regions, and the least in Boreal, northern Alpine and Atlantic regions (Bakkenes et al. 2002; Thuiller et al. 2005).

In 2000, *A. retroflexus* is found in dry hot low land areas, and by 2020, and further by 2050, the species tends to move to higher altitudes, having very small population size. In Germany, the same species in 2000 lives in landscapes with average temperature and precipitation rates, whereas in 2020, it moves at higher altitudes, where the temperature is lower and the precipitation higher. By 2050, the species in Germany becomes extinct.

A. retroflexus and *E. crus-galli* are C_4 summer annual weeds, that prefer low land areas with average or low precipitation rates, and average or high temperatures, in Central and

Southern Europe, correspondingly. By 2020, these species tend to move to higher altitudes, whereas the species in Germany, by 2050 decrease their population size.
S. halepense is a C_4 summer perennial weed. In 2000, the species is found at low land areas of Greece, but also at the western more wet part of Greece. According to Pyankov et al. (2010), only the abundance of C_4 dicots is correlated with precipitation. By 2020 and 2050, the species moves to higher altitudes with very small population size.
The abundance of C_4 plants decreases with increasing temperature and expression of aridity (decreasing aridity index) (Pyankov et al. 2010). Moreover, according to a study by Parker-Allie et al. (2009), there was seemingly greater disparity among the C_4 grasses in their modeled distributions with climate warming. However, all species studied displayed habitat loss of relatively similar magnitude with climate warming and shifts in their distribution ranges into higher elevations. This was most apparent in the drier and warmer savannas. They also exhibited similar shifts in their distribution ranges into predominantly colder and moister grasslands at higher elevations in the countries interior.
Proceeding with C_3 weeds, in this study, *C. album* and *C. canadensis* (summer annual C_3 dicots) have the same pattern of distribution both as species in Greece and as species in Germany. Therefore, in Greece, they almost co-exist in dry hot low land and artificial areas. By 2020, and 2050 these species tend to move in much smaller populations at higher altitudes. Additionally, in Germany, these species in 2000 are found in hot low land areas, with average precipitation values. By 2020, they expand to almost all Germany, having a medium probability though. The same pattern is recorded for both species in 2050, with an exception of *C. canadensis* that is not found at the western most cold part of Germany. However, the probability of *C. canadensis* to expand all over Germany by 2050 is higher than for *C. album*. Compared to other studied species, *C. canadensis* in Germany might be a more tolerant species to climatic changes.
Concerning, C_3 winter annual plants, such as *L. rigidum* and *P. rhoeas*, the distribution pattern and the projections of the species do not follow the same pattern as C_3 summer annual plants. *L. rigidum* in Greece is found at agricultural areas, low land areas, and mainly in the Greek islands, hot and dry areas. By 2020, and 2050, the species has a higher probability of expanding at the drier and more hot east part of Greece. However, It is not found at high altitude areas. Additionaly, *P. rhoeas* in 2000 prefers dry and hot areas. By 2020, and 2050, the species tends to expand allover Greece, with lower though probability. In Germany, the same species prefers in 2000, low land areas, with average temperature, and low precipitation. The pattern is similar in 2020, but the probabilities of expansion become lower. By 2050, the species seems to expand allover Germany, but the prediction is not so strong.
Therefore, C_3 summer annual weeds, in central Europe seems to be tolerant to climatic changes, whereas in drier and more hot climates they tend to move to higher altitudes in small populations. On the contrary, the distribution of winter annual C_3 weeds in Greece, such as *L. rigidum* and *P. rhoeas* is not deteriorated by climatic changes.
According to Lee (2011), C_3 plants with a long growth period, such as *C. album*, they would not be able to avoid high temperature stress during late growth periods, such as the reproductive stage. Under conditions with an increase in temperature without elevated CO_2 levels, the biomass production of C_3 plants is expected to be strongly depressed by high temperature stress. However, because elevated CO_2 stimulates photosynthesis in C_3 plants during high temperature seasons, the loss in biomass and in

the reproductive index in response to elevated temperature would be greatly attenuated in C_3 plants under conditions with elevated CO_2. If they successfully compete during the initial establishment stage in early spring, their survival ability will be greatly enhanced by elevated CO_2 and high temperature stress. This could be an explanation of why C_3 winter annual weeds are more survival competitive than C_3 summer annual weeds, since their phenological stages are established much earlier than summer weeds. Therefore, C_3 summer annual weeds will not probably have a high chance of survival, while growing at late spring, they would not be able to avoid high temperature stress.
Northern Europe, as well as some mountainous areas, were climatically more suitable and southern Europe less suitable for the weed species studied. However, variation among species was substantial, indicating the importance of species-specific responses to climate change (Hyvönen et al. 2012). Finally, suitable climate conditions are the first prerequisite for the establishment of a permanent weed population. The accuracy of the projections could be improved by including the distribution of suitable cropping systems for the weed species (McDonald et al. 2009).

5.5 Comparing HS with GARP models

A main difference between these two models is that Habitat Suitability Models are factor analysis models, whereas GARP models combine variables to predict and project species distribution in current and future conditions. GARP models do not specify the factors affecting species distribution. Their analysis is related with the representation of species distribution on maps. Moreover, in these models, mainly bioclimatic conditions are used as explanatory variables. On the contrary, in Habitat Suitability Models a number of Eco-geographical variables are used to specify current species distribution and the factors affecting it, in an area of study.
Therefore, Habitat Suitability models seem to be more reliable than GARP models, since they can combine a larger number of variables, and also explain which are the main factors of species distribution. However, HS models cannot be used to speculate the species distribution in future climatic conditions, since HS models cannot perform projections in the future.
In this study, combining GARP and HS models, it is found that more marginal weed species tend to become locally extinct or to be found in small populations at higher altitudes. This is more obvious for species landing in the Greek landscape, with the exception of *P. rhoeas.* On the contrary, the same weed species in Germany are less marginal, and more tolerant to climatic changes. Therefore, species adapted strictly to certain climate conditions may lose those conditions and face increased extinction risk in a changing climate (Root et al. 2003; Thuiller et al. 2005). In contrast, species most likely to benefit from climate change in terms of extending range sizes are those with distributions independent of specific climatic conditions and land use patterns (Hyvönen et al. 2011).

6. Conclusions

The climatic conditions, especially mean temperature and annual precipitation, are the main factors affecting weed species distribution both in Greece and Germany. NDVI and landscape heterogeneity are in cases related with the presence of some of the studied species. However, the models applied to examine the relationship of NDVI with species presence need improvements concerning the scale that they are performed. Smaller geographical areas could improve the performance of the models. Working in larger scale maps that present smaller areas in detail, could make possible to use field data of weed species populations (in situ). Improvements of the study could be made by using crop, and cultivation techniques information, in relation with weed species presence/absence. Moreover, information of natural weed populations that are positioned in the field and recorded as herbicide resistant could upgrade the picture that we have of the species performance in differing environmental conditions. For the prediction of species future distribution, carbon dioxide emission trends per country could also add to this study. Hence, further study is needed to come up with prediction models of weed species future distribution. Finally, the models and maps produced in this study could be used as a monitoring tool by farmers that would like to find new cultivation areas and/or alternatively detect how their current agricultural areas will respond and influenced by weeds in the next few decades. Moreover, crop science and crop protection companies could be supplied by useful maps and information that combined with crop maps could become a powerful management tool for finding new innovative crop protection solutions.

7. References

Anderson, L.E., 1971. Chloroplast and cytoplasmic enzymes. II. Pea leaf triose phosphate isomerases. *Biochimica et Biophysica Acta*, 235, 237–244.

Bakkenes, M., Alkemade ,J.R.M., Ihle, F., Leemans, R, and Latour, J.B., 2002. Assessing effects of forecasting climate change on the diversity and distribution of European higher plants for 2050. *Global Change Biology*, 8, 390-407.

Black, C.C., 1973. Photosynthetic carbon fixation in relation to net CO2 uptake. *Annual Review of Plant Physiology*, 24, 253–286.

Bonham-Carter, G.F., (ed.), 1994. *Geographic Information Systems for Geoscientists – Modelling with GIS*. Computer Methods in the Geosciences, 13. Pergamon Press, New York, 398.

Boyce, M.S., Vernier, P.R., Nielsen, S.E. and Schmiegelow, F.K.A., 2002. Evaluating resource selection functions. *Ecological Modelling*, 157, 281–300.

Chander, G., Markham, B. L. and Helder, D. L., 2009. Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. *Remote Sensing of Environment*, 113, 893-903.

Chavez, P.S.Jr., 1988. An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. *Remote Sensing of Environment*, 24 (3), 459–479.

Davis, M.B., Shaw, R.G. and Etterson, J.R., 2005. Evolutionary responses to changing climate. *Ecology*, 86, 1704–1714.

de Souza Muñoz, M. E., De Giovanni, R., de Siqueira. M. F., Sutton, T., Brewer, P., Pereira, R. Si., Canhos, DAL. and Canhos, V.P., 2011. OpenModeller: a generic approach to species' potential distribution modelling. *Geoinformatica*, 15 (1), 111-135.

Edwards, G.E., Franceschi, V.R. and Voznesenskaya, E.V., 2004. Single-cell C(4) photosynthesis versus the dual-cell (Kranz) paradigm. *Annual Review of Plant Biology*, 55, 173–196.

Edwards, E.J., Osborne, C.P., Strömberg, C.A., Smith, S.A., Bond, W.J., Christin, P.A., Cousins, A.B., Duvall, M.R., Fox, D.L., Freckleton, R.P., , Ghannoum, O., Hartwell, J., Huang, Y., Janis, C.M., Keeley, J.E., Kellogg, E.A., Knapp, A.K., Leakey, A.D., Nelson, D.M., Saarela, J.M., Sage, R.F., Sala, O.E., Salamin, N., Still, C.J. and Tipple, B., 2010. The origins of C4 grasslands: integrating evolutionary and ecosystem science. *Science,* 328, 587–591.

Ehleringer, J.R., Sage, R.F., Flanagan, L.B. and Pearcy, R.W., 1991. Climate change and the evolution of C4 photosynthesis. *Trends in Ecology & Evolution*, 6, 95–

99.

Garrigues, S., Allard, D., Baret, F. and Weiss, M., 2006. Quantifying spatial heterogeneity at the landscape scale using variogram models. *Remote Sensing of Environment*,103, 81–96.

Gillespie, T.W., 2006. Predicting woody-plant species richness in tropical dry forests: a case study from South Florida, USA. *Ecological Applications*, 15, 27–37.

Gowik, U. and Peter Westhoff, P., 2011.The Path from C3 to C4 Photosynthesis. *Plant Physiology,* 155, 56-63.

Grinnell, J., 1917. Field tests of theories concerning distributional control. *The American Naturalist*, 51 (602), 115–128.

Hall, K., Reitalu, T., Sykes, M. T. and Prentice, H. C. , 2012. Spectral heterogeneity of QuickBird satellite data is related to fine-scale plant species spatial turnover in semi-natural grasslands. *Applied Vegetation Science*, 15, 145–157.

Hirzel, A.H. and Arlettaz, R., 2003. Modelling habitat suitability for complex species distributions by the environmental-distance geometric mean. *Environmental Management*, 32, 614-623.

Hirzel, H. A. and Gwenaëlle, L. L., 2008. Review, Habitat suitability modelling and niche theory. *Journal of Applied Ecology*, 45, 1372–1381.

Hirzel, A. H., Hausser, J., Chessel, D. and Perrin, N., 2002. Ecological - Niche factor analysis: How to compute Habitat - Suitability maps without absence data? *Ecology*, 83 (7), 2027–2036.

Hirzel, A.H., Helfer, V. and Métral, F. , 2001. Assessing habitat-suitability models with a virtual species. *Ecological Modelling*, 145 (2), 111–121.

Hirzel, A. H., Le Lay, G., Helfer, V., Randin, C. and Guisan, A., 2006. Evaluating the ability of habitat suitability models to predict species presences. *Ecological Modelling* 199, 142–152.

Huang, C., Yang, L., Homer, C., Wylie, B., Vogelman, J. and DeFelice, T., 2001. At-Satellite Reflectance: A First Order Normalization of Landsat 7 ETM+ Images. U.S. Geological Survey, Unpublished draft white paper. http://landcover.usgs.gov/pdf/huang2.pdf

Hutchinson, G.E., 1957. Concluding remarks. *Cold Spring Harbour Symposium on Quantitative Biology*, 22, 415–427.

Hyvönen, T., Glemnitz, M., Radics, L. and Hoffmann, J., 2011. Impact of climate and land use type on the distribution of Finnish casual arable weeds in Europe. *Weed Research*, 51, 201-208.

Hyvönen, T., Luoto, M. and Uotila, P. 2012. Assessment of weed establishment risk in a changing European climate. *Agricultural and Food Science,* 21, 348-360.

Janssen, L.L.F. and Huurneman, G.C. (eds.), 2001. *Principles of Remote Sensing*. ITC Educational Textbook Series 2, Second Edition, ITC, Enschede, The Netherlands.

IPPC (2001) Climate Change, 2001: Working Group II: Impacts, Adaptation and Vulnerability. www.grida.no/climate/ipcc_tar/wg2/.

Kumar, R. and Silva., L., 1973. Light ray tracing through a leaf cross- section. *Applied Optics* 12, 2950-2954.

Lee, J.-S., 2011. Combined effect of elevated CO2 and temperature on the growth and phenology of two annual C3 and C4 weedy species. *Agriculture, Ecosystems and Environment, 140*, 484–491.

Levin, N., Shmida, A., Levanoni, O., Tamari, H. and Kark, S., 2007. Predicting mountain plant richness and rarity from space using satellite-derived vegetation indices. *Diversity and Distributions*, 13, 692–703.

Long, S.P., 1999. Environmental responses. In: Sage, R.F. and Monson, R.K., eds, *C4 Plant Biology.* Academic Press, San Diego, pp 215–249.

MacArthur, R.H., 1957. On the relative abundance of bird species. *Proceedings of the National Academy of Sciences of the United States of America*, 43, 293–295.

Magurran, A.E., (ed.), 1988. *Ecological Diversity and Its Assessment*. Princeton University Press, Princeton, NJ.

Mannetje, L.'t., 2006. Climate change and grasslands through the ages - an overview. *Grassland Science in Europe*, 11, 733-738.

McDonald, A., Riha, S., DiTommaso, A., DaGaetano, A., 2009. Climate change and the geography of weed damage: Analysis of US maize systems suggests the potential for significant range transformations. *Agriculture, Ecosystems and Environment*, 130, 131-140.

Moss, S. R., 2002. Herbicide-Resistant Weeds. In: Naylor, R. E. L., (ed.), Chapter 11, *Weed Management Handbook*, 9th edition, British Crop Protection Council, Blackwell Publishing.

Oaks, A., 1994. Efficiency of nitrogen utilization in C3 and C4 cereals. *Plant Physiology*, 106, 407–414.

Oindo, B. O. and Skidmore, A. K., 2002. Interannual variability of NDVI and species richness in Kenya. *International Journal of Remote Sensing*, 23 (2), 285-298.

Ozturk, M., Rehder, H. and Zeigler, H., 1981.Biomass productionof C3 and C4 plants species in pure and mixed culture with different water supply. *Oecologia*, 50, 73-81.

Parker-Allie, F. , Musil, C. F. and Thuiller, W., 2009. Effects of climate warming on the distributions of invasive Eurasian annual grasses: a South African perspective. *Climatic Change*, 94, 87–103.

Parviainen, M., Luoto, M. and Heikkinen, R. K., 2009. The role of local and landscape level measures of greenness in modelling boreal plant species richness. *Ecological Modelling*, 220, 2690–2701.

Patterson, D. T., 1995a. Effects of Environmental Stress on Weed/Crop Interactio *Weed Science*, 43 (3), 483-490.

Patterson, D. T., 1995b. Weeds in a Changing Climate. *Weed Science*, 43 (4), 685-701.

Pearson, R.G. and Dawson, T.P., 2003. Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Global Ecology and Biogeography*, 12(5), 361–371.

Pielou, E.C., 1969. *An Introduction to Mathematical Ecology*. John Wiley, New York.

Shannon, C.E. and Weaver, W., 1962. *The Mathematical Theory of Communication*. University of Illinois Press, Urbana, IL.

Pulliam, H.R., 2000. On the relationship between niche and distribution. *Ecology Letters,* 3, 349–361.

Pyankov, V. I., Ziegler, H., Akhani, H., Deigele C. and Lüttge, U., 2010. European plants with C4 photosynthesis: geographical and taxonomic distribution and relations to climate parameters. *Botanical Journal of the Linnean Society*, 163, 283–304.

Raes, N. and ter Steege, H., 2007. A null-model for significance testing of presence-only species distribution models. *Ecography*, 30, 727-736.

Raines, G.L., 1999. Evaluation of Weights of Evidence to Predict Epithermal-Gold Deposits in the Great Basin of the Western United States. *Natural Resources Research*, 8 (4), 257-276.

Raines, G.L. and Bonham-Carter, G.F., 2006, Exploratory spatial modelling: demonstration for Carlin-type deposits, central Nevada, USA, using Arc-SDM. In: Harris, J.R. , (ed.), *GIS for the Earth Sciences,* Geological Association of Canada, Special Publication 44, 23-52.

Rocchini, D. and Neteler, M., 2012. Spectral rank–abundance for measuring landscape diversity. *International Journal of Remote Sensing*, 33 (14), 4458-4470.

Rocchini , D. and Vannini , A., 2010. What is up? Testing spectral heterogeneity versus NDVI relationship using quantile regression. *International Journal of Remote Sensing*, 31 (10), 2745-2756.

Root, T.L., Price, J.T., Hall, K.R., Schneider, S.H., Rosenzweig, C. and Pounds, J.A., 2003. Fingerprints of global warming on wild animals and plants. *Nature*, 421, 57-60.

Rouse, J. W., Haas, R. H., Schell, J. A., and Deering, D. W., 1973. Monitoring vegetation systems in the Great Plains with ERTS. *Proceedings Third ERTS Symposium*, 1, 48– 62.

Seto, K. C., Fleishman, E. , Fay, J. P. and Betrus, C. J., 2004. Linking spatial patterns of bird and butterfly species richness with Landsat TM derived NDVI. *International Journal of Remote Sensing*, 25:20, 4309-4324.

Shreeve, T.G., Dennis, Ri.H. and Pullin., A.S., 1996. Marginality: scale determined processes and the conservation of the British butterfly fauna. *Biodiversity and Conservation,* 5, 1131-1141.

Soberon, J., 2007. Grinnellian and Eltonian niches and geographic distributions of species. *Ecology Letters*, 10 (12), 1115–1123.

Soberon, J. and Peterson, A.T., 2005. Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodiversity Informatics*, 2, 1–10.

Stockwell, D. and Peters, D, 1999. The GARP modeling system: problems and solution to automated spatial prediction. *International Journal of Geographical Information Science*, 13 (2), 143-158.

Thuiller, W., Lavorel, S., Araùjo, M.B., Sykes, M.T. and Prentice, I.C., 2005. Climate change threats to plant diversity in Europe. *Proceedings of the National Academy of Sciences*, USA 102, 8245-8250.

Walther, G.R., Post, E. and Convey, P., 2002. Ecological responses to recent climate change. *Nature*, 416, 389-395.

Whittaker, R.H., 1965. Dominance and diversity in land plant communities. *Science,* 147, 250–260.

Woodward, F. I., 1988. Temperature and the distribution of weed species. In: S. P. , Long, and F. I., Woodward ,(eds.), *Plants and Temperature. Company of Biologists*, Cambridge. 59-75.

8. Internet sources

- International Survey of Herbicide Resistant Weeds, http://www.weedscience.org/summary/UniqueCountry.asp?lstCountryID=19, 12-Jun-13
- Landsat org., http://landsat.org/ortho/index.php, 03-Jul-13
- GADM, http://www.gadm.org/home, 20-Jul-13
- GBIF, http://data.gbif.org/species/, 20-Jul-13
- Colorado State University, http://ibis.colostate.edu/WebContent/WS/ColoradoView/TutorialsDownloads/CO_RS_Tutorial10.pdf, 20-Jul-13
- WorldClim, http://www.worldclim.org/bioclim, 20-Jul-13
- EEA europa, http://www.eea.europa.eu/data-and-maps/data/dominant-land-cover-types-2000-1, 22-Jul-13
- European Soil Portal, http://eusoils.jrc.ec.europa.eu/ESDB_Archive/ESDBv3/legend/LegendData.cfm (soil data), 22-Jul-13
- Biomapper, http://biomapper.wikispaces.com/Ecogeographical+variables, 03-Aug-13
- GCM, http://www.ccafs-climate.org/form.php, 17-Aug-13
- Wolframalpha, http://www.wolframalpha.com/widgets/view.jsp?id=ec78b0700064223c5cd7898812ecffae,19-Sept-13
- USGS, http://landsat.usgs.gov/band_designations_landsat_satellites.php, 10-Oct-13a
- USGS, https://lta.cr.usgs.gov/LETMP, 10-Oct-13b
- CRISP, http://www.crisp.nus.edu.sg/~research/tutorial/intro.htm, 13-Oct-13
- Sanderson, R., 2012, ftp://ftp.wsl.ch/downloads/babst/Fernerkundung_WS2012/literatur/remote_sensing.pdf, 20-Oct-13
- SEINet, http://swbiodiversity.org/imglib/seinet/Chenopodiaceae/photos/Chenopodium-album-FL-web-.jpg, 03-Nov-2013a

- SEINet, http://swbiodiversity.org/imglib/seinet/genfield/palexander/set001/Amaranthuret_10Sep06_6557.jpg,3-Nov-2013b
- SEINet,http://swbiodiversity.org/imglib/seinet/Papaveraceae/500px-Poster_Papaver_2a1.jpg, 3-Nov-2013c
- SEINet, http://swbiodiversity.org/seinet/imagelib/imgdetails.php?imgid=19331, 3-Nov-2013d
- SEINet, http://swbiodiversity.org/imglib/seinet/genfield/palexander/set003/Echinochloa_c-g_19Aug06_1595.jpg, 3-Nov-2013e
- SEINet, http://swbiodiversity.org/seinet/imagelib/imgdetails.php?imgid=11147, 3-Nov-2013f
- SEINet, http://swbiodiversity.org/imglib/seinet/genfield/palexander/set004/Sorghum_hal_4Sep06_5007.jpg, 3-Nov-2013g
- SEINet, http://swbiodiversity.org/imglib/seinet/genfield/palexander/set004/Sorghum_hal_4Sep06_5046.jpg, 3-Nov-2013h
- Wikimedia, http://upload.wikimedia.org/wikipedia/commons/c/c0/Lolium_rigidum_1.JPG, 3-Nov-2013
- Worldatlas, http://www.worldatlas.com/geography/greecegeography.htm, 5-Feb-14a
- Worldatlas, http://www.worldatlas.com/geography/germanygeography.htm, 6-Feb-14b
- DATA BASIN, http://databasin.org/datasets/5f7f3a247e2d468cb4e43517baa931b7, 08-Feb-2014
- IPCC, http://www.ipcc.ch/ipccreports/tar/wg1/029.htm, 08-Feb-2014
- CCAFS, http://www.ccafs-climate.org/statistical_downscaling_delta/, 08-Feb-2014
- Emorsgateseeds, http://wildseed.co.uk/species/view/8,15-February-2014

Appendices

Appendix 1 Weed species photos

Photo 1. Chenopodium album (SEINet 2013a)

Photo 2. Amaranthus retroflexus (SEINet 2013b)

Photo 3.Papaver rhoeas (SEINet 2013c)

Photo 4. Conyza canadensis (SEINet 2013d)

Photo 5. Echinochloa crus-galli plant (SEINet 2013e)

Photo 6. Echinochloa crus-galli dried mature flower (SEINet 2013f)

Photo 7. Lolium rigidum (Wikimedia 2013)

Photo 8. Sorghum halepense (SEINet 2013g)

Photo 9. Sorghum halepense flower (SEINet 2013h)

Appendix 2 List of other LUMA-GIS theses

Series from Lund University
Department of Physical Geography and Ecosystem Science

Master Thesis in Geographical Information Science (LUMA-GIS)

1. *Anthony Lawther:* The application of GIS-based binary logistic regression for slope failure susceptibility mapping in the Western Grampian Mountains, Scotland. (2008).
2. *Rickard Hansen:* Daily mobility in Grenoble Metropolitan Region, France. Applied GIS methods in time geographical research. (2008).
3. *Emil Bayramov:* Environmental monitoring of bio-restoration activities using GIS and Remote Sensing. (2009).
4. *Rafael Villarreal Pacheco:* Applications of Geographic Information Systems as an analytical and visualization tool for mass real estate valuation: a case study of Fontibon District, Bogota, Columbia. (2009).
5. *Siri Oestreich Waage:* a case study of route solving for oversized transport: The use of GIS functionalities in transport of transformers, as part of maintaining a reliable power infrastructure (2010).
6. *Edgar Pimiento:* Shallow landslide susceptibility – Modelling and validation (2010).
7. *Martina Schäfer:* Near real-time mapping of floodwater mosquito breeding sites using aerial photographs (2010)
8. *August Pieter van Waarden-Nagel:* Land use evaluation to assess the outcome of the programme of rehabilitation measures for the river Rhine in the Netherlands (2010)
9. *Samira Muhammad:* Development and implementation of air quality data mart for Ontario, Canada: A case study of air quality in Ontario using OLAP tool. (2010)
10. *Fredros Oketch Okumu*: Using remotely sensed data to explore spatial and temporal relationships between photosynthetic productivity of vegetation and malaria transmission intensities in selected parts of Africa (2011)
11. *Svajunas Plunge:* Advanced decision support methods for solving diffuse water pollution problems (2011)
12. *Jonathan Higgins:* Monitoring urban growth in greater Lagos: A case study using GIS to monitor the urban growth of Lagos 1990 - 2008 and produce future growth prospects for the city (2011).
13. *Mårten Karlberg:* Mobile Map Client API: Design and Implementation for Android (2011).
14. *Jeanette McBride:* Mapping Chicago area urban tree canopy using color infrared imagery (2011)
15. *Andrew Farina:* Exploring the relationship between land surface temperature and vegetation abundance for urban heat island mitigation in Seville, Spain (2011)
16. *David Kanyari*: Nairobi City Journey Planner An online and a Mobile Application (2011)

17. *Laura V. Drews:* Multi-criteria GIS analysis for siting of small wind power plants - A case study from Berlin (2012)
18. *Qaisar Nadeem:* Best living neighborhood in the city - A GIS based multi criteria evaluation of ArRiyadh City (2012)
19. *Ahmed Mohamed El Saeid Mustafa:* Development of a photo voltaic building rooftop integration analysis tool for GIS for Dokki District, Cairo, Egypt (2012)
20. *Daniel Patrick Taylor*: Eastern Oyster Aquaculture: Estuarine Remediation via Site Suitability and Spatially Explicit Carrying Capacity Modeling in Virginia's Chesapeake Bay (2013)
21. *Angeleta Oveta Wilson:* A Participatory GIS approach to *unearthing* Manchester's Cultural Heritage '*gold mine'* (2013)
22. *Ola Svensson:* Visibility and Tholos Tombs in the Messenian Landscape: A Comparative Case Study of the Pylian Hinterlands and the Soulima Valley (2013)
23. *Monika Ogden:* Land use impact on water quality in two river systems in South Africa (2013)
24. *Stefan Rova:* A GIS based approach assessing phosphorus load impact on Lake Flaten in Salem, Sweden (2013)
25. *Yann Buhot:* Analysis of the history of landscape changes over a period of 200 years. How can we predict past landscape pattern scenario and the impact on habitat diversity? (2013)
26. *Christina Fotiou:* Evaluating habitat suitability and spectral heterogeneity models to predict weed species presence (2014)