Spatial modeling of plant distributions

Coupling Remote Sensing with GIS-based Models

Dissertation

To Fulfill the Requirements for the Degree of

„Doctor rerum naturalium“(Dr. rer.nat.)

Submitted to the Council of the

Faculty of Biology and Pharmacy

Of the Friedrich Schiller University Jena

Master. Science. Hamidreza Keshtkar

born on 08/04/1981 in Iran

Jena, April 2016

Gutachter:

1: Prof. Dr. Stefan Halle, Institute of Ecology, University of Jena, Jena, Germany.

2: PD Dr. Sören Hese, Institute of Remote Sensing, University of Jena, Jena, Germany.

3: Prof. Dr. Frank Schurr, Institute of Landscape and Plant Ecology, University of Hohenheim, Stuttgart, Germany.

Tag der öffentlichen Verteidigung: April 28, 2016

Contents

Summary

Spatial species distributions and the relationship between species and environmental factors have been studied for several years. Climate change and habitat fragmentation can be considered as the factors effective in biodiversity changes. Landscape fragmentation can have a significant effect in reducing migration capabilities of plants by lessening suitable habitats and separating them from each other, as well as, via increasing, decreasing, or eliminating dispersal barriers and climate change affects plant species distributions and may create new environments and forces organisms have to react to. Dynamic climate variables changing with a projection time frame and static landscape variables have been adopted by most studies to project future species distribution. This method of application ignores the dynamics landscape changes over time. Also, many studies on climate change and ecosystems consider unlimited dispersal or no dispersal. Neglecting dispersal process is usually seen as a defect and limitation in projects of species distribution models under climate change. Therefore prediction of species range shifts under climate change and other physical processes is a crucial challenge for the management of natural resources.

The major objective of this thesis was to integrate MigClim, SDM and CA-Markov chain models so as to assess the effects of future landscape fragmentation and climate change scenarios on the geographic distributions of three open-land plant species in 21st century.

To simulate future landscape changes, we used a hybrid (CA-Markov) model (chapter 3). This model was derived by examining past trends in land change and projected changes in basic environmental and human driven factors. The end result clearly showed if the current trends of change continue regardless of the actions of sustainable development, drastic natural area decline will ensue.

A remote sensing study was conducted to identify the best classifier algorithm (three pixel-based and one object-based method) (chapter 2). This study demonstrates that object-based support vector machine classifier is the most accurate classifier with an overall classification accuracy of 93.54% and a kappa value of 0.88 for the image 2010. The best method (i.e. object-based support vector machine) was then used in chapter 3 to obtain the land-cover maps corresponding to different years.

In chapter 4, we compare six species distribution models (three machine-learning and three regression models) constructed with different subsets of environmental predictors (climate-only, topography-only, and topo-climate variables). The generalized additive model (GAM) showed the best accuracy when projected with a set of topo-climate variables. The best model and the best set of predictors were used to produce habitat suitability maps of plant species in chapter 5.

For all target plants, simulations were performed for four dispersal events (full dispersal, no dispersal, regular dispersal (short-distance dispersal), and regular dispersal along with long-distance dispersal), two landscape (static and dynamic change) and two climate change (RCP4.5

and RCP8.5) scenarios (chapter 5). The choice of dispersal events is in line with the choice made in Engler et al. (2009). In this investigation, it was shown that the predicted distribution areas for all the three species under RCP8.5 scenario will largely increase in the coming decades. Also, a significant difference appears to be between the simulations of realistic dispersal limitations and those considering full or no dispersal for projected future distributions during the 21st century. Besides, the results obtained by the limited projections of future plant distributions via realistic dispersal restrictions showed to be generally closer to no-dispersal than to full-dispersal scenario when compared with real dispersal scenarios.

Overall, the results of this study indicate that dispersal limitations can have an important impact on the outcome of future projections of species distributions under climate change scenarios. Also our findings clearly showed that change in landscape fragmentation is more effective than the climate change impacts on species distributions in our study area.

Zusammenfassung

Die räumliche Verteilung von Arten sowie ihre Beziehung zu den herrschenden Umweltfaktoren sind seit vielen Jahren Gegenstand der Forschung. Klimaänderung und Lebensraumfragmentierung werden als Hauptbedrohungen der Biodiversität angesehen. Landschaftszerschneidung kann die Migrationsfähigkeit von Pflanzen einschränken, indem sie den geeigneten Lebensraum für eine erfolgreiche Kolonisierung verringert. Sie kann ebenso durch Vergrößern, Verringern oder Beseitigen von Ausbreitungsschranken und Klimaänderungen die Verbreitung von Pflanzenarten beeinträchtigen und vermag neue Umgebungen zu schaffen und Organismen zum Reagieren zu zwingen. Die meisten Studien haben den Ansatz dynamischer Klimavariablen in Abhängigkeit eines Vorhersagezeitraums in Zusammenhang mit statischen Landschaftsvariablen angenommen, um die zukünftige Artenverbreitung vorherzusagen. Diese Anwendungsmethode vernachlässigt dynamische Landschaftsänderungen mit der Zeit. Außerdem gehen viele Studien über Klimaänderung und Ökosysteme entweder von unbegrenzter oder keiner Ausbreitung aus. Die Vernachlässigung des Ausbreitungsprozesses wird gewöhnlich in Projekten über Artenverteilungsmodellen unter Klimaeinfluss als Fehler und Einschränkung betrachtet. Deshalb ist die Vorhersage einer Veränderung des Verbreitungsgebiets von Arten bei Klimaänderungen oder anderen physikalischen Prozessen eine entscheidende Herausforderung für die Verwaltung natürlicher Ressourcen.

Das Hauptziel dieser These war die Integration von MigClim, SDM und CA-Markov-Ketten-Modellen, um die Auswirkungen zukünftiger Landschaftszerschneidungs- und Klimaänderungs-Szenarien auf die geografische Verbreitung von drei Offenlandpflanzen im 21. Jahrhundert abzuschätzen.

Um zukünftige Landschaftsänderungen zu simulieren, benutzten wir ein hybrides (CA-Markov) Modell (Kapitel 3). Dieses Modell wurde aus der Untersuchung vergangener Trends in Landschaftsänderungen und vorhergesagten Änderungen in grundlegenden Umwelt- und anthropogenen Faktoren abgeleitet. Das Endergebnis zeigt ganz klar, dass eine drastische Verringerung naturnaher Gebiete folgt, wenn der aktuelle Trend der Änderungen anhält und die Aktionen einer nachhaltigen Entwicklung nicht berücksichtigt werden.

Eine Fernerkundungsstudie wurde durchgeführt, um die besten Klassifikatoralgorithmen zu identifizieren (drei Pixel-basierte und eine Objekt-basierte Methode) (Kapitel 2). Diese Studie zeigt, dass ein Objekt-basierter Support-Vektormaschinen-Klassifikator der genaueste Klassifikator ist, mit einer Gesamtklassifizierungs-Genauigkeit von 93.54% und einem Kappawert von 0,88 für das Bild von 2010. Die beste Methode (d.h. Objekt-basierte Support-Vektormaschinen-Klassifikator) wurde anschließend in Kapitel 3 verwendet, um die Bodenbedeckungskarte verschiedener Jahre zu erhalten.

In Kapitel 4 vergleichen wir sechs Artenverteilungsmodelle (drei Maschinenlern- und drei Regressions-Modelle), die mit verschiedenen Untermengen von Umwelt-Einflussfaktoren (Variablen nur für Klima, nur für Topographie, für Klima und Topographie) erstellt wurden. Das Allgemeine Additive Modell (GAM – "generalized additive model") zeigt die höchste Genauigkeit, wenn die Abschätzung nur mit Variablen für Klima und Topographie erfolgt. Das beste Modell und die beste Auswahl an Prädiktorvariablen wurden in Kapitel 5 verwendet, um Landkarten für geeignetem Lebensraum der Pflanzenarten zu erstellen.

Für jede Pflanzenart wurden Simulationen von vier Ausbreitungsszenarien (Vollausbreitung, keine Ausbreitung, realistische Ausbreitung, realistische Ausbreitung mit Fernausbreitungsereignissen), zwei Landschafts- (statisch und dynamische Änderungen) und zwei Klimaänderungs-Szenarien (RCP4.5 und RCP8.5) ausgeführt (Kapitel 5). Die Festlegung der Dispersal-Ereignisse erfolgte in Übereinstimmung mit Engler et al. (2009). In dieser Untersuchung wurde gezeigt, dass das vorhergesagte Verbreitungsgebiet aller drei Arten unter dem RCP8.5 Szenario in den drei kommenden Jahrzehnten stark zunehmen wird. Außerdem zeigt sich ein wesentlicher Unterschied für die vorhergesagte Verbreitung im 21. Jahrhundert zwischen Simulationen realistischer Ausbreitungsgrenzen und solchen mit voller oder abwesender Ausbreitung. Im Übrigen waren die erhaltenen Ergebnisse der begrenzten Vorhersage zukünftiger Pflanzenverteilungen mit realistischen Ausbreitungsgrenzen i.A. dichter am Szenario ohne Ausbreitung als an der Voll-Ausbreitung dran, verglichen mit realistischen Ausbreitungs-Szenarien.

Zusammengefasst zeigen die Ergebnisse dieser Studie, dass Ausbreitungsbegrenzungen einen wesentlichen Einfluss auf das Resultat von Zukunftsabschätzungen der Artenverteilung unter Klimaänderungs-Szenarien haben können. Außerdem zeigen unsere Befunde ganz klar, dass Landschaftszerschneidung in unserem Studienbereich einen größeren Einfluss auf die Artenverbreitung hat als Auswirkungen von Klimaänderungen.

Manuscript overview and author contribution to the chapters of this thesis

Chapter 2

Analysis of Landscape Changes Using Multi-Temporal Remote Sensing Imagery and Machine-Learning Classifier. Hamidreza Keshtkar & Winfried Voigt

This study aims to use and compare two different classification methods to monitor land-use change during last two decades.This study demonstrates that object-based support vector machine classifier is the most accurate classifier.

The authors have made the following declarations about their contributions: HK and WV conceived the project. WV provided the data and commented on the paper. HK analyzed the data and wrote the first draft. Overall contribution of HK was 90% of the work.

Manuscript status: ready for submission

Chapter 3

A spatiotemporal analysis of landscape change using an integrated Markov chain and cellular automata models. Hamidreza Keshtkar & Winfried Voigt

This manuscript aims to analysis and modeling of land-use change in our study area in the central Germany. An integrated Cellular Automata-Markov Chain land change model was carried out to simulate the future landscape change during the period of 2020–2050. As a consequence, land change model predicts very well a continuing downward trend in grassland, farmland and forest areas, as well as a growing tendency in built-up areas. Hence, if the current trends of change continue regardless of the actions of sustainable development, drastic natural area decline will ensue.

The authors have made the following declarations about their contributions: HK and WV conceived the project. WV provided the data and commented on the paper. HK analyzed the data and wrote the first draft. Overall contribution of HK was 95% of the work.

Manuscript status: Published as an article in Model. Earth Syst. Environ Journal, 2:1-13, 2016. DOI: 10.1007/s40808-015-0068-4

Chapter 4

Comparison of statistical models to predict the spatial distribution of plant species in Central Germany. Hamidreza Keshtkar & Winfried Voigt

The objective of this study is to compare the performance of some of the most common methods of presence-absence distribution models using data on the distribution of three plant species in central Germany. In this study, three different sets of explanatory variables (climate-only, topography-only and topography-climate combination) for each species were quantified and the interaction of differences among predictor variables with differences in the modeling approaches

in determining the accuracy of predictions was tested. Regression-based approaches showed better performances than machine-learning methods. The results showed that topography-climate variables are the most important variables for mapping potential suitable habitats of target species.

The authors have made the following declarations about their contributions: HK and WV conceived the project. WV provided the data and commented on the paper. HK analyzed the data and wrote the first draft. Overall contribution of HK was 90% of the work.

Manuscript status: ready for submission

Chapter5

Potential Impacts of Climate and Landscape Fragmentation Changes on Plant Distributions: Coupling Multi-Temporal Satellite Imagery with GIS-based Cellular Automata Model. Hamidreza Keshtkar & Winfried Voigt

The aim of this study is to test probable alteration of future species distribution based on the association of land-use and climate change scenarios compared to the classical approach that assumed an unchanged land-use. A significant difference appears to be between the simulations of realistic dispersal limitations and those considering full or no dispersal for projected future distributions. Although simulations accounting for dispersal limitations produced, for our study area, results that were closer to no dispersal than to full dispersal. Additionally, our results revealed that future land-use change is more effective than the climate change impacts on species distributions in this study.

The authors have made the following declarations about their contributions: HK and WV conceived the project. WV provided the data and commented on the paper. HK analyzed the data and wrote the first draft. Overall contribution of HK was 90% of the work.

Manuscript status: Published as an article in Ecological Informatics Journal, 32, 145-155, 2016. DOI: 10.1016/j.ecoinf.2016.02.002

Chapter 1

General Introduction

Ecologists have been studying for a long time to find out why some species are found only in certain habitats and how they found their way to these regions (Darwin, 1859; Connell, 1961). Ecologists often consider the geographical ranges of species, which are explained by the intricate interactions of each species with the environment it establish in as a basis for answering such questions. Species distribution models (SDMs) generally specify spatially and temporally suitable habitats for species by combining species occurrence data with environmental predictors (Brown et al., 1996; Guisan and Zimmermann, 2000).

1.1. Ecological niche

The conceptual framework of modeling approaches for evaluating and predicting relationships between living organisms and their environment in a given place and at a given time is referred to as the ecological niche concept. This concept was first formally defined by Hutchinson (1957). According to Hutchinson, the ecological niche of a species is the set of environmental conditions under which the species can sustain a positive growth rate. Hutchinson later modified this concept by introducing a distinction between “fundamental” and “realized” niche. The first term can only be considered for lifeless factors such as climate, topography or soil properties, but the second term can also be considered for biotic interactions such as competition and facilitation (Engler, 2009). Plant species that grow in mountainous areas can be a good example to better describe the concepts of “fundamental” and “realized” niche. Climatic constraints limit the habitat of these plants to below a certain altitude, and competition severely limits the possibility of their presence in low-altitude mountain ranges (Brown et al., 1996). So the low-altitude mountain ranges are a part of the fundamental niche of this species (i.e. in the absence of competition, they can grow in these habitats), but not a part of their realized niche because competition does not allow them to be present in these habitats.

Nowadays, most SDM models are based on Hutchinson’s niche concept. However, some researchers (e.g., Pulliam, 2000) believe that the presence or absence of a species in a habitat is a more important factor than the suitability of the habitat or interspecies competition. Limited distribution, local extinctions, and recolonization are events that constantly occur in nature, and thus, one can regularly see species not present in their suitable habitat but present in an unsuitable habitat.

Engler (2009) suggests that "meta-population and source-sink dynamics must also be considered in the concept of ecological niches because undoubtedly factors like the factor of competition, affect

the formation of ecological niches". For instance, Pulliam (2000) has shown that how dispersal can maintain population in unsuitable areas. Studies have shown that when dispersal is high, a significant portion of the population (up to 30%) can live outside of fundamental niche. However until now most studies have ignored metapopulation and source-sink dynamics.

SDMs assume that by investigating the environmental conditions in which a species exists, it is possible to satisfy its ecological needs (Pearson and Dawson, 2003). This implies that, the assumption of equilibrium between a species and its habitat is a basic principle for modeling full potential distribution of species in SDMs (Guisan and Zimmermann, 2000; Engler, 2009). That is why sometimes SDMs are called “equilibrium models”. The condition of an invasive species that is newly arrived to an environment is a good example of non-equilibrium condition. Such a species is in migration stage, and therefore does not occupy all suitable habitats (Engler, 2009).

1.2. Methodological improvements of SDMs

The geographical ranges of mobile species (such as most animals) may be dynamic and change in a relatively short period of time (e.g., in summer and winter).Changes in the geographical ranges of non-mobile species (such as plants), however, often occur at longer intervals, because these species often live in a single location throughout their lives, and thus, their dispersal and distribution are limited to their reproduction. Of course, factors such as wind, animals, and humans may also play a role in this regard (Nathan and Muller-Landau, 2000; Meier, 2011).

SDMs for predicting distribution of living organisms under the influence of climate change in future generally face two major limitations. One is that they do not consider probable changes in intragroup and extragroup biological interactions and evolutionary interactions, and the other is that they cannot study and utilize dispersal-related processes (Pearson and Dawson, 2003, Pearman et al., 2008).

The biggest limitation of SDMs is that they do not consider the dispersal of a species or the dynamicity of a population. Although this is not necessarily an issue in modeling for current climate (assuming that the desired species has reached an environmental equilibrium), it becomes an important issue when modeling for future climate. The reason is obvious: projections obtained from SDMs emphasize locations that are expected to be potentially suitable for the species in the future, but these models do not consider whether these species would actually be able to reach these new suitable locations. The issue is exacerbated by a significant increase in fragmentation which is an obstacle for species dispersal (Pitelka et al., 1997; Engler, 2009).

When not examining distribution-limiting factors, studies examining the impact of climate change on plant species distribution can only examine two scenarios: "no dispersal" and "full dispersal". The former assumes that species cannot move to new environments. For this reason, climate change can affect their survival. Unlike the no-dispersal scenario, the full-dispersal assumes that

species can be established in all areas where the environmental conditions are suitable for their growth based on the model's prediction (Bateman et al., 2013). This scenario is widely used and in some references it is called "unlimited dispersal" and "universal dispersal" (Engler et al., 2009; Thomas et al., 2004). These two assumptions can lead to significantly different results. The difference between these two assumptions highlights the need for more dynamic models that consider the limits of dispersal.

1.3. Evaluating the potential effects of climate change

Studies have shown climate change in previous decades had considerable effects on living organisms, and sometimes led to migration of species or threatened them with the risk of extinction. Scientists believe climate change in the twenty first century will have deeper effects on both biotic and abiotic conditions of the environment. Therefore, researchers are interested in studying the effects of these changes on distribution of living organisms in the coming years (Lenoir et al., 2008, Parolo and Rossi, 2008).

Since the climate was relatively stable during the Holocene period, the current ranges are expected to be the result of adaptation of species with the current environment (Kullman and Kjallgren, 2000). Today, it is assumed that global change is rapidly changing the distribution of plant species (Thomas, 2010). Many species also have to adjust their ranges with a rate as high as the rate of climate change to be able to adapt to the new environmental conditions (Parmesan et al., 1999). In this situation, if a species cannot migrate, it either quickly adapts to the new ecological conditions of its current habitats or becomes extinct (Channell and Lomolino, 2000).

There are two groups of studies that evaluated the potential impacts of the predicted climate change on species using the modeling of species distribution methods. The first group that called global scale was performed over a big geographical area but at a low spatial resolution. The second group that termed local scale was performed with high spatial resolution but over a small geographical area. Global scale studies were usually performed at a continental scale or word-wide, with a spatial resolution of ~10-50 km, sometimes a little fewer. A good example for global scale SDMs is the project by Thuiller et al (2005) who standardized Europe-wide models at a resolution of 50 km and projected species habitats at a resolution of ~15 km. Local scale studies perform with high spatial resolution (usually in the scope of 20-100 m), but over a small spatial area (a few hundred or a few thousand square kilometers). The works form Dirnbök et al. (2003) or Randin et al. (2006) are two typical examples of local scale SDMs.

Although global scale studies have been conducted in various regions of the world, local scale studies are mainly restricted to the region under study with limited spatial ranges. Moreover, since different studies usually use diverse methods and climate change scenarios, their results are difficult to compare. Therefore, to compensate for this limitation, various studies should be

conducted on a small scale but in a wide geographical range using the factor of climate change (Engler, 2009).

1.4. Objectives and structure of the thesis

Species distributions are predicted based on various concepts and assumptions, most of which have limitations that must be considered. For example, it is expected that the range of a species is strongly influenced by factors such as landscape fragmentation, climate change scenarios, the species' dispersal rate and capacity, and, finally, by the interactions between these processes. Few of the extensively-used models, however, consider the effects of these factors. Therefore, it is important to study how climate factors affect species distribution when combined with other factors (e.g., dispersal limitations, landscape fragmentation, etc.). Such studies would better explain the effects of these factors and eventually enable the management of the most important one(s).

To improve the prediction of migration potential of plants under future climate change scenarios I used a dynamic model –MigClim (Engler et al., 2012)– to couple dispersal processes and landscape fragmentation scenarios with SDMs in projecting distributional changes under climate change (Chapter 5). This model has a cellular-automata structure which presents the study field as a regular network of stations and simulates the changes in species distribution. The main core of MigClim is based on potential distribution maps of plant species produced by SDMs. For this reason, Chapter 4 is an evaluation of some of the most common presence-absence distribution models using data on the distribution of three plant species in central Germany; three different sets of explanatory variables (climate-only, topography-only and topography-climate combination) are quantified. MigClim model lets that individual species respond to changes in their habitats, but the model ignores the dynamics landscape fragmentation changes over time. To solve this deficiency I applied a cellular automata-Markov chain model (CA-Markov) as an alternative to simulate landscape change scenarios during the 21st century (Chapter 3). Effective analysis of landscape changes require a considerable amount of data about the Earth's surface. Remote sensing prepares a great source of data from which updated land-cover maps, and changes can be analyzed and predicted efficiently. Several image classifier techniques have been developed; the selection of an appropriate classification method is very important for increasing the accuracy and precision of land-cover mapping. Thus, Chapter 2 focuses on comparing various machine-learning image classifier algorithms with land-cover mapping. The framework of this thesis is illustrated in Figure 1.

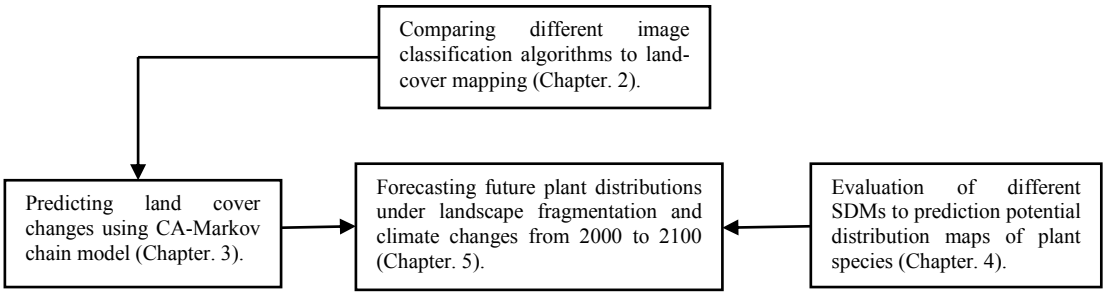


Figure 1. Flowchart of different chapters in this PhD thesis.

Based on the aforementioned issues, this dissertation examines the results of landscape fragmentation and climate changes on spatiotemporal prediction of three plant species. The goal of my work was to attempt to refine SDMs for better incorporate main processes that are neglected in last studies. Ultimately, my goal was to apply improved models to develop optimal strategies for species management. Specifically my study has four specific objectives:

1- To develop a general understanding of the recent historical distribution of land-cover structures in the study area.

This objective will employ remotely sensed imagery as well as change detection methods. I will evaluate various machine-learning algorithms and comparing pixel-based and object-based methods to land-cover mapping.

2- To simulate future land changes using CA-Markov techniques.

This objective will perform CA-Markov model to derive future lands-cover change. Land-cover change simulation models basically examine changes in land at two different periods (t0 and t1) and then simulate land changes for future periods (i.e. t2). I will simulate future land-cover changes in time and space based on their current state and on ancillary information which may drive future transitions among land-cover classes.

3- Comparing species distribution models constructed with different subsets of environmental predictors.

This objective seeks to compare some of the most common methods of presence-absence distribution models using data on the distribution of three plant species in central Germany. Specifically, I will compare three regression methods and three machine-learning methods. I will also quantify three different sets of explanatory variables for each species, and will test differences in the interaction of predictor variables with differences in modeling approaches in determining the accuracy of predictions.

4- To predict future plant distributions under landscape fragmentation and climate changes.

The major objective of this PhD thesis is to integrate MigClim, CA, and Markov chain models so as to assess the effects of future landscape fragmentation and climate change scenarios on the geographic distributions of three open-land plant species in 21st century. For each species, simulations will run for four dispersal scenarios (full dispersal, no dispersal, realistic dispersal, and realistic dispersal with long-distance dispersal events), two landscapes (static and dynamic change) and two climate change (RCP4.5 and RCP8.5) scenarios. To the best of our knowledge, this is the first study to combine the CA-Markov model with the models of species distribution to investigate species migration in the future.

Specifically, this study attempted to answer the following four research questions:

1- Can regression-based models as well as machine-learning models be used for prediction suitable habitat maps?

2- What is the potential impact of climate change on plant species?

3- How much does the inclusion of dispersal limitation events affect projections?

4- To what extent does landscape fragmentation prevent the movement of species toward suitable habitats?

Chapter 2

Analysis of Landscape Changes Using Multi-Temporal Remote Sensing Imagery and Machine-Learning Classifier

Hamidreza Keshtkar and Winfried Voigt

2.1. Summary

Frequent human activities resulted by fast urbanization lead to an assortment of environmental issues. Therefore, for efficient environmental management and urban planning, monitoring land-cover change is critical. We sought to pursue two objectives: first, to compare pixel-based random forest (RF) and decision tree (DT) classifier methods and support vector machines (SVM) algorithm both in pixel-based and object-based approaches to classification of land-cover in a heterogeneous landscape for 2010; and second, to examine the spatio-temporal land-cover change in two last decades (1990-2010) using Landsat data. This study demonstrates that object-based support vector machine classifier is the most accurate classifier with an overall classification accuracy of 93.54% and a kappa value of 0.88. The post-classification change detection algorithm was used to determine the trend of changes between land-cover classes. As a consequence, the most significant change occurred from 1990 to 2010 is caused by the expansion of built-up area. In addition to the net changes, the rate of annual change for each phenomenon was calculated in order to have a better understanding of the process of change. Between the years 1990 to 2010, on average, about 4.53% of lands annually turned to the built-up land, while there is an annual decrease of about 0.81% in natural lands. Hence, if the current trends of change continue regardless of the actions of sustainable development, drastic natural area decline will ensue. The results of this study can be a valuable baseline for land-cover managers in the region to better understanding the current situation as well as adopt appropriate strategies for management of land-cover.

Keywords: land-cover; decision tree; random forest; support vector machines; object-based; Landsat.

2.2. Introduction

Land-cover is a key variable in the Earth system as it is related to most of the human and physical environments. Change in the situation of land-cover is an important variable among the global changes that affect environmental systems (Otukei and Blaschke, 2010).

Land-cover changes play a main role in the global carbon cycle, and in the exchange of greenhouse gases between the atmosphere and the earth surface (Loveland and Belward, 1997). In general, intense human activities have an increase in construction and farmland and destruction of forests, meadows and other natural resources (Lambin and Geist, 2003; Lawrence et al., 2012). Agricultural lands increase the entry of nutrients into the water in the basin and in particular affect the amount of nitrogen (Robson, 2014). Human activities associated with construction pollute atmosphere, water and soil (Kang et al., 2010; Li et al., 2009). Destruction of forests and meadows in addition to the loss of biodiversity causes the release of carbon into atmosphere and changes land-surface albedo, and as a result affects climate change (Foley et al., 2005; Hua and Chen, 2013). Conversely, afforestation and reforestation remove carbon from the atmosphere. New evidences show changes in land-cover that have human origin during the past 150 years have led to the release of large amounts of carbon into the atmosphere. Although the combustion of fossil fuels is the main source of the release of carbon into the atmosphere, land-cover has still significant share (~20 percent) in the pollution of the atmosphere. Increase in the level of greenhouse gases and emission of the heat caused by urbanization have increased the Earth temperature significantly. In addition, temperature change affects the amount of humidity and precipitation, which in turn reduces recovery of forests and pastures (Davin and de Noblet-Ducoudré, 2010; Hasler et al., 2009; Snyder et al., 2004). In general, we can say that change in land-cover have positive and negative effects on human health and can have favorable or unfavorable consequences (Hansen and DeFries, 2004). For example, the conversion of forests and pastures into agricultural land leads to the provision of food, fiber, fuel and hosting of other crops, which has increased population throughout the human history. At the same time, destruction of forests and pastures reduces biodiversity, degrades catchment areas and increases soil pollutions.

Data of land-cover are considered as one of the main layers of information for a variety of scientific activities, managerial affairs and administrative tasks. Land-cover maps have a set of essential information that should be provided in a clear and reliable manner. Thus detailed and timely information about land-cover is essential for land change monitoring, management of ecosystems and urban planning. Technological advances, the availability of data in appropriate intervals and high-resolution satellite images have made techniques such as remote sensing and GIS very useful for conducting researches such as the analysis of detection of changes in land-cover and prediction of future scenario (Lambin et al., 2001).

Today the classification of satellite images in general is a common method for extracting information related to land-cover patterns and their changes. Several image classifier techniques have been developed, a recent comprehensive review of which can be found in Lu and Weng (2007). Selection of appropriate classification methods and imagery is an important factor for increasing the accuracy and precision of classification.

Land-cover classification by satellite images can be done based on two criteria: 1) pixels, 2) objects. While analysis based on pixels has been the dominant approach in the classification of remote sensing images for a long time, object-based image analysis has become very common in the last years (Blaschke, 2010). Some of the past studies have tested the efficiency of different image classifier methods that are based on pixels or objects.

After preparing the land-cover maps at different times, the differences between these maps must be determined using appropriate techniques. Change detection has been specified as a process of "identifying differences in the state of a phenomenon by comparing it at times A and time B" (Singh, 1989). Despite a rich archive containing the satellite images from a number of decades, it is possible to easily detect and analyze the changes that occurred over the years (Hostert et al., 2003; Kennedy et al., 2007; Röder et al., 2008). Different techniques have been employed using satellite data for land-cover change detection for many years (Rogan and Chen, 2004; Singh, 1989). Those methods divided into two groups: (1) pre-classification change detection; (2) post-classification comparison (Yuan et al., 1998). A set of comparison methods has been developed for analysis of multi-temporal satellite images and pre-classification change detection. For example it can be pointed to image regression (Ridd and Liu, 1998), vegetation index differencing (Howarth and Boasson, 1983), principal components analysis (Gong, 1993), change vector analysis (Chen et al., 2003; Johnson and Kasischke, 1998), artificial neural networks (Dai and Khorram, 1999), and classification tree (Rogan et al., 2003). These methods basically create "no-change" vs. "change" maps, but do not recognize the type of change (Lu et al., 2004; Singh, 1989). Post-classification change detection techniques identify land change by comparing produced land-cover maps at different time periods (Singh, 1989; Yuan et al., 1998). Post-classification techniques not only locate the changes, but also generate "from-to" change data (Jensen, 2005; Yuan et al., 2005). This method minimizes the problems created by difference in sensors and atmospheric situation, since images from different times are separately classified (Singh, 1989; Yuan et al., 1998).

This study focuses on comparing various machine learning algorithms, i.e. random forest (RF), decision tree (DT), support vector machines (SVM), and object-based support vector machines (OSVM), to classifying Landsat image 2010. Furthermore land-cover maps of 1990 to 2010 produced using multi-temporal Landsat images (TM and ETM+) and the best classifier method. For this purpose, multispectral images of the study area over a period of two decades have been chosen to indicate the changes in land-cover phenomena. Our specific objectives were to: (1) comparing pixel-based and object-based classification methods for land change detection, (2) mapping land-cover change in our study area between 1990 and 2010, from Landsat TM/ETM+ images, and (3) comparing satellite-based land-cover trends in two last decades (1990-2000 and 2000-2010).

2.3. Materials and methods

2.3.1. Study area

The study area is located in central Germany and covers 690000 hectares (Fig. 1). Elevation ranges from 114 to 982 m.a.s.l, with higher elevations concentrated in the Grosser Beerberg Mountain located in the Thuringian Forest. The predominant climate is of the continental type with an average annual rainfall of 604 mm, and an average annual air temperature of 8.6 °C (based on monthly recording data of 18 stations, in Free State of Thuringia from 1960–1990). The soil parent material is mainly calcareous. The landscape maps presented five classes: forest, built-up area, grassland, farmland, and water bodies (lakes, rivers, ponds, and reservoirs).

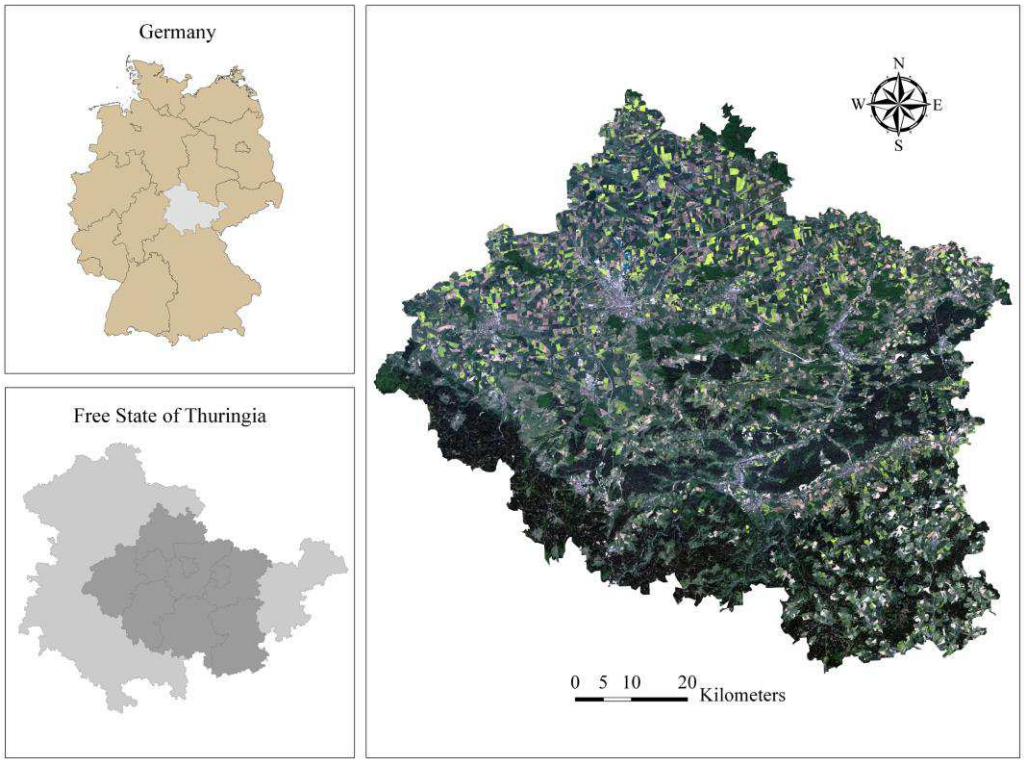


Figure 1 Location and land-cover (true color image) of study area for the year 2010.

2.3.2 Landsat image collection and pre-processing

In this study temporal coverage Landsat TM and ETM+ images from 1990 to 2010 with a standard resolution of 30×30 meters were obtained from the United States Geological Survey (USGS) archive (http://earthexplorer.usgs.gov/). Since Landsat ETM+ images from 2003 and beyond have high rates of no data values (stripping), images from 2010 were collected from Landsat TM.

Data must be preprocessed and image quality must be checked before classification and change detection analysis.Image registration should be considered before any analytical process is performed on images (in particular, change detection analysis), because slight differences created by the spatial offset directly impact the image analysis results. In this study, image registration was

performed by selecting an appropriate number of well-recognized ground control points (i.e. road intersections). A second order polynomial transformation was used to hold down the root-mean-square error (RMSE). Eventually, images with an error lower than a half-pixel (about 15 meters) were registered.

Each sensor that records visible or near-visible electromagnetic radiation generally registers a combination of two types of energy (Hadjimitsis et al., 2010; Jensen, 2013), so the numerical value of any pixel in a satellite image does not show the true ground-leaving radiance of that point on the Earth's surface. Part of this brightness is derived from the reflectance of the intended point, and the rest is due to atmospheric effects. Absorption and scattering of radiation are two major atmospheric effects that influence satellite images. In this study, the PCI Geomatica ATCOR model was used for atmospheric correction (Geomatica, 2013). Atmospheric and terrain effects were removed to determine the true ground reflectance of the Earth's surface. This model requires information, some of which is available in the metadata file, including date and time of data acquisition, sensor type, coordination of the image center, and atmospheric definition area. Atmospheric definition area was selected as “rural” regions and atmospheric condition was limited to “mid-latitude summer”.

2.3.3. Image Classification

Satellite Images classification is one of the most important steps in capturing detailed land-cover information. Determination of a suitable classification algorithm is a critical prerequisite for performing a precise classification. Therefore, the model used in this study operates in two stages. At the first stage we compared several classifiers that are considered to be suitable for land-cover image classification. Training samples were selected for training process before classification. In this study, the Jeffries-Matusita distance method is used to determine the spectral separability among different phenomena taken from training samples. After the selection of training samples, different classification algorithms were used to create the classified maps from Landsat 2010 image. Then the accuracy of classified maps was compared not only by visual observation but also by statistical methods. The second stage involved classifying all images (1990, 2000 and 2010) using the best classification algorithm identified in the first stage, and will have a process similar to that of first stage. Ultimately all classified maps will undergo an accuracy evaluation.

2.3.4. Collection of training data

A suitable classification system and adequate representative training samples are very important for a successful classification (Lu and Weng, 2007). We determined five land-cover categories (Built-up area, Forest land, Framland, Grassland, Water bodies) with visual interpretation and analysis of the satellite images. We gathered ground truth data (training and validation data) based

on Quickbird images available in Google Earth (http://earth.google.com). For whole of study area, a sample of ground truth points randomly collected within the area covered by high resolution Quickbird images, overlaid selected samples on the Quickbird images, and then grouped these points to appropriate classes based on visual interpretation. A point was assumed as an especial class if land-cover patches included at least one pixel. Based on visual interpretation of the Landsat images, the training sites were carefully determined and restricted to homogeneous regions where class membership was permanent from 1990 to 2010. We checked the separability of the training samples by Jeffries-Matusita distance measure and optimized the sample dataset until we achieved maximum stable accuracy. This optimizing task was carried out by removing training samples that may have been sources of error or collecting new samples to obviously misclassified categories. Finally, we used a sample of 1374 points were mapped from Quickbird images. We split all ground truth points into training (75%) and evaluation (25%) data. We finally used 100% of the ground truth data to produce land-cover maps of the whole study area.

2.3.5. Pixel-based image classification

Pixel-based image classification approaches either automatically allocate all the pixels in an image to land-cover types or classify them thematically pixel by pixel. In this study, three different pixel-based machine learning classifiers were applied on each data set, namely: (1) Random Forest (Breiman, 2001; Ghimire et al., 2010); (2) Decision Tree (Keshtkar et al., 2013; Quinlan, 1987); and (3) Support Vector Machines (Duro et al., 2012a; Vapnik, 1995).

2.3.5.1. Decision tree

Decision Tree (DT) is a non-parametric classification method which can deal with various types of datasets containing categorical variables. DT represents a set of constraints or conditions that are hierarchically organized and is composed of one root node (containing all data), a number of internal nodes (splits), and a number of terminal nodes (leaves). Each node in this model has only one parent node and two or more child nodes (Breiman et al., 1984). In this method, the specified decision tree induction approach determines the variable from which classification begins (root node). Then the data splitting process repeats until either all samples are assigned to a specific class or a predetermined stop condition is reached. Once the decision tree is constructed, it must undergo a cutback (or pruning) process aimed at reducing its structural complexity and improving its accuracy. Factors such as the number of leaf nodes, tree-depth, or the number of internal nodes can be considered as the measure of pruning.

This model was run in R software (R Development Core Team, 2009) using the rpart package (Therneau and Atkinson, 1997). In this study, “information gain” measure was considered for deciding between alternative splits. The rpart package has two main parameters to be adjusted: The

minimum number of observations in a node (minsplit), and maximum depth of tree (maxdepth). Although complex trees are more expressive and potentially allowing higher accuracy, but they do not generalise the data well and are more likely to overfit. Pruning the model by setting the minimum number of observations in a node or setting the maximum depth of the tree can avoid this problem. Therefore, we tried to examine several decision trees to achieve a strong model.

2.3.5.2. Random forests

RF is a nonparametric algorithm which is considered as an improved version of Classification and Regression Tree (CART). This model consists of a large number of classification trees, each constructed by taking an individual bootstrap sample from the original data set (sampling with replacement). Ultimately, the trees will be aggregated and a majority vote rule will be applied to determine the final category (Breiman, 2001), and that is why this method is also considered as an ensemble model.

To estimate the misclassification error and variable importance the samples that are not in the bootstrap sample (out-of-bag data, OOB) were used. The OOB sample composed of approximately 37% of the original data set at each bootstrap iteration. To ensure less similarity (i.e. more diversity) between the individual trees, no pruning is performed and this model allows all the trees in the forest to become fully grown (Genuer et al., 2010). RF method has two key parameters that must be adjusted: The number of tree (ntree) and the number of input variables (mtry). These two parameters must be optimized to improve the classification accuracy (Breiman, 2001).

The splitting criterion used in this study was the Gini coefficient, and the stop criteria to stop splitting, i.e. the minimum number of samples in a node and the minimum impurity in a node were set 1 and 0, respectively, in which values the decision trees will be full grown. We used a grid-search approach based on the OOB estimate of error to figure out the optimal combination for ntree and mtry parameters (Tian et al., 2009). Finally, the optimized parameters were entered into ImageRF in the EnMAP-Box to classify satellite image (Waske et al., 2012).

2.3.5.3. Support vector machines

A Support Vector Machine (SVM) is a discriminative method that classify data according to the statistical learning theory (Vapnik, 1995). The operation of the SVM algorithm is based on fitting a separating hyperplane that gives the best separation between training samples in a multidimensional feature space. On the other world, the optimal hyperplane provides the largest minimum distance (margin) between the training samples. In case of an optimal separating hyperplane, distance between the hyperplane and the nearest positive and negative training example, called the margin, and data points on the margin are known as support vectors. Instead of using all training samples to separate classes, SVM use only support vectors that describe class

boundaries (Foody and Mathur, 2004). To separate classes with non-linear boundaries, a set of kernel functions including the sigmoid, the radial basis function (RBF), and the polynomial can be used, which transfer training samples into a higher-dimensional space, where linear class separation is possible and the problem can be solved in the new space (Huang et al., 2002).

In this study, SVM was implemented using the radial basis function (RBF) kernel. The SVM implementation of ENVI 4.8 software (ITT Visual Solutions Inc., http://www.ittvis.com/) has four parameters to be adjusted: the kernel width ‘‘gamma (γ)’’, the penalty parameter (C), the number of pyramid levels to use, and the classification probability threshold value. Classification probability threshold is an important value for the SVM classifier since all rule probabilities less than this threshold are unclassified. We set zero value for this threshold that means all pixels had to be classified into one category. Also, we set zero value for the pyramid parameter, which force the model to processes the image at full resolution.

Gamma and penalty parameters are two important factors that influence the accuracy of SVM classification. By default, the inverse of the number of bands is set for the value of gamma. This default value seems to be reasonable but it is not perfect. Studies have shown that the best combination of γ and C depends on the training data and cannot be known by default (Kuemmerle et al., 2009). Also, SVM needs the normalization of numeric inputs and thus there is a data normalization prior to SVM classification (Ben-Hur and Weston, 2010). Two methods are usually well known to data normalization: linear normalization and Gaussian normalization. In this study, the linear normalization was performed as pre-processing in SVM.

2.3.6. Object-based image classification

In object-based classification methods, every classification is related to a specific scale. This means that an image can be introduced in different scales. For example, an image may be represented based on the average size of image objects, whereas a similar image may be divided into different size of objects (Walsh et al., 2008). Image segmentation illustrates an essential primary step in object-based image analysis, as the image objects (termed “image segments”) resultant from this process produce the core of an object-based classifier (Castilla and Hay, 2008). The process that identifies pixels with a similar characteristic (such as spectral similarity) and allocates them to a certain group is called segmentation (Petropoulos et al., 2012). The spectral characteristics of each object are obtained from the mean value of the pixels forming the objects in each spectral band and from different statistical information such as minimum, maximum, and standard deviation. In addition to their spectral characteristics, objects can be described and recognized by their tone, size, shape, and texture (Bock et al., 2005). All these characteristics can be used in the classification and discrimination processes of objects.

In this study, image segmentation technique consists of two key steps: (1) edge-based segmentation (2) Full Lambda Schedule merging method. This procedure starts with a multiscale edge-based segmentation to divide the images into image objects on the basis of similar spatial, spectral, and textural characteristics. Two different types of errors may happen in image segmentation including over-segmentation and under-segmentation (Kampouraki et al., 2008; Möller et al., 2007). Using a low segmentation level generally is resulted in many small segments, which in turn brings about over-segmentation. On the contrary, high segmentation levels result in a few large segments that accord with different land-cover classes. Hence, a precise analysis seems necessary to choose segmentation scale (Liu and Xia, 2010). Preventing formation of over-segmented statements, which can be a very difficult task, is one of fundamental phases in this process. The full lambda schedule model was used to solve over-segmentation problem; hence segmentation is used in the integration stage where all adjacent segmentations, given their range and location features, are integrated (Robinson et al., 2002). Merging continues if the algorithm catches a pair of adjacent regions, i and j, such that merging cost, $t_{i,j}$ is less than a described threshold lambda value, of 0 to 100. The full lambda schedule algorithm is estimated as,

$$t_{i.j} = \frac{\frac{|o_i| \times |o_j|}{|o_i| + |o_j|} \times \|u_i - u_j\|^2}{length\left(\partial(O_i, O_j)\right)}$$

where O_i is region i of the image, $|O_i|$ denotes the area of region i, u_i is the average value in region i, u_j is the average value in region j, $\|u_i - u_j\|$ is the Euclidean distance between the spectral values of regions i and j, and length ($\partial(O_i, O_j)$) denotes the length of the common boundary of O_i and O_j.

Table 1 Image object features used in object-based classification

Object features	Description
Spectral-Mean	Mean value of the pixels comprising the region in a specific band
Spectral-STD	Standard deviation value of the pixels comprising the region in a specific band
Texture-Mean	Average value of the pixels comprising the region inside the kernel
Texture-Variance	Average variance of the pixels comprising the region inside the kernel
Intensity	Intensity using the *Spectral-Mean* attributes, and is measured in floating-point values from 0 to 1.
NDVI	Normalized Difference Vegetation Index: (band4 – band3)/(band4 + band3)

We classified several image segmentations of different scales to identify the one with the highest overall accuracy. This trial-and-error approach is often utilized in object-based classifications (e.g. Dingle Robertson and King, 2011; Duro et al., 2012b; Myint et al., 2011). Following the image segmentation process, object features were selected for use in the object-based classification. There are several types of image object features that could be potentially incorporated into image analysis (Duro et al., 2012a). Selecting object features for use in object-based classification can be

based on user experience and previous studies (e.g. Duro et al., 2012a; Pu et al., 2011; Yu et al., 2006), or a feature selection method can be used prior to final classification (e.g. Qian et al., 2014; Van Coillie et al., 2007). In this study, the inclusion of object features was based on our knowledge and previous studies. Consequently, we selected out 16 object features (Table 1).
These 16 features included 12 features calculated based on the 6 multispectral bands, which is mean value and standard deviation of these bands. In addition, we chose intensity, texture-variance, texture-mean, and NDVI (Normalized Difference Vegetation Index) for classifications. Finally, for executing SVM classification method, we selected training samples for each land-cover type based on the previously segmented and merged objects. All these processes have been done in ENVI ZOOM (Version 4.8) software (ITT Visual Solutions Inc., http://www.ittvis.com/).

2.3.7. Accuracy assessment

In this study, accuracy evaluation was based on the calculation of the overall accuracy, producer's accuracy, user's accuracy, and the Kappa index. The likelihood that the classifier labels a pixel of a known class into an accurate class is called the producer's accuracy (Congalton, 1991), while the user's accuracy is a quantity representing the chance that any classified pixel to represent the accurate class. We also used McNemar test to assess the statistical significance of superiority of each classification algorithm over another. This test is based on a chi-square (χ^2) statistics with 1 degree of freedom, computed from two error matrices and given as,

$$\chi^2 = \frac{(f_{12} - f_{21})^2}{f_{12} + f_{21}}$$

where *f12* shows the number of cases that are wrongly classified by classifier one but correctly classified by classifier two, and *f21* shows the number of cases that are correctly classified by classifier one but wrongly classified by classifier two (Manandhar et al., 2009).

2.3.8. Analyzing land-cover change

Change detection is the process of identifying the alterations made in an object or phenomenon by observing it at different times. In this study, we calculated the net changes and annual changes in the land-cover within the study area to compare the status of this factor at different time periods. The net changes were obtained by pixel based post-classification change detection algorithm. The post-classification change detection method not only maps the changes magnitude, but also determines the trend of changes (from-to) between land-cover classes (Yuan et al., 2005). Net change was calculated as the difference in land-cover (in ha) between 1990 and 2010, whereas annual change rates (ACR) were calculated for each time period j as:

$$ACR_j = \left[\left(SC_j / CPB_j\right) \times 100\right] / Y$$

where SC is the sum of changes in time period j, CPB is the cover of each phenomenon at the beginning of time period j, and Y donates the number of years between image A and image B.

To assess whether land-cover change varied with altitude and slope, we classified the digital elevation model (DEM) into four classes (class-1 (<255 m), class-2 (255-393 m), class-3 (393-561 m) and class-4 (>561 m)) using Jenks natural breaks classification method and calculated percentage of net land change for each class. Likewise, we summarized net land-cover changes for three slope classes: gentle (<5°), moderate (5-10°), and steep (>10°). The Jenks optimization method has been developed to clustering data into different classes. This model tries to classify the data in a manner that the classes, while having the least internal variance, can have the greatest possible variance with respect to each other (Morris and Simon, 2012).

2.4. Results

2.4.1. Class Separability

In this study, the six reflective bands of the Landsat images were used as the reference basis for the calculation of the separability index of the collected spectra from the training sites indicating the different classes. Table 2 shows pairwise spectral separability values of different classes of training samples for 2010 image classification. Values range from 0 to 2. The closer to 2, the more separable training samples have been selected. Values more than 1.8 indicate that class pairs have good separability, while values less than 1 represent that the class pairs must be joint into one class (Petropoulos et al., 2010). Observing the values shown in Table 2, most of the class pairs are well separated from each other with values more than 1.8. Farmland and grassland have comparatively lower value (1.65); and class separability value of pair of farmland and built-up area (1.63) is also relatively lower than other pairs. Thus, no class has to be combined into others because all values are greater than 1. The selected training samples are satisfactory to be used for classification.

Table 2 Class separability of training samples of 2010 image

Separability Values	Forest	Water	Farmland	Grassland	Built-up area
Forest	*	1.89	1.92	1.97	1.96
Water	1.89	*	1.99	1.99	1.99
Farmland	1.92	1.99	*	1.65	1.63
Grassland	1.97	1.99	1.65	*	1.83
Built-up area	1.96	1.99	1.63	1.83	*

2.4.2. Tuning of machine learning algorithm parameters

For DT classifier, the minsplit was set to 5 and examined a set of maxdepth from 2 to 7, which yielded a total of 6 classified images. We did not select the minsplit value less than 5 since by

setting minsplit to too small a value, the model may run toward the risk of overfitting. Results showed that the highest overall classification accuracy (i.e., the percentage of correctly classified samples) achieved by a maximum depth value of 6.

In order to optimize ntree and mtry parameters for RF classification model, we investigated a set of mtry from 1 to 6, while the range of the value for the ntree parameter was set between 100 and 1000 with an interval of 100, which resulted in 60 different classifications. Results indicated that the ntree value of 900 combined with a mtry value of 2 produced the lowest OOB error rate (4.54%). On the other hand, the highest OOB error rate (5.81%) was produced by the combination of mtry value of 1 with ntree value of 100 (Figure 2b).

For SVM-based classifications, a grid-search approach was used to find the optimal combination for γ and C parameters. Therefore, the gamma was adjusted by considering a nested cross-validation process, where γ $\{10^{-3}, 10^{-2}, 10^{-1}, 1, 10, 10^{2}, 10^{3}\}$. Also, we set the C parameter by considering a nested cross-validation with C $\{10^{-2}, 10^{-1}, 1, 10, 10^{2}, 10^{3}\}$. Results from the grid search indicated that the γ value of 1 combined with a C value of 10 produced the highest accuracy for the SVM-based classifications (pixel and object-based methods) (Figure 2a).

Also, for object-based classifier, an iterative trial-and-error approach was used to identify the best image segmentation scale based on the highest overall accuracy. Results show that the scale value of 50 without merging to reduce the number of segments produced the highest accuracy for the object-based classification method (Table 3).

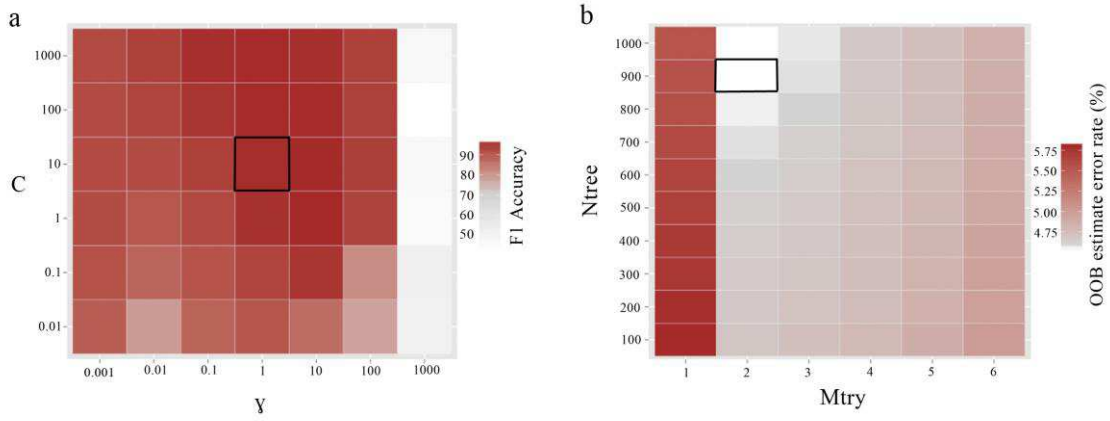


Figure 2 Heat maps resulted from grid search procedure: (a) Optimization of the SVM parameters (C and γ). The F1 measure was used to determine the best accuracy for the different combinations (n = 42) of parameters; (b) Optimization of the RF parameters (mtry and ntree). The OOB sample was used to determine the error rate for the different combinations (n = 60) of parameters.

2.4.3. Accuracy assessment and statistical comparisons

Classification was conducted on 2010 image using four different machine learning classifiers, which were DT, RF, SVM, and OSVM. The classification maps are shown in Figure 3. Analyzing

the classification maps from Figure 3 (a) to (d) visually, indicate that all classifiers can generate useful land-cover maps and produce consistent classification results.

Table 3 The classification result (overall accuracy) of values for image segmentation parameters used in the object-based method. Values in bracket are the kappa coefficient.

	Scale level						
Merge level	20	30	40	50	60	70	80
0	93.35 (0.882)	93.35 (0.882)	93.35 (0.882)	93.54 (0.884)	92.09 (0.853)	83.56 (0.715)	67.92 (0.352)
20	93.20 (0.879)	93.20 (0.879)	93.20 (0.879)	93.35 (0.880)	91.80 (0.847)	83.51 (0.697)	67.99 (0.352)
40	92.77 (0.870)	92.77 (0.870)	92.77 (0.870)	92.55 (0.863)	91.63 (0.843)	83.01 (0.691)	66.63 (0.326)
60	90.36 (0.822)	90.36 (0.822)	90.36 (0.822)	90.34 (0.819)	91.46 (0.840)	81.27 (0.678)	65.93 (0.321)
80	89.02 (0.794)	89.13 (0.794)	89.15 (0.797)	89.61 (0.803)	88.83 (0.789)	80.64 (0.656)	66.09 (0.313)

In addition to visually observing the classification maps, accuracy assessment was also performed on the classification maps to quantitatively compare the performance of these classifiers. The associated classification accuracy statistics are summarized in Table 4. According to the result, in general, classification using OSVM has highest overall accuracy and Kappa coefficient, which are 93.54% and 0.88 respectively, while DT generates the least accurate classification map with 86.36% overall accuracy and 0.76 Kappa coefficient. Classification maps generated by RF and SVM have much higher overall accuracy (90.28% and 90.93% respectively) than DT, but relatively slightly lower than OSVM.

For the OSVM the classes with the highest producer's accuracy were those of water (96.58%), forest (96.31%) and farmland (96.13%) followed by the built-up class (94.06%), whereas the lowest producer's accuracy was obtained for the class of grassland (63.36%). User's accuracy was higher for the water bodies (98.61%), the forests (96.71%) and farmland areas (94.57%) followed by grassland areas (87.74%), the lowest user's accuracy was found for the built-up areas class (78.44%).

All classes were obviously separable in all classifier algorithms applied here. But classes with relatively poor or indistinct producer's and user's accuracy were for the case of SVM and OSVM classifiers in the grassland areas, whereas for RF classification method the grassland and built-up areas and for DT classifier method the water, grassland and built-up areas.

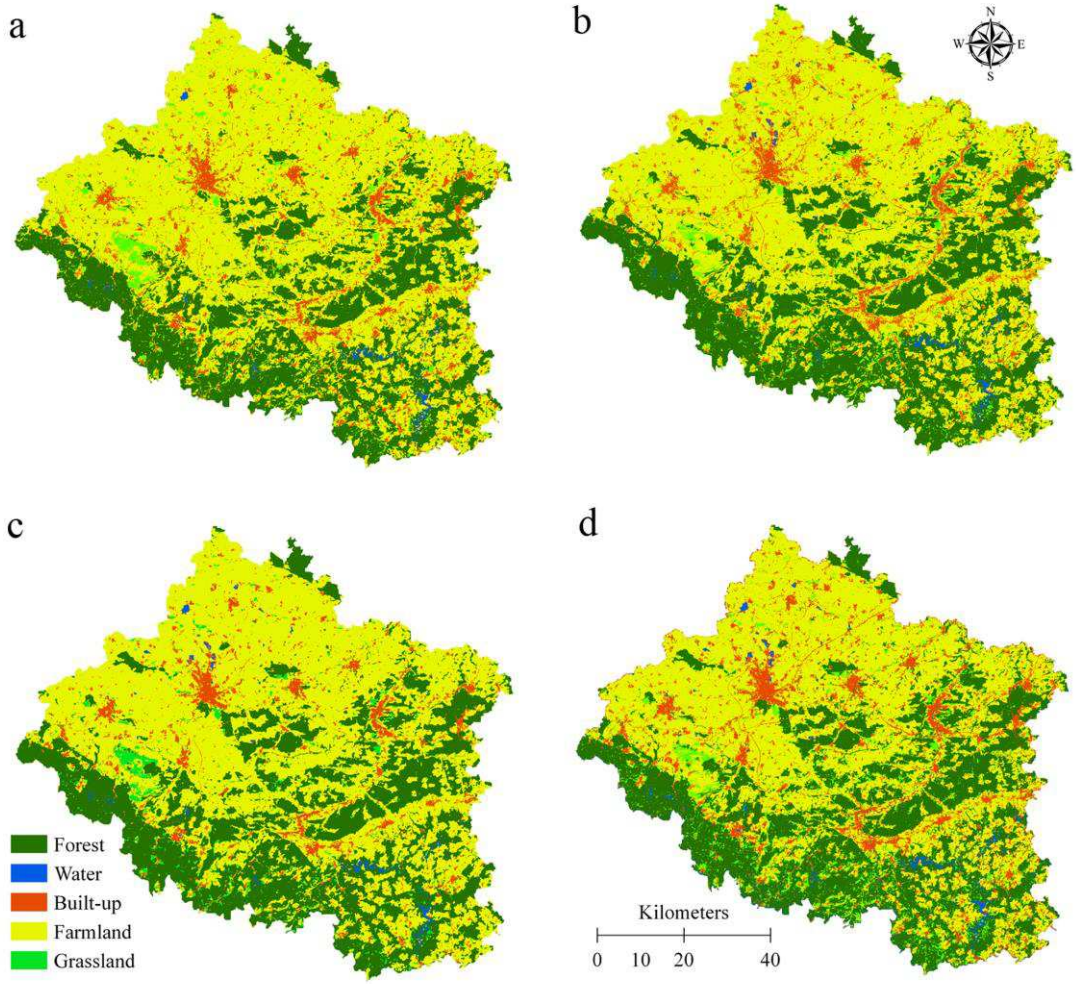


Figure 3 Land-cover maps of 2010 generated by (a) DT, (b) RF, (c) SVM, and (d) OSVM classification algorithms.

Also, we used the McNemar test to figure out whether a statistically significant difference exists between different machine learning algorithm. The McNemar test indicated that the observed difference between pixel-based image classifications was not statistically significant ($p>0.05$). For pixel-based classifier methods and object-based classifier, a statistically significant difference ($p<0.05$) between DT and OSVM algorithms ($p=0.004$) was observed, while RF and SVM algorithms did not show significant difference with OSVM method ($p>0.05$).

2.4.4. Analysis of land-cover change

Object-based classification (i.e. OSVM algorithm) was performed on three Landsat images of 1990, 2000 and 2010. Accuracy assessment result of each classification map is summarized in Table 5. By counting the number of pixels of each phenomena for each year, land-cover coverage information can be obtained, which is shown as Figure 4.

Table 4 Summary of mapping accuracy obtained by different classifiers to the Landsat TM 2010 image. OA= overall accuracy; K= kappa coefficient; UA= user's acxcuracy; PA= producer's accuracy.

	DT		RF		SVM		OSVM	
	OA (86.36)	K (0.76)	OA (90.28)	K (0.85)	OA (90.93)	K (0.86)	OA (93.54)	K (0.88)
Class	UA (%)	PA (%)	UA (%)	PA (%)	UA (%)	PA (%)	UA (%)	PA (%)
Forest	95.76	84.41	96.19	95.49	97.26	94.05	97.71	96.31
Water	99.17	56.70	99.00	93.86	99.20	94.07	99.61	96.58
Farmland	89.47	89.24	86.24	93.65	88.97	93.31	94.57	96.13
Grassland	66.39	81.53	83.54	61.06	80.77	56.33	87.74	63.36
Built-up	56.09	81.23	63.68	89.77	78.34	90.07	78.44	94.06

According to the change detection results, the most significant change occurred from 1990 to 2010 is caused by the expansion of built-up area. Analysis of land-cover area changes indicate that during this time period, built-up areas increased from 2.8% to 5.5%. The built-up land was continuously increased, and the farmland, grassland and forest were continuously decreased. Grasslands decreased significantly from 4.89% to 4.02% during 1990–2010. During this period, forest area decreased from 32.38% to 32.26%. Also, the coverage of farmlands reduced from 59.21% to 57.58% in the same time. The area of water increased a little.

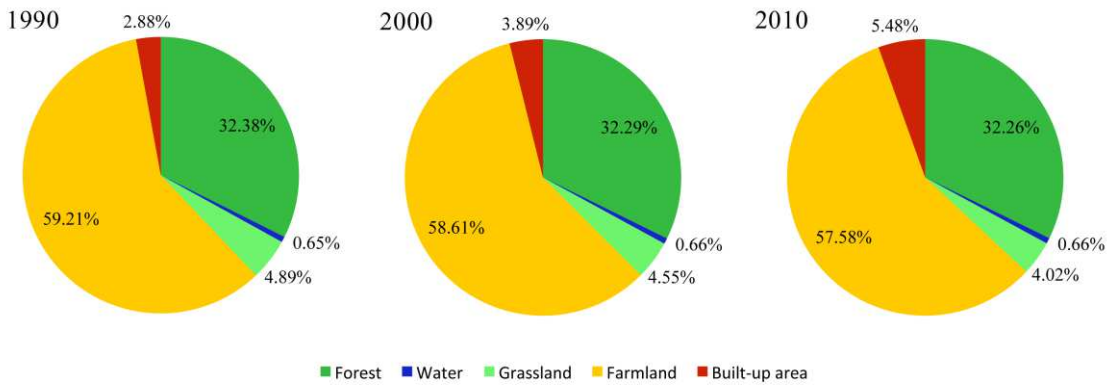


Figure 4 Pie chart of land-cover coverage (%) from 1990 to 2010.

Figure 4 only illustrates the static state of each phenomenon in 1990, 2000 and 2010. Table 6 depicts the summarized specific "from-to" change information. This table represents the amount of change from one class detected in 1990 to another class detected in 2010. The diagonal values in table represent the area with no change. Forest and farmland are moderately stable classes that don't have significant change, keeping 91.23% and 94.53% unchanged respectively. About 5.04% of forest change to farmland and 2.54% change to grasslands. As for farmland, a small portion (2.94%) of the area changes to built-up area and another small portions, i.e. 1.69%, 0.8% and 0.04%, changes to grassland, forest and water bodies, respectively. Comparatively, grasslands

experience the most dramatic change. Only about 56.35% of grasslands were kept unchanged. 19.76% and 16.3% of grasslands alter into farmland and forest, respectively.

Table 5 Summary of mapping accuracy obtained by object-based SVM classifier to the Landsat 1990, 2000 and 2010 images. OA= overall accuracy; K= kappa coefficient; UA= user's accuracy; PA= producer's accuracy.

	1990		2000		2010	
	OA	K	OA	K	OA	K
	(89.75)	(0.84)	(92.36)	(0.85)	(93.54)	(0.88)
Class	UA (%)	PA (%)	UA (%)	PA (%)	UA (%)	PA (%)
Forest	97.35	94.23	97.14	96.51	96.71	96.31
Water	99.66	96.99	96.97	96.76	98.61	96.58
Farmland	77.39	92.87	91.87	97.64	94.57	96.13
Grassland	97.16	61.65	84.35	65.37	87.74	63.36
Built-up	86.93	89.28	79.46	93.56	78.44	94.06

Table 6 Change detection classification matrix for 1990–2010 based on post classification comparison to specify 'from–to' transitions. The amount of changes is demonstrated by percentage (%). Bold type denotes that there is no change in land-cover between two dates.

		1990				
	Class	Forest	Water	Built-up	Farmland	Grassland
2010	Forest	**91.23**	15.98	2.20	0.80	16.30
	Water	0.35	**75.94**	0.88	0.04	0.57
	Built-up	0.84	3.64	**94.26**	2.94	7.03
	Farmland	5.04	3.25	0.96	**94.53**	19.76
	Grassland	2.54	1.19	1.79	1.69	**56.35**

Table 7 Distribution of land-cover classes (in hectare) and annual change rates (ACR) for 1990-2010.

	Year			ACR in land-cover structure		
	1990	2000	2010	Δ%1990–2000	Δ%2000–2010	Δ%1990-2010
Forest	223400	222810	222550	-0.26	-0.12	-0.38
Water	4460	4530	4550	1.57	0.44	2.02
Grassland	33750	31390	27750	-6.99	-11.60	-17.78
Farmland	408490	404370	397270	-1.01	-1.76	-2.75
Built-up area	19840	26840	37820	35.28	40.91	90.63

Table 7 shows the ACR of land-cover classes for three time periods, 1990-2000, 2000-2010, and 1990-2010. This table indicates that mean annual deforestation rates were three times higher in 1990–2000 compared to 2000–2010. The maximum rate of annual change in water bodies belonged to the years 1990-2000 and was about 4 times higher than the rate of the succeeding 10 years (2000-2010). Mean annual degradation rates of farmland and grassland in 1990-2000 were almost two times higher than the same rates in 2000-2010. Also, the results showed that the ACR of built-up area for 1990-2000 and 2000-2010 was 3.53% and 4.09%, respectively. Between the years 1990 and 2010, an annual average of about 4.53% of lands turned to the built-up area.

2.4.5. Land-cover changes in relation to topographic factors

The elevation distribution of each land-cover class is shown in Table 8. Elevation in the study area is mostly less than 400 m. The areas with an elevation of <255 m (Class-1), 255-393 m (Class-2), 393-561 m (Class-3), and >561 m (Class-4) accounted for 25.4%, 36.5%, 25.6%, and 12.4% of the whole area, respectively.

The results show that more than 80% of Class-1 and 60% of Class-2 lands are allocated to farmland; forest and farmland occupy approximately 47% and 42%, respectively, of Class-3 lands; and more than 70% of Class-4 land is covered by forest; only 20% of Class-4 land is farmland. The mean elevation of each land-cover class is, in ascending order, built-up area (about 290 m) < farmland (about 310 m) < water bodies (about 350 m) < forest and grassland (over 430 m). During the years 1990-2010, changes in Class-1 land reduced the area of farmland by 98.6 km^2 and increased built-up areas by 81.7 km^2. The most major change was related to the increase in built-up lands, which was as much as 60.3 km^2. The decrease of 38.5 km^2 in grassland reflected the greatest change in Class-3 land. Forest land underwent the largest change with a loss of 16.4 km^2.

The results obtained from the division of the study area based on slope show that about 61% of the area under study has a slope below 5 degrees (4220.56 km^2), and over 70% of this slope category is farmland. Additionally, over 22% of the study area (1532.28km^2) has a slope of 5-10 degrees. Forest and farmland cover about 46% and 44% of this slope gradient category, respectively. About 17% of the study area has a slope greater than 10 degrees (1146.56 km^2), about 75% of which is covered by forest (Table 9). The mean slope gradients of built-up land and farmland are less than 3.3°, while those of forest, grassland, and water bodies are more than 5°.

2.5. Discussion and Conclusion

This study compared various machine-learning algorithms (i.e. RF, DT, SVM, and OSVM) with classifying Landsat image 2010. Furthermore, land-cover maps of 1990, 2000, and 2010 were produced using multi-temporal Landsat images.

The results showed that SVM has the highest accuracy among pixel-based methods compared with two other methods (i.e. RF and DT), although the McNemar test did not show a significant difference in the performance of these three models ($p > 0.05$). It is noteworthy that both RF and SVM algorithms can obtain similar overall classification accuracies which are usually greater than those acquired using DT-based algorithms (Table 4). All in all, the classification results reported here are generally in agreement with results reported by other authors. For example, Duro et al. (2012) reported that the RF model achieved higher classification accuracies than the DT model. Pal (2005) found that both RF and SVM image classifiers produced similar classification accuracies. Conversely, the results disagreed with those described by Otukei and Blaschke (2010),

who discovered that DT generally achieved better classifications than those obtained using SVM. Adam et al.(2014) and Rodriguez-Galiano and Chica-Rivas (2012) also reported that the RF model achieved higher classification accuracies than the SVM model.

Although among the investigated models in this study, the DT method showed the least overall accuracy (83.36%), but all in all it provided an acceptable quality. Since the maps produced from the satellite data with an overall classification accuracy of 85%, are determined as acceptable (Anderson, 1976). One of the most important reasons for high accuracy of classification in the SVM method is its capability in separating hyperplanes for different classes. This model can generalize this separation system to unseen samples with the minimal error for all the separating hyperplanes and consequently can offer the best separation compared to the other models (Licciardi et al., 2009). Moreover, SVMs use training data in an appropriate space for self-assessment which are explained by a kernel function.

The ability of pixel-based and object-based methods to classify the land-cover classes was compared. Both pixel-based and object-based classifiers produced comparable overall accuracies. Object-based classifier achieved higher classification accuracy than other methods by about 3-7% (Tables 4). In this study, the classifications resulted from the pixel-based or object-based image analysis yielded almost similar results of different phenomena from the visual point of view. According to the results of the McNemar test, there is no statistical reason for preferring pixel-based methods to the object-based classifier. As expected, the object-based classifier approach (i.e. OSVM) in comparison to the pixel-based classifications obtained a more generalized visual appearance and more contiguous representation of land-cover, which possibly better shows how land-cover interpreters and analysts recognize the landscape (Stuckens et al., 2000).

One of the weaknesses of pixel-based methods is the ‘‘salt and pepper’’ effect (Fung et al., 2008). The restriction of the mentioned effect is not a problem in the object-based method. A combined usage of segmentation using information gained from image objects was resulted in rather more precision in the object-based method. Segmentation plays a key role in declining the "salt and pepper" effect considerably. Class discrimination was comparatively higher using the object-based method compare to pixel-based methods, as showed in the higher user accuracy for different classes (Table 4). Parallel with our results, different authors have pointed to superiority and advantages of the object-based methods rather pixel-based classifiers (e.g. Benz et al., 2004; Fung et al., 2008; Petropoulos et al., 2012). Although the accuracy of classification is an important feature for selecting the classification methods, choosing an image analysis approach is not always based on its accuracy (Duro et al., 2012a). In situations in which the statistical difference among the classification algorithms is low the end-user in selecting these models may consider other factors. For example, being cost free, user-friendly and availability may be considered which might encourage the user in selecting a specific model.

Our results show that, overall accuracies derived from the implementation of the OSVM classifier to the Landsat images of 1990, 2000 and 2010 were, respectively, 89.75%, 92.36% and 93.54% (Table 5), thus indicating the suitability of the classified remote sensing images for effective and reliable land-cover change analysis and modeling. Figure 6 illustrates the produced land-cover maps.

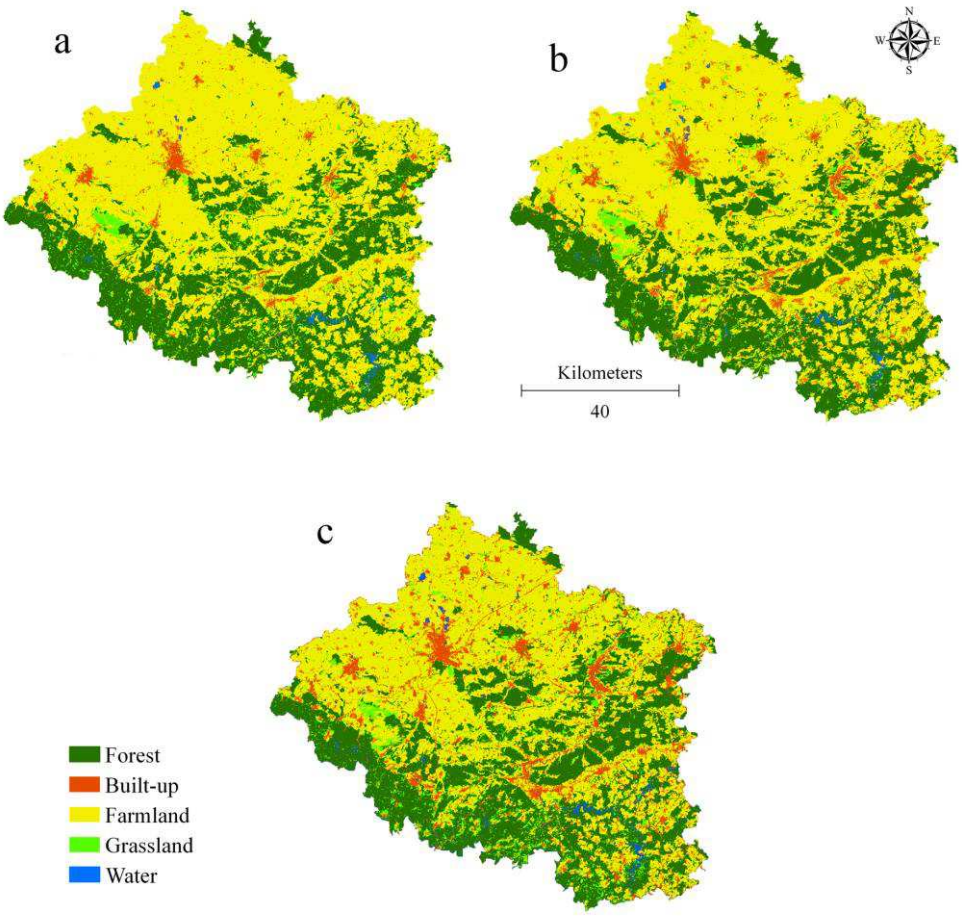


Figure 6 Time series of detailed land-cover maps for (a) 1990, (b) 2000, and (c) 2010.

The results of this study showed that natural lands affected by human activities are rapidly transforming and being damaged. The net built-up cover growth for 1990–2000 and 2000–2010 period was 7000 and 10980 hectares, respectively. Built-up area has continued to expand since 1990 at the average rate of 4.53% per year. The main driving force for built-up area expansion would appear to be the implement of the urban development, termed “Critical Reconstruction”, in the region since 1990 (Tölle, 2010). This in turn created stress to natural lands like forests and grasslands.

The results showed that between the years 1990 to 2010, the grasslands are significantly converted to farmlands, forests and built-up lands and only 56% of the grasslands are remained unchanged (Table 6). However, during the same time only about 7% of other lands are converted to the

grassland, which mainly is a consequence of deforestation. Table 7 shows that in the first decade (1990- 2000), on average, about 0.7% of grasslands' area is reduced annually, a number which has become more than two times (%1.16) in the second decade (2000-2010). These results confirmed the high vulnerability of grasslands in the study area.

Post-classification change detection results showed that the destructed forests mainly converted to farmland and grassland. Noting the data in table 6 it can be concluded that deforestation and reforestation are happening concurrently, but the speed and amount of deforestation is higher. Although, the results of the study on two different decades showed that the amount of deforestation in the second decade experienced a significant reduction and reduced to one third of the first decade.

As seen in Table 8, results show that with increasing altitude, land-cover changes were reduced. Major changes happened in Class-1 and Class-2; approximately 100 km^2 and 66 km^2, respectively, were transformed from one state to another. In the two classes at higher elevations, a total of 61 km^2 of land has changed. Built-up areas grew significantly from 198 km^2 to 378 km^2 between the years 1990-2010 (Table 7). This phenomenon covered 81.7 and 60.3 km^2 of the total land area of Class-1 and Class-2, respectively. In the same period, farmland lost a total of 2.75% of total area; in Class-1 and Class-2 alone, 98.6 and 32.8 km^2 of farmland were transformed into other types of land, respectively. At higher elevations, however, the surface of the area increased 19.2 km^2, because the development in built-up land changed the use of a large part of farmland (Table 8). However, the results do indicate that about 81% of Class-1 lands remained under farmland use by 2010.

Table 8 Land-cover changes in relation to elevation during 1990-2010.

		Forest	Built-up	Farmland	Grassland	Water
Mean (m)	1990	475	277	306	434	350
	2010	472	295	310	430	340
	Changes	-3	18	4	-4	-10
1990 (km^2)	Class 1	81.3	105.2	1526.2	30.4	10.0
	Class 2	690.4	66.3	1642.6	108.1	14.2
	Class 3	837.5	22.0	750.2	138.4	19.3
	Class 4	624.8	4.8	165.9	60.7	1.2
2010 (km^2)	Class 1	79.9	186.9	1427.6	46.2	12.5
	Class 2	696.6	126.7	1609.8	76.0	12.6
	Class 3	840.6	54.9	753.9	99.9	18.0
	Class 4	608.3	9.7	181.4	55.5	2.4
Land-cover changes (km^2)	Class 1	-1.4	81.7	-98.6	15.8	2.5
	Class 2	6.1	60.3	-32.8	-32.1	-1.6
	Class 3	3.1	32.9	3.8	-38.5	-1.3
	Class 4	-16.4	5.0	15.4	-5.2	1.2

Most grasslands were located at high altitudes, but their area decreased in the years 1990-2010 in all height classes (except Class-1).Grassland in Class-2 and Class-33 has largely been replaced by built-up land. This may indicate that the development of built-up areas and their penetration into

grassland is legally much easier than the penetration into forest or farmland; perhaps grasslands have the features necessary for urbanism.

The results show that as slope increased, forest areas also increased (Table 9). For example, about 40% of the forest land is located on a slope greater than 10 degrees. Unlike forestland, however, the area covered by other phenomena decreased as slope increased. This is especially pronounced in the case of farmland and built-up areas. More than 75% of farmlands are located at a slope below 5 degrees, and only about 5% appear on slopes greater than 10 degrees. Moreover, more than 85% of built-up areas occupy slopes below 5 degrees. Although built-up areas spread significantly in areas with a slope of less than 5 degrees from 1990 to 2010 (157 km^2), the development of these lands has expanded into areas with higher slopes. The development of built-up areas and farmland has caused the destruction of natural lands, especially grasslands and forests, which have lost 7 and 15 km^2 of their area, respectively.

Table 9 Land-cover changes in relation to slope during 1990-2010.

		Forest	Built-up	Farmland	Grassland	Water
Mean (degree)	1990	9	3	3	7	6
	2010	9	3	3	7	5
	Changes	0	0	0	0	-1
1990 (km^2)	<5°	681.9	168.5	3194.6	151.6	23.9
	5°-10°	679.9	26.4	713.4	101.8	10.8
	>10°	592.5	2.7	150.9	58.9	6.7
2010 (km^2)	<5°	686.5	325.5	3054.0	127.9	26.7
	5°-10°	681.9	38.6	728.8	72.3	10.6
	>10°	583.4	11.9	161.7	48.0	6.5
Land-cover changes (km^2)	<5°	4.6	157.0	-140.6	-23.6	2.7
	5°-10°	2.1	12.2	15.5	-29.5	-0.2
	>10°	-15.1	10.6	13.0	-6.8	-1.6

In general, it should be noted that during the years 1990-2010, changes in land-cover occurred generally in areas with a slope of less than 5 degrees; that includes about 54% of total changes in these years.

Chapter 3

A spatiotemporal analysis of landscape change using an integrated Markov chain and cellular automata models

Hamidreza Keshtkar and Winfried Voigt

This manuscript has been published in 2015, Model. Earth Syst. Environ. 2: 1-13.

3.1. Summary

Spatially land-cover models are necessary for sustainable land-cover planning. The expansion of human-built land involves the destruction of forests, meadows and farmlands as well as conversion of these areas to urban and industrial areas which will result in significant effects on ecosystems. Monitoring the process of these changes and planning for sustainable use of land can be successfully achieved by using the remote sensing multi-temporal data, spatial criteria and predictor models. In this study, land-cover change analysis and modeling was performed for our study area in central Germany. An integrated Cellular Automata-Markov Chain land change model was carried out to simulate the future landscape change during the period of 2020–2050. The predictive power of the model was successfully evaluated using Kappa indices. As a consequence, land change model predicts very well a continuing downward trend in grassland, farmland and forest areas, as well as a growing tendency in built-up areas. Hence, if the current trends of change continue regardless of the actions of sustainable development, drastic natural area decline will ensue. The results of this study can help local authorities to better understanding the current situation and possible future conditions as well as adopt appropriate strategies for management of land-cover. In this case, they can create a balance between urban development and environmental protection.

Keywords: Land-cover change, Markov chain, Cellular automata, Multi criteria evaluation.

3.2. Introduction

Many interacting components affect the global environment change and land-cover change is probably one of the most important components which has a significant impact on ecological systems (Vitousek, 1994). Land-cover has long been faced with changes and probably will change in the future as well (Ramankutty and Foley, 1998). These changes are occurring in different scales (local to global) and in different time periods (days to millennia) (Townshend et al., 1991). Given that regional and/or local land-cover changes can be used as an important data layer in ecological and environmental models (such as species distributions, climate change, sustainable

development policies, spatial planning and flood risk assessment), the update presence of this data is of great importance (Castella et al., 2007; Leinenkugel et al., 2013; Funkenberg et al., 2014; Kuenzer et al., 2014).

The unprecedented rate of land change has become a major concern around the world that's why this issue has affected the environmental services and biodiversity at the global level. Both anthropogenic and natural forces are responsible for these changes in Earth's surface. Anthropogenic forces such as urban expansion and the destruction of forests and meadows for economic purposes (development of agricultural land); and natural forces such as fire, flood and tsunami; have changed the type of land-cover and land-cover all over the world. In recent decades, the changes caused by anthropogenic forces have found a faster pace than natural variations. This is because technological development and population growth are the two main factors which are responsible for the anthropogenic changes and has been unprecedented growth in past two decades. As a result, human has significantly changed almost all the world's ecosystems or is going to change them; and therefore the capacity of ecosystems to provide goods and services is going to be reduced (Lambin and Meyfroidt, 2011).

Rapid landscape changes in the recent decades have negative effects on biodiversity (Pimm and Raven, 2000; Sala et al., 2000; Balmford et al., 2001), soil erosion (Sidle et al., 2006), destabilization of watersheds (Rai and Sharma, 1998), increasing levels of greenhouse gas emissions (Macedo et al., 2013), water pollution, and air pollution (Houghton, 1994). Considerable evidence has also shown that these changes have an observable impact on the geographical distribution of species. Land-cover changes can affect distribution of plant species directly through changing the quality and quantity of habitat suitability and indirectly via increasing, decreasing, or eliminating dispersal barriers.

Understanding land change trends is a subject of interest and concern among environmental managers and ecologists (Bagan and Yamagata, 2012). Predicting landscape change is a necessary but difficult process because it needs access to a large amount of information about the relationship between phenomena and the factors which influence changes, and also sufficient information about the current and past status of the landscape. Remote sensing prepares a great source of data, from which updated land-cover maps and changes can be analyzed and predicted efficiently. With recent advances in geographic information systems (GIS) and remote sensing tools and modules enable researchers to predict future land-cover changes effectively.

Several statistical and geospatial models have been used to model land-cover change, including logistic regression models (Hu and Lo, 2007), neural networks (Pijanowski et al., 2002; Basse et al., 2014), Markov chains (Kamusoko et al., 2009), and cellular automata(CA; Poelmans and Van Rompaey, 2010). These techniques are often combined together to produce a hybrid model.

In this research, we applied a cellular automata-Markov chain model (CA-Markov) to simulate future land-cover changes. Both cellular automata (CA) and the Markov chain model have great advantages in the study on land-cover changes (Sang et al., 2011). Markov-Chain model is one of the most widely used methods for quantifying the probability of land-cover change from state A to state B (e.g. forest to built-up area) in discrete time stages. These probabilities then enter into the cellular automata (CA) model to predict spatial changes over a specific time period (Mitsova et al., 2011; Yang et al., 2012). CA-Markov model is based on the initial distribution and transition matrix; it assumes that the drivers, which have created the current situation for the region land-cover, will continue to operate as before in the future (Guan et al., 2011). In many studies, the combination of remote sensing and GIS are effectively used in CA-Markov model (Mitsova et al., 2011; Subedi et al., 2013).

The objective of this study is to simulate future land-cover changes based on the CA-Markov model in our study area which is located in central Germany. Firstly, the Markov model is used to estimate transition matrices from the land-cover maps (1990, 2000 and 2010) to predict changes of land-cover. Secondly, an integration evaluation procedure is used to generate transition suitability maps based on change drivers. Finally, transition suitability map and transition matrix are carried out in the CA-Markov model to predict spatial distribution of land-cover from 2010 to 2050.

3.3. Materials and methods

3.3.1. Study areas

The study area is located in central Germany and covers 690000 hectares (Fig. 1). Elevation ranges from 114 to 982 m.a.s.l, with higher elevations concentrated in the Grosser Beerberg Mountain located in the Thuringian Forest. The predominant climate is of the continental type with an average annual rainfall of 604 mm, and an average annual air temperature of 8.6 °C (based on monthly recording data of 18 stations, in the Free state of Thuringia from 1960–1990). The soil parent material is mainly calcareous. The landscape maps presented five classes: forest, built-up area, grassland, farmland, and water bodies (lakes, rivers, ponds, and reservoirs).

3.3.2. Modeling framework

In this section, we describe the main components used for the land-cover changes in future. The process occurs in a raster data environment, most often a grid of uniform cells of a specified resolution. The workflow that was carried out in this study consists of: 1) land-cover mapping of 1990, 2000 and 2010 using the classification of satellite images, 2) computation of transition area matrix derived from a Markov process, indicating the number of pixels to be expected to change each land-cover class to another class over a specified time interval (1990-2000, 2000-2010); 3) getting transition suitability images by Markov chain and multi-criteria evaluation (MCE) model

(These suitability images imply the suitability of each cell for a particular land-cover); 4) Evaluating the predictive power of the model by comparing the difference between the actual and projected maps of year 2010; and finally, 5) land-cover change simulation using CA-Markov module for 2020, 2030, 2040, and 2050. Land-cover mapping (first step) was run in ENVI ZOOM (Version 4.8), and other steps were calculated in IDRISI-Selva software (https://clarklabs.org/).

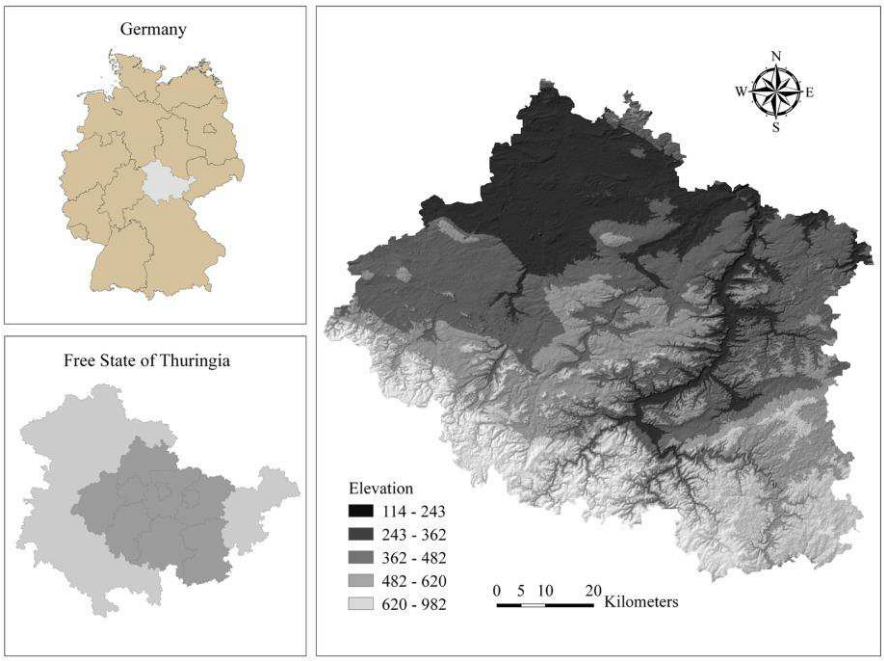


Figure 1 Location and elevation model of study area.

3.3.2.1. land-cover mapping

A temporal coverage of Landsat TM and ETM+ images (USGS Global Visualization Viewer) from 1990 to 2010 was collected. Using the landsat images in 1990, 2000 and 2010, the land-cover maps were generated for the three corresponding years. To remove the distortions, noises, and errors produced during the imaging process, pre-processing techniques (both geometrically and atmospherically) were applied to all the images. After geometric and atmospheric corrections, the land-cover maps were derived from object-based support vector machine (SVM) classification method (Duro et al., 2012). One of the advantages of object-based classification methods is that in contrast to pixel-based classification methods, there is no "salt-and-pepper" effect in images classified by them. The landscape maps presented five classes: forest, farmland, grassland, built-up area, and water bodies.

3.3.2.2. Generating transition area matrix

In this study, two pairs of land-cover images (1990-2000 and 2000-2010) were applied to calculate the transition area matrices of land-cover types during the two corresponding periods. Each matrix records the number of pixels that are expected to vary from a class to another class in a specified

period in the future. This part of the model according to the trends observed in the past, is used to estimate the replacement rate of one class by another class. These matrices are obtained using the Markov Chain model with a proportional error of 0.1.The transition area matrices for the year 2010 were created by overlaying the 1990 and 2000 classifications and delineating the change between the two time periods on a class-by-class basis. This information is used as the input of the Markov model to assist in determining the possibility of conversion of any pixels of a land-cover class (e.g., forest) to other land-cover classes (e.g., farmland) and vice versa.

3.3.2.3. Generating transition potential maps

Since access to information as well as sufficient and accurate data for drawing transition potential maps for land-cover classes is difficult, drawing these maps is generally a difficult task. Incorporating all types of factors or constraints that exist within the study area seems is impossible. In this paper, transition potential maps of land-cover types were extracted by using GIS algorithms, multi-criteria evaluation (MCE), and fuzzy membership functions. Firstly, two drivers including neighborhood interaction (Euclidean distance to the same type cell) and conditional probability image were selected to compute transition potential maps of forest, grassland and farmland areas. As a rule of thumb, the pixel closer to an existing land-cover class has the higher possibility to change into that particular class. Since this rule cannot be applied to all situations (Ahmed and Ahmed, 2012), the conditional probability images are used for each category to reduce uncertainty in the transition potential maps. The conditional probability images show, to what extent, each pixel in the next time period will likely belong to the designated category; and since this probability is conditional on their current state, they are referred to as conditional probability images. Therefore, these images are a visual presentation of the transition probability matrix (El-Hallaq and Habboub, 2015). Restrictions for forest, grassland, and farmland were the built-up areas and water bodies. Finally, four typical biophysical and proximate drivers including slope, distance from nearest road, neighborhood interaction, and distances from water bodies were selected to compute transition potential map of built-up areas (Table 1). Also, water bodies considered as restriction area for built-up lands. Studies have shown that these ancillary data are closely related to the probability of urban changes (He et al., 2013; Yang et al., 2014).

Since Markov chain does not locate the occurrence of land-cover transitions, GIS algorithms, multi-criteria evaluation (MCE), and fuzzy membership functions were applied to determine the suitability and locations of transitioning cells. The fuzzy algorithms can create a standard index and prevent the selection of unknown Boolean constraints or cut-off values (Eastman, 2006). Hence, the fuzzy membership functions (e.g., sigmoidal monotonic decrease function) were used to rescale driver maps into the range 0–255, where 0 represents unsuitable sites and 255 represents the most suitable sites. Also, in this study, we use an Analytic Hierarchy Process (AHP) to

determine the weight of driving factors with the use of pairwise evaluations (Malczewski, 1999). The AHP enables land-cover transition potential based on a collection of potential maps (e.g., distance from water sources or magnitude of altitude), and considers growth limitations. The AHP affords a comprehensive and logical method to solve the decision problem, characterizing and quantifying its components, correlating the related components towards overall targets, and assessing alternative solutions. This GIS-based AHP is a strong tool because of its high ability to incorporate different types of heterogeneous variables and its simplicity to gain the weights of suitable variables (Hafeez et al., 2002; Ying et al., 2007). This model has a unique advantage when the quantification and comparison of important variables is difficult, or where the establishment of communications between working team members becomes problematic by their various specializations, terminologies, or perspectives. Because the areas of water is small, transition potentials to water is not computed. A set of transition potential maps are displayed in Fig. 2.

Table 1 Extracted weights based on AHP and fuzzy standardization for built-up areas

Factors	Functions	Control points	Weights
Distance from roads	J-shaped	0-50 m highest suitability 50-1500 m decreasing suitability >1500 m no suitability	0.28
Distance from water bodies	Linear	0-100 m no suitability 100-7500 m increasing suitability >7500 m highest suitability	0.15
Distance from built-up areas	Linear	0-100 m highest suitability 100-5000 m decreasing suitability >5000 m no suitability	0.38
Slope	Sigmoid	0% highest suitability 0-15% decreasing suitability >15% no suitability	0.19

3.3.2.4. Model evaluation

Performance evaluation of the model is one of the most important stages of the modeling although so far there has been no consensus on the evaluation criteria of landscape change models (Pontius, 2000). The model is evaluated to detect whether the projected land-cover map is giving any abrupt result or not. To validate the operation of a model the simulated map compares with the real conditions. This method has been favored in other studies such as by Araya and Cabral (2010) who used it to verify the accuracy of a model predicting land-cover change. The comparisons between the actual map of 2010, which was obtained through remote sensing techniques, and the projected map of year 2010, which observed using changes between 1990 and 2000 images, has been performed using Kappa variation statistics. The kappa statistics assess the model accuracy in terms of the quantity of cells properly classified along with the location of the cells. Its range is from 0 (random location) to 1 (perfect location specification) (Pontius, 2000).

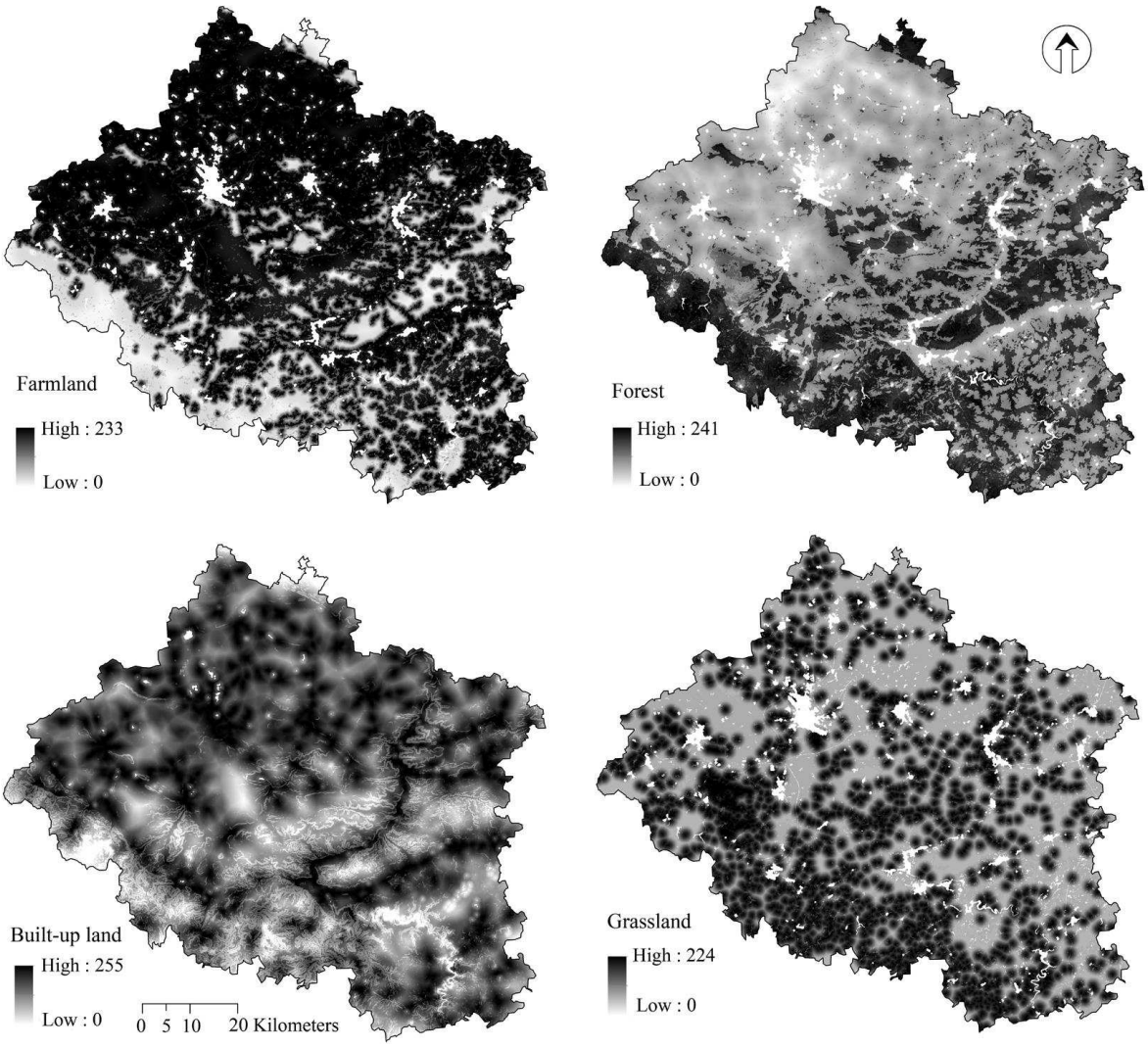


Figure 2 Transition potential maps of land-cover type in 2010.

3.3.2.5. CA–Markov Model

CA-Markov modeling is a hybrid modeling technique that binds the strengths of a spatially explicit, deterministic modeling framework with a stochastically based temporal framework. This model is a combination of Markov chain and cellular automata (CA) models which has become a robust method in terms of dynamic spatial phenomenon's simulation and future land-cover change prediction in time and space based on their current state and on ancillary information which may drive future transitions among land-cover classes. These results can in turn be used for theoretical constructions and for scenario-based projections by recalibrating the ancillary data. Markov chain is a powerful model and when the description or procedure of landscape changes is ambiguous, this model can predict the demand for and probability of landscape change using the history of

changes happened in the past. This model is one in which the future state of environment can be analyzed solely according to the previous state. Markov chain model is a stochastic process model that describes how likely one state is to change to another state and use this as the basis to project future changes. This task can be done by the transition probability matrix of landscape change through a period of time which shows that the nature of changes in the past years can be used to predict future landscape change. In this model, the transition probability can be seen for each phenomenon, but no information is supplied on the spatial distribution of these phenomena. Thus, the CA is used to characterize the spatial characters. CA model is comprised of a regular lattice framework where any cell in the lattice is in one of a defined number of states. These states either remain in their current state or change at every iteration or time step into a different state (O'Sullivan and Unwin, 2003). The changes are initiated by a set of deterministic rules that are defined prior to the execution of this process. There are four parameters needed to run cellular automata: 1) a cellular (or grid) space, 2) a neighborhood definition, 3) a set number of states, and 4) a set of transition rules. The strength of cellular automata is that it can robustly simulate processes that play out across time and space in human and in natural systems. As such they offer a useful framework for exploring system interactions (White and Engelen, 1994). Hence, CA manages spatial dynamics via local transition rules, while Markov processes depict temporal dynamics of land-cover classes based on transition probabilities (see Appendix A and Eastman 2006; for more information).

To generate future land-cover maps, the suitability images are coupled with base land-cover and the transition matrices in a process called multi-object land allocation. This process compares all pixels and their suitability for each land-cover class. Each pixel has the potential to be populated by each land-cover class during the simulation (except by restricted and unchangeable area). The class which has the highest suitability at that pixel will be the class that is chosen given the prior spatial constraints of the cellular automata and the temporal step to be classified for the stated time period. The process executes for each land-cover class and runs through the process several times at each time period. By subtracting the least likely pixels to be included in each land-cover class, the process continues until the correct number of pixels has been identified for the land-cover class under investigation. Because this process has random elements, an iterative process was used to create the potential land-cover class for each period. In order to gauge which areas are most likely to be another area, several iterations of this process were run and then combined into a frequency image. This image is the overlay of all the iterations for a given land-cover class at a given time period and shows the proportion of times each cell was classified as a given land-cover class.

3.4. Results

3.4.1. Land-cover Classification and Accuracy Assessment

An object-oriented image analysis was applied to produce a multi-temporal land-cover geographic database for the three years under study. In order to use the derived maps for further change analysis, the classification accuracy were estimated. To assess the accuracy of classified images, we gathered ground truth data (training and validation data) based on Quickbird images available in Google Earth (http://earth.google.com). For whole of study area, a sample of ground truth points randomly collected within the area covered by high resolution Quickbird images, overlaid selected samples on the Quickbird images, and then grouped these points to appropriate classes based on visual interpretation. A point was assumed as an especial class if land-cover patches included at least one pixel. Based on visual interpretation of the Landsat images, the training sites were carefully determined and restricted to homogeneous regions where class membership was permanent from 1990 to 2010. We optimized the training sample dataset until we achieved maximum stable accuracy. This optimizing task was carried out by removing training samples that may have been sources of error or collecting new samples to obviously misclassified categories. Finally, we used a sample of 1374 points were mapped from Quickbird images. We split all ground truth points into training (75%) and evaluation (25%) data. Overall accuracies for the extract land-cover maps of 1990, 2000 and 2010 were, respectively, 89.75%, 92.36% and 93.54%, thus indicating the suitability of the classified remote sensing images for effective and reliable land-cover change analysis and modeling. We finally used 100% of the ground truth data to produce land-cover maps of the whole study area. Fig. 3 illustrates the produced land-cover maps.

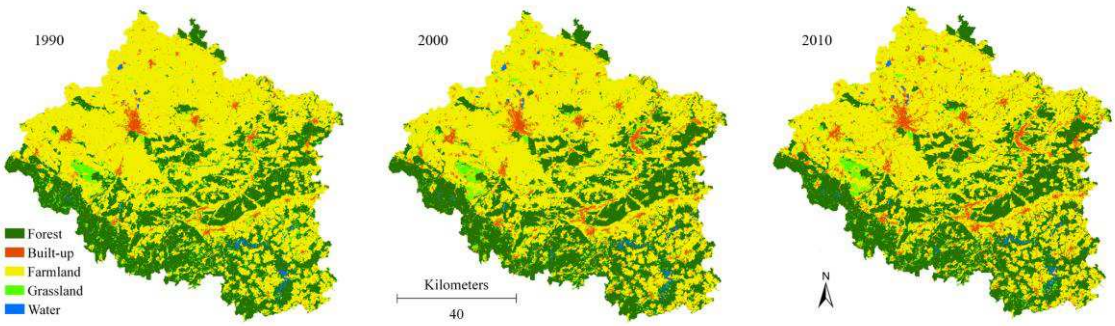


Figure 3 Time series of land-cover maps for 1990-2010.

3.4.2. Analysis of Landscape Metrics

Analysis of land-cover area changes in table 2 indicate that from 1990 to 2010, built-up areas increased from 2.8% to 5.5%. For the period between 1990 and 2000, around 7000 ha have been

changed to built-up lands, and 10980 ha within the period 2000–2010. The built-up land was continuously increased, and the farmland, grassland and forest were continuously decreased. Grasslands decreased significantly from 337.5 ha (4.89%) to 277.5 ha (4.02%) during 1990–2010. During this period, forest area decreased from 223400 ha (32.38%) to 222550 ha (32.26%). Also, farmlands reduced by 2.75% from 1990 to 2010. Overall, farmlands lost around 11220 ha in this time period. The area of water increased a little.

Table 2 Distribution of land-cover classes (in hectare) and percentage of changes for 1990-2010.

	Year			Change in Land-cover Structure		
	1990	2000	2010	Δ%1990–2000	Δ%2000–2010	Δ%1990-2010
Forest	223400	222810	222550	-0.26	-0.12	-0.38
Water	4460	4530	4550	1.57	0.44	2.02
Grassland	33750	31390	27750	-6.99	-11.60	-17.78
Farmland	408490	404370	397270	-1.01	-1.76	-2.75
Built-up land	19840	26840	37820	35.28	40.91	90.63

3.4.3. Land-cover Modeling and Validation

Evaluation of model was performed by comparing the simulated map of 2010 with the real land-cover map of 2010 based on Kappa variations. The change trajectories between the observed and simulated land-cover classes for the year 2010 are shown in Fig. 4, in which five land-cover classes have relative errors lower than 5%.

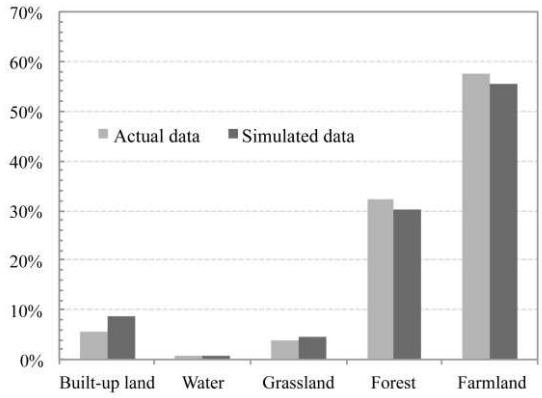


Figure 4 Actual and simulated land-cover classes for the year 2010.

Models with accuracies in excess of 80% are typically considered very strong predictive tools (Araya and Cabral, 2010). The $K_{standard}$ value was 87.6%, which verifies the accuracy of this model. Pontius (2000) states that the K_{no} value is a better alternative than $K_{standard}$ for assessing the overall accuracy of the model. The model performed very well in its overall ability to predict land-cover map of 2010 (K_{no}=91.5%), and the $K_{location}$ value of 92.2% indicates that the model provides a reasonable representation of location. Also, visual interpretation of the results (Fig. 5) shows that there is an evident similarity between the real and simulated maps for the year 2010. Therefore,

based on the Kappa values obtained, the CA-Markov model can be used to simulate future land-cover conditions.

In this research, patterns and tendency of land-cover changes were modeled according to the preceding land-cover states. Although the probability of transformation of a phenomenon to another phenomenon is determined in this model, it does not allow for prediction of spatial distribution of phenomena. Hence, the Markov model needs to be integrated with the CA model in order to add spatial characters to the model and to overcome this inherent limitation. In effect, by defining the land-cover map of 2010, the transition suitability maps derived from MCE analysis and Markov model (conditional probability images), transition area matrices of the land-cover maps of 2000–2010, a contiguity filter selection (5×5 Moore neighborhood kernel) to define neighborhood interactions, and one iteration per year were employed to predict the future changes in 2020-2050. The contiguity filter down-weights the suitability of pixels that are far from existing areas of each land-cover class. The role of this filter is to ensure that the best choices for land-cover transformation are limited to cells that are both inherently suitable and in close proximity to existing areas of that land-cover class; this gives preference to contiguous suitable areas. In each iteration, pixels with the most transition probability to transfer from one category to another category turn into a new category; while pixels with lower probabilities remain unchanged. If 10 iterations are selected for the model, the model allocates one tenth of all cells which are expected to be transferred to another category during each repetition (Eastman, 2006). The multi-objective land allocation (MOLA) procedure was used to resolve the land allocation conflicts. All land-cover classes act as claimant phenomena and contend for land within the host class (Eastman, 2006).

Table 3 Transition probability matrix of land-cover types for the periods 1990–2000 and 2000–2010.

	Land-cover type	Forest	Built-up land	Farmland	Grassland	Water bodies
1990-2000	Forest	0.9346	0.0087	0.0193	0.0349	0.0025
	Built-up land	0.0102	0.9276	0.0428	0.0118	0.0076
	Farmland	0.0046	0.0602	0.8686	0.0657	0.0008
	Grassland	0.1332	0.0417	0.061	0.7627	0.0014
	Water bodies	0.078	0.0185	0.0056	0.0001	0.8978
2000-2010	Forest	0.9247	0.003	0.0404	0.0295	0.0024
	Built-up land	0.0094	0.9643	0.0076	0.0122	0.0065
	Farmland	0.0107	0.0341	0.9076	0.0475	0.0001
	Grassland	0.1095	0.0424	0.0496	0.7982	0.0003
	Water bodies	0.0135	0.0083	0.005	0.0017	0.9715

The outcome of this process was a rendering of a potential land-cover distribution at the specified time of 40 years into the future at four steps of 10 years. Ten years for each time step was chosen as it corresponded to the time step by which the transition matrix was constructed (between the years 2000-2010). Firstly, 2010 year is set as starting year; transition area matrix of 2000–2010 periods is used to simulate 2020 year land-cover change; then, 2020 year is set as starting year;

transition area matrix of 2000–2010 periods is used to simulate 2030 year land-cover change; thirdly, 2030 year is set as starting year; transition area matrix of 2000–2010 periods is used to simulate 2040 year land-cover change; finally, 2040 year is set as starting year; transition area matrix of 2000–2010 periods is used to simulate 2050 year land-cover change. The forecasted land-cover maps for 2020 to 2050 are displayed in Fig. 6.

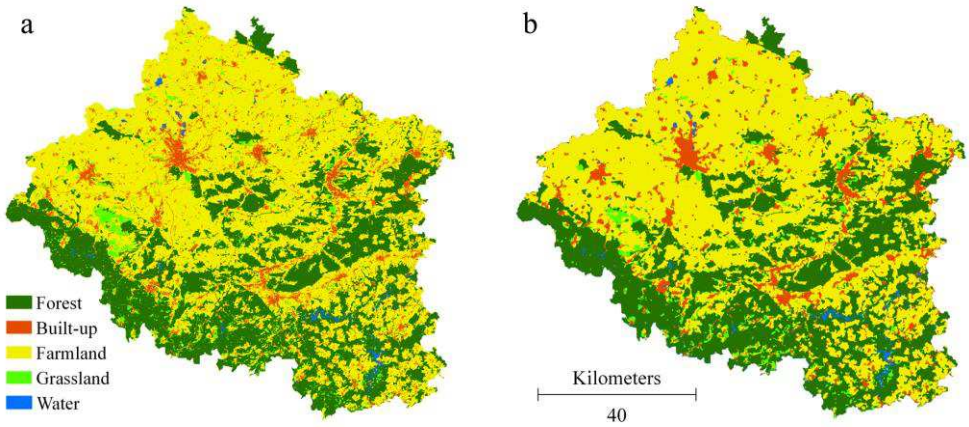


Figure 5 (a) Actual map and (b) simulated map of land-cover type in 2010.

3.4.4. Analysis of Simulation Results

Our results indicate that 5.5% of the entire study area has been occupied as a built-up area in 2010, which will increase to 10.5% by 2050, while for the other land-cover types (except water class), descending rate will observe by 2050 (Table 4). For example, grassland area was seen to decline from 27750 ha (4%) to 15730 ha (29.3%) during 2010-2050. Also, for the other two land-cover classes, similar trends were observed, i.e. from 222550 ha (32.3%) to 217410 ha (32.5%) and 397270 ha (57.6%) to 379510 ha (55%) for forest and farmland, respectively.

Table 4 Absolute quantities for land-cover classes (in hectare) for 2010-2050.

	Built-up land	Forest	Water	Farmland	Grassland
2010	37820	222550	4552	397270	27750
2020	43740	220870	4630	394710	25990
2030	52690	219510	4680	388940	24120
2040	64120	218660	4730	382150	20280
2050	72480	217410	4810	379510	15730

3.5. Discussion and Conclusion

The study reported here investigated land-cover changes over three time periods, 1990–2000, 2000–2010 and 2010-2050 using multi-temporal remote sensing data and GIS. Our results indicate that built-up areas dramatically increased by 90.6% from 1990 to 2010. Overall, 17980 ha have

been changed to built-up areas in this time period. This suggests that the development of urban and rural areas in the past two decades has been a high pace. Araya and Cabral (2010) reported such a high rate of growth in their study area between the years 1990 to 2006. It highlights the fact that an increase in built-up area could be interpreted as a decrease in natural lands (Nature land = Total land area – (Farmland area + Built-up area); Lambin and Meyfroidt, 2011). Table 2 shows that natural areas decreased from 254852 ha in 1990 to 237950 ha in 2010. For example forest lost 850 ha of its cover from 1990 to 2010. Degradation and loss of natural and semi-natural lands has become a profound concern which almost has affected the entire Western and Central Europe (Poschlod et al., 2005; Riecken et al., 2008; CBD, 2010; GBO3, 2010).

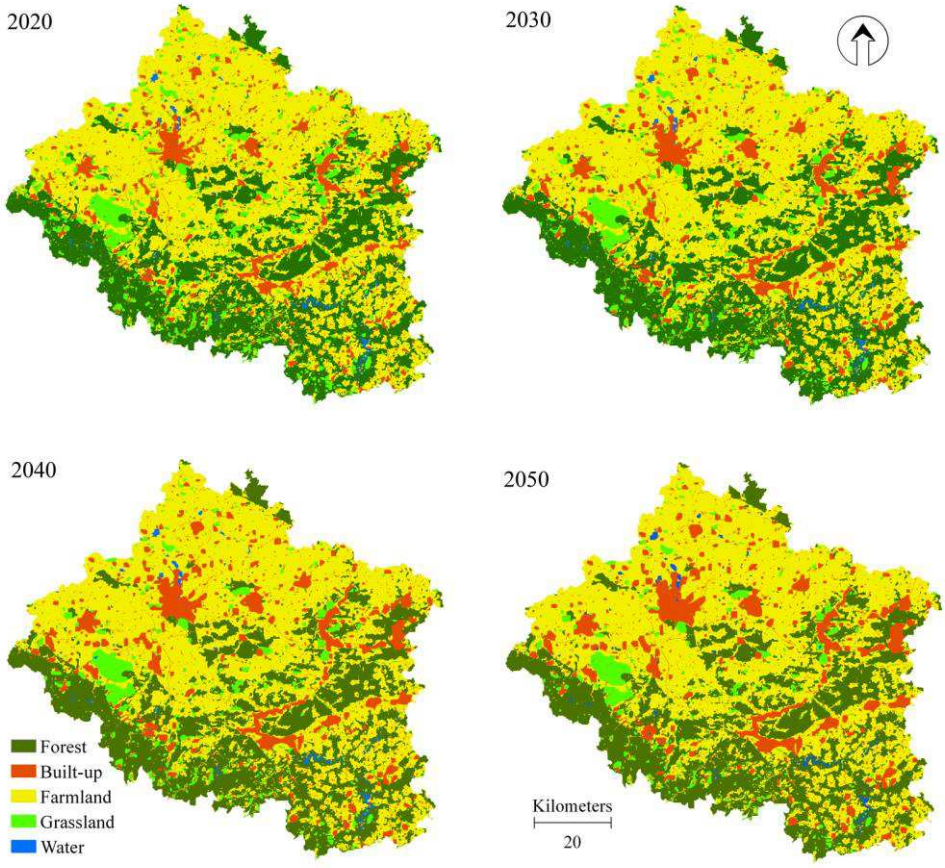


Figure 6 Simulated map of land-cover type from 2020 to 2050.

The results of this study revealed that between the years 1990 to 2010, grasslands have lost a greater percentage of their area compared with forest lands. Table 2 shows that grasslands have lost 17.7% of their land, while forests have lost just 0.38 percent in the same period. These results confirmed the high vulnerability of grasslands in European regions. The grasslands are decreasing in our study area, while previous studies warned that grassland deterioration could have a

significant impact on ecosystem services (i.e. the carbon cycle, regional economy and climate) (Le Houérou, 1996; Angell and McClaran, 2001; Wen et al., 2013). Despite the fact that grasslands are the habitat for more than 50% of vascular plant species in Central Europe (Lind et al., 2009), the European Topic Centre for Biological Diversity (ETC-BD) reports that grasslands are among the endangered habitats in the European regions and only 20 percent of them are in a favorable conditions (EU-COM, 2009; Siehoff et al., 2011).

The U.S. office of Military Government (1946) reported that after the Second World War, timber exports from Germany were particularly heavy, and forest area dramatically decreased consequently. But with change of national and regional policies the rate of deforestation started to decline (FAO, 2011). The effect of this policy change is also visible in the results of this study, so that deforestation in the second decade (2000-2010) was almost half (0.44) of the first decade (2000-1990), while a downward trend has accelerated in grasslands so that in the second decade, this area declined approximately 1.67 times more than the first decade.

Seen from Fig. 7, area change results show that built-up patches will increase in area by the year 2050. The built-up areas are predicted to gain about 34660 ha. Whiles grasslands and forest areas would lose 12020 and 5140 ha, respectively, in the same period. As can be seen in figure 6, the built-up areas have spread towards the suburban side, because farmlands in the suburbs easily change into built-up areas.

In total, results from CA-Markov models indicated a decrease in natural areas. The natural areas are expected to cover 31.5% of our study area in year 2050, which means a 4.7% decrease in comparison with its current distribution. These results clearly showed the high degree of habitat loss and landscape fragmentation in study area which can break habitat connectivity and create a landscape mosaic of suitable, less suitable, and unsuitable habitat patches for species (Wiens et al., 2009). Due to the occurrences of less suitable habitats for establishment of species, the competition between species might increase. It is expected that competition among species significantly reduce their migration speed (Meier et al., 2012; Urban et al., 2012). Increasing competition and declining emigration, can lead to disappearing endangered species. Previous studies (Kinezaki et al., 2010; Meier et al., 2012) indicated that landscape changes will have a strong role in reducing the distribution areas of species in the coming decades, especially at a local scale (Pearson and Dawson, 2003; Engler et al., 2011) like our study area.

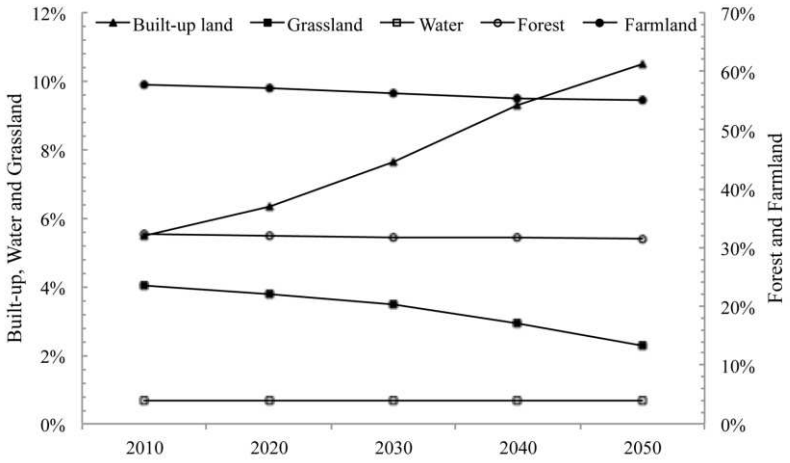


Figure 7 Area changes of land-cover classes from 2010 to 2050.

Although the model used in this study has so well performed the simulation, there are a number of uncertainties in the projections of land-cover classes in the future, which are described as follows: First, it is important to note that some differences are evident between the observed and simulated land-cover maps of 2010. Accuracy of simulated land-cover changes will undoubtedly undergo image classification results (Araya and Cabral, 2010). Although, object-based SVM is a very efficient classification method in handling complex class distributions (Huang et al., 2002; Pal and Mather, 2005), but the classified images are somewhat erroneous. This misclassification can be considered as an uncertainty source in such studies.

Second, inadequate suitability maps for modeling the land-cover classes and the shape of the contiguity filter used in this study have been another source of uncertainty discussed in various studies (Sun et al., 2007; Araya and Cabral, 2010). The suitability maps used in this study have a great influence on the land-cover simulations. This is because they are used as rules during the modeling process. Different suitability maps will lead to different rules that in turn may produce utterly different results. Further research is required to investigate the sensitivity of the predictions to the suitability maps.

Third, although this study confirmed that the procedure used in the analysis of the Markov chain is an effective method to calculate the transition probabilities of land-cover classes; these procedures assumed that transition probabilities do not change over time. In other words, the land-cover changes in the future in this model form up on the basis of land-cover patterns that have been identified in the past. This issue causes uncertainty in the simulation of land-cover changes because the model is not able to assess the new processes occurring on land-cover structures. For example our results show that from 1990 to 2010, the transition probabilities from various lands to built-up areas were extremely high and future land changes simulated based on these transition probabilities. But the evidence suggests that the reality will be something else. After the fall of the Berlin Wall in 1990, the need for modernization and spatial expansion in eastern regions (Former

East Germany) was absolutely essential(Braun et al., 2012). As a result, a new framework for planning the urban development, termed “Critical Reconstruction”, was implemented in these areas (Neill and Schwedler, 2001; Tölle, 2010). The implementation of this policy caused open and empty areas within cities and many lands (i.e. agricultural lands) in the countryside to be converted quickly into industrial and urban areas after the year1990(Loeb, 2006). Therefore, transition probability from other lands to built-up areas was extremely high in eastern regions (such as our study area) during these years. But, at present, the majority of these changes are over and it predicts that urban development (transition probability to built-up area) would significantly reduce in the coming years.

Finally, predicting future landscape changes would be full of uncertainty due to unpredictable events (such as fires and floods), effects of climate change(Vittoz et al., 2009), possible changes in managerial attitudes, and potential uncertainty which coming from simulator models.

The obtained results in this study show that the integration of remote sensing, GIS techniques and land change modeling can expand our understanding of the future trends which landscape will face. In dealing with the problems of the loss of natural land, the simulated maps for future land-cover can prepare better understanding of land-cover changes. Ultimately, these findings can be presented to local authorities (policy makers, urban planners, and natural resource managers) as a comprehensive guide for planning and managing land-covers so that they can have a better understanding of the relationship between landscapes; consequently, this can result in a balance between the interactions of urban development and environmental factors. To determine whether the patterns of projected landscape change are specific to our study area, this technique should be empirically repeated and needs more comparative studies.

Chapter 4

Comparison of statistical models to predict the spatial distribution of plant species in Central Germany

Hamidreza Keshtkar and Winfried Voigt

4.1. Summary

A variety of statistical techniques has been used in species distribution modeling that attempt to predict occurrence of a given species in respect to environmental conditions. We compared the performance of three regression-based models (GAM, GLM, and MARS) with three machine-learning algorithms (RF, ANN, and GBM). Also in this study, three different sets of explanatory variables (climate-only, topography-only and topography-climate combination) for each species were quantified and the interaction of differences among predictor variables with differences in the modeling approaches in determining the accuracy of predictions was tested. Model accuracy was evaluated using the area under Receiver Operating Characteristics curve (AUC) and true skill statistics (TSS). Regression-based approaches showed better performances than machine-learning methods. The results showed that topography-climate variables are the most important variables for mapping potential suitable habitats of target species.

Keywords: Machine-learning algorithm, regression methods, plant distribution, explanatory variable.

4.2. Introduction

Spatial species distributions and the relationship between species and environmental factors have been studied for several years (Guisan and Zimmerman, 2000). Linking environmental variables with physiological tolerance threshold of species makes it possible to model the impacts and consequences of environmental changes on species and ecological systems. To implement such schemes, a proper understanding of the relationship between the species and environment is required. This understanding is usually achieved through theoretical and statistical methods which can relate the environmental variables and emergence of the species. Secondly, we should relate the relationships between species and environment with structural characteristics of the habitat in terms of GIS data.

Ecologists commonly assume that the ranges of current geographic species represent the characteristics of species habitats, which support or limit their presence in a specific location. Accordingly, a range shift is justified by measuring changes in the bioclimatic envelope (a set of

biological and physical conditions suitable for the development and establishment of particular species). In fact, they are under the influence of environmental change processes (Dullinger et al., 2012; Bateman et al., 2013).

Predicting geographical species distributions using statistical models has become essential in several aspects of biogeography, ecology and biology. Species distribution models (SDMs) have been widely used to predict the impact of ecological factors on distribution of species plants (Hijmans and Graham, 2006; Elith et al., 2010; Dirnböck et al., 2011) that rely on the estimation of statistical species–environment relationships (Guisan and Zimmermann, 2000, Guisan and Thuiller, 2005, Zimmermann et al., 2010). Using these models, researchers can predict a probability of existence species in a location where no occurrence data is known.

Various models and methods are used to examine maps which represent the status of organisms in their habitat. The type of required data is one of the most important differences between these models; therefore, they can be divided into two main groups: the first group of these models requires presence/absence data to create statistical tasks or discriminative rules that permit habitat suitability to be graded based on distributions of presence and absence of species (e.g. generalized additive models and artificial neural networks) (Guisan and Zimmerman, 2000); the second group simply requires presence data and they are generally used when the knowledge of the samples of absences is insufficient or inaccessible; these methods include the Ecological Niche Factor Analysis (ENFA), Maximum Entropy Modeling (MaxEnt), BIOCLIM and DOMAIN (Farber and Kadmon, 2003).

Recently, regression and machine learning techniques have been used more than other methods. Most of the regression models used to predict the geographic species distribution presents the highly interpretable and meaningful results. These models are usually restricted to binary data organizations that have a precise and regular sampling strategy; GAM is one such modeling method that has a specifically forceful performance when modeling species presence/absence data (Lehmann et al., 2002). Machine learning techniques include a variety of non-parametric methods able to compute regression or classification tasks using available information. These methods show some benefits with reference to statistical methods: they are capable of handling non-linear relationships among predictors, able to deal with complicated relationships among predictors that can occur in big data sets and capable of managing complicated and noise data (Recknagel, 2001).

Selection of environmental factors to apply as predictors is one of the big challenges in SDMs (Araujo and Guisan, 2006). Selection of predictors with direct effect on species distribution is the best solution for this problem (Austin, 2002, 2007). But most of the time it is not possible to include different types of predictor factors because of the limitation in availability of data, time and/or resource limitations (Bucklin et al., 2015). Since climate is an effective driver of species distribution, one subclass of SDMs comprises only climate (hereafter climate-only) predictors

(Pearson and Dawson, 2003). Climate-only SDMs are essential implements for guiding future conservation efforts (Elith and Leathwick, 2009), although due to the lack of enough information for determination of climate range in species distribution, some scientists have criticized climate-only models (Beale et al., 2008). If non-climatic variables are used along with climatic information, this problem will be solved (Austin and Van Niel, 2011). Although various studies have incorporated climatic and topographic variables in modeling, a few of these studies have examined the climate-only and topography-only models versus combination models.

The objective of this study is to compare the performance of some of the most common methods of presence-absence distribution models using data on the distribution of three plant species in central Germany. Specifically, we compare three regression methods and three machine-learning methods. Considering previous evidence that differences in predictive accuracy among models is depend on the explanatory variables (Bucklin et al., 2015; Oppel et al., 2012; Tsoar et al., 2007; Zhang and Zhang, 2012), we also quantified three different sets of explanatory variables (climate-only, topographic-only and topographic-climate combination) for each species and tested the differences among predictor variables interact with differences in the modeling approaches in determining the accuracy of predictions. Such knowledge could assist to selection of predictors for practical SDM applications, and providing information on which modeling techniques are useful for a group of species.

4.3. Methods

4.3.1. Study areas

The study area is located in central Germany and covers 690000 hectares. Elevation ranges from 114 to 982 m.a.s.l, with higher elevations concentrated in the Grosser Beerberg Mountain located in the Thuringian Forest. The predominant climate is of the continental type with an average annual rainfall of 604 mm, and an average annual air temperature of 8.6 °C (based on monthly recording data of 18 stations, in Free State of Thuringia from 1960–1990). The soil parent material is mainly calcareous. The study area is covered by 5 major phenomena: forest, built-up area, grassland, farmland, and water bodies (lakes, rivers, ponds, and reservoirs).

4.3.2. Species and data preparation

In this study we chose three non-woody species that are native in grasslands of the study area: *Festuca rupicola* (*F. rupicola*), *Achillea millefolium millefolium* (*A. millefolium*), and *Centaurea jacea ssp. Angustifolia* (*C. jacea*). The selected species represent a balanced mix of occurrence frequency, with species being respectively very common (*F. rupicola*), relatively frequent (*A. millefolium*) and relatively rare (*C. jacea*). Species-presence data were obtained from an

unpublished data (Christiane Roscher) which were sampled during the summers of 1994-1996. Out of 201 available vegetation plots, we randomly selected occurrences separated by a minimum distance of 250 metres to minimize spatial autocorrelation effect. The process was separately performed for the observation points of each species. This yielded 144, 102, and 58 occurrences of *F. rupicola*, *A. millefolium*, and *C. jacea*, respectively.

4.3.3. Environmental predictors

We selected a set of quantitative topographic and climatic predictors that cover the main physiological requirements (i.e. water, nutrients, and energy) of species (Körner, 2003; Pearson et al., 2002).In total, 11 predictor variables were calculated at a 25 m spatial resolution: six climatic and five topographic predictors (Table 1). The environmental variables used in this study are described in Zimmermann and Kienast (1999) and Parviainen et al. (2008) and thus we only discussed here in brief. All topographic predictors were derived from the digital elevation model (DEM). The topographic wetness index (TWI) is a popular measure to infer information about the local relative differences in moisture conditions (Grabs et al., 2009). Topographic position index (TPI) or difference from mean elevation (DIFF) is a useful measure that increasingly used to express the exposure of a central point in space compared to the surrounding terrain (Wilson and Gallant, 2000).

Table 1 List of environmental variables which tested for multi-collinearity.

Variables	Unit	Details
Topography		
Slope*	Degrees	Slope inclination
Aspect*	Degrees	The compass direction that a slope faces
Elevation*	m	The elevation of a geographic locations
Topographic wetness index*	m	It quantifies the role of topography for redistributing water in the landscape
Topographic position Index	Unitless	Identification of topographic features at various spatial scales
Climate		
Mean annual temperature	°C	Average of annual temperature
Mean summer temperature*	°C	Average temperature from April to September
Sum annual precipitation	mm	Sum of annual precipitation
Sum summer precipitation*	mm	Sum of precipitation from April to September
Summer solar radiation*	$kJ \times m^{-2} \times day^{-1}$	Sum of monthly average of daily global solar radiation from April to September
Soil moisture index*	mm	Difference between precipitation and potential evapotranspiration

The long-term daily climatic data were derived from the German Meteorological Service for the period 1961-1990. Given that topography strongly affects temperature and precipitation (Fontaine et al., 2002; Hong et al., 2005) a co-kriging technique was executed to interpolate the long-term values over the whole study area using DEM as a co-variable. Soil moisture index was calculated as the monthly difference between precipitation and potential evapotranspiration (see Zimmermann et al., 2007 for details). Solar radiation is a direct ecological factor which can affect the habitat conditions by influencing soil moisture, soil temperature and near surface air temperature (Bennie et al., 2008).

4.3.4. Multi-collinearity analysis

To minimize the effect of multicollinearity, explanatory variables were examined through Pearson correlation coefficient, and predictors that showed strong correlation with other variables (>|0.7|) were left out of the model before running. Finally, 8 environmental predictors were kept for model calibration (marked with '*' in Table 1), out of the original 11 variables.

4.3.5. Calibration of statistical models

Six predictive models (three regression methods and three machine-learning methods) were run and compared to predict species distributions, which are known to provide good predictions: (1) generalized linear models (GLMs; McCullagh and Nelder, 1989; Meynard and Quinn, 2007); (2) generalized additive model (GAM; Hastie and Tibshirani, 1990; Yee and Mitchell, 1991); (3) multivariate adaptive regression splines (MARS; Friedman, 1991; Muñoz and Felicísimo, 2004); (4) generalized boosted models (GBMs, also known as boosted regression trees (BRT); Ridgeway, 1999; Friedman et al., 2000); (5) random forest (RF; Breiman, 2001; Williams et al., 2009); (6) artificial neural networks (ANNs; Ripley, 1996; Segurado and Araújo, 2004). By relating the independent and dependent variables, all the models mentioned here can specify at what probability percentage a pixel will be hosting the target species.

GLMs and GAMs were fitted for each species with a binomial variance and a logit transformation. In both models, the selection of significant variables was done with an Akaike information criterion-based stepwise method (Akaike, 1973) in forward and backward directions. GBMs were calibrated with a maximum number of 2500 trees and internal 3-fold cross-validation procedure to select the optimal numbers of trees to be kept and a value of 7 as maximum depth of variable interactions. MARS models were calibrated using a maximum interaction degree equal to 2. For RF model, we set 500 for the number of trees to grow (ntree), and for the number of input variables (mtry) we used the default value, which is the square root of variables' number. For ANN model, the model optimized the number hidden layer (size) and the weight decay (decay) factor by cross validation based on area under the curve (AUC) of the receiver operating characteristic (ROC). To optimizing, the model test different values for "size" and "decay", respectively, (2, 4, 6, 8) and (0.001, 0.01, 0.05, 0.1), and the one given the best AUC will be selected. All models were run in R software (R Development Core Team, 2009) using the biomod2 package (Thuiller et al., 2013).

For each species, above models were fitted using three different sets of explanatory variables: (1) Topographic variable only (hereafter abbreviated ''Topo''; Table 1); (2) climate only (abbreviated ''Clim''; Table 1); and (3) Topo + Clim variables (abbreviated ''ALL'').

4.3.6. Assessment of model performance

As no independent data were available to evaluate the predictive power of the model, we used a repeated data-splitting procedure. To determining the optimum proportion of training and testing data, we used the formula which provided by Huberty (1994). This heuristic formula is restricted to presence/absence models, and the ratio of testing data should be $[1 + (p - 1)^{½}]^{-1}$, where p is the number of predictors. Accordingly, a model was trained using a random data sample (70% of presences and pseudo-absences data) and evaluated on the remaining 30% with using both the widely reported true skill statistics (TSS, Allouche et al., 2006), and the AUC (Hanley and McNeil, 1982). The TSS evaluation value varies from 0 to 1, where a value of 0 can be interpreted as random predictions and value 1.0 indicates a perfect agreement (Franklin, 2009). The AUC value varies from 0 to 1, where a value below than 0.5 interprets that predictions are no better than random, values of 0.5–0.7 indicate low predictions, 0.7-0.9 indicate useful predictions, and >0.9 indicate excellent predictions (Franklin, 2009; Swets, 1988). This index is calculated as specificity (proportion of correctly predicted presences) + sensitivity (proportion of correctly predicted absences)-1 (Franklin, 2009). The resampling technique (data-splitting) was repeated 25 times for the models and the evaluation metrics averaged. We weighted the presence and pseudo-absence data in the modeling procedure so that both gave prevalence of 0.5. This equal prevalence prevents the model bias towards over-prediction of either presences or pseudo-absences data (Isabelle et al., 2014).

For the final calibration of each model the whole of data used to carry out spatial projections. The predictive maps were developed for the target species after calibration of the models. While continuous predictions need conversion to a binary map (i.e. a species is either predicted present or absent), we used threshold classification according to the sensitivity-specificity sum maximization approach (Liu et al., 2005). To test whether the probability of occurrence values for each species predicted by different predictive models and three different sets of explanatory variables differed from each other, we used a non-parametric Wilcoxon's signed-rank test (Phillips et al., 2009; Randin et al., 2006).

4.4. Results

4.4.1. Multi-collinearity among variables

All of the variables that were used for model calibration had correlation values <0.7. Table 1 shows variables that kept for modeling. All in all, one topography and two climate variables were deleted from study. Topography and climate variables showed almost no correlation with each other.

4.4.2. Performance of distribution models

The selection of the best model is obtained by considering AUC and TSS measures. The mean AUC values of the six models ranged from 0.64 (ANN) to 0.94 (GAM and MARS) and from 0.42 (ANN) to 0.80 (GAM) for TSS index. All three species were classified useful and excellent with all models, except ANN model that classified poorly (Fig. 1). Tables 2 present the results of the different statistical techniques constructed with different subset of environmental variable which applied to the *F. rupicola*, *A. millefolium* and *C. jacea* data sets. The statistical tests by Wilcoxon signed rank test depicted that there were no statistically difference between performance of regression models (GLM, MARS and GAM; P>0.05, Table 3).

Table 2 Mean evaluation values of TSS (true skill statistics) and AUC (the area under the receiver-operated characteristic curve) of six modeling techniques for predicting the distribution of three plant species based on three set of explanatory variables. TOPO= topography-only variables, CLIM= climate-only variables, All= topo-climate variables. See Fig. 1 for the techniques abbreviations.

Models	TOPO		CLIM		ALL	
	TSS	AUC	TSS	AUC	TSS	AUC
C. jacea						
RF	0.62	0.78	0.61	0.76	0.66	0.84
ANN	0.42	0.65	0.47	0.67	0.55	0.73
GBM	0.61	0.74	0.68	0.84	0.67	0.85
GAM	0.65	0.78	0.70	0.87	0.71	0.90
GLM	0.67	0.81	0.71	0.84	0.73	0.92
MARS	0.63	0.76	0.67	0.79	0.70	0.87
A. millefolium						
RF	0.63	0.77	0.65	0.80	0.69	0.84
ANN	0.57	0.69	0.57	0.72	0.61	0.74
GBM	0.62	0.75	0.68	0.80	0.72	0.89
GAM	0.71	0.78	0.73	0.85	0.74	0.93
GLM	0.66	0.76	0.71	0.84	0.69	0.83
MARS	0.68	0.84	0.69	0.79	0.71	0.85
F. rupicola						
RF	0.65	0.79	0.70	0.81	0.71	0.83
ANN	0.49	0.64	0.53	0.68	0.56	0.73
GBM	0.68	0.82	0.69	0.87	0.72	0.86
GAM	0.72	0.87	0.74	0.93	0.80	0.94
GLM	0.73	0.83	0.76	0.89	0.75	0.89
MARS	0.71	0.89	0.76	0.92	0.79	0.94

With the *F. rupicola* data set the best performance was achieved by the MARS with AUC=0.92, and TSS = 0.76, although GAM results were very similar. For *C. jacea,* the best performance was obtained by the GLM with AUC=0.92 and TSS= 0.73. The best projection was carried out using the GAM with AUC=0.93 and TSS= 0.74 for *A. millefolium.* Our results illustrated that the GAM method provided significantly more robust predictions than all the machine-learning algorithms (Wilcoxon signed rank test; P< 0.05; Table 3 and Fig. 1).

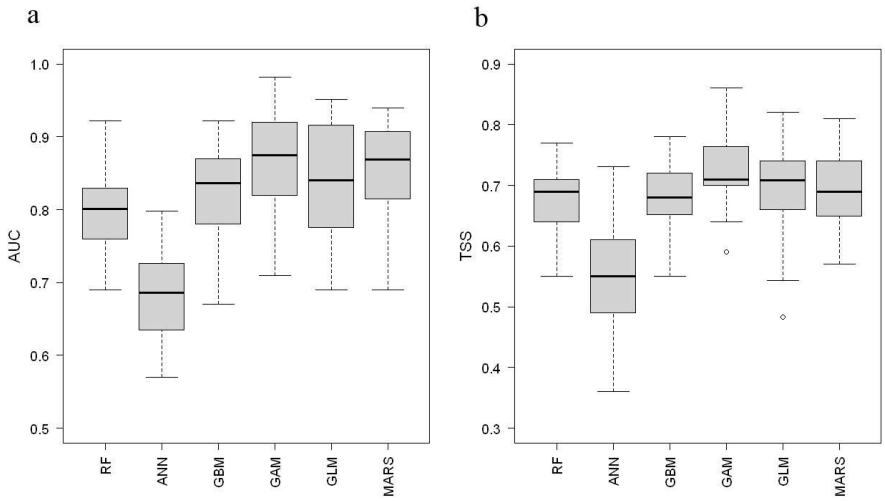


Fig. 1 Comparison of AUC (a) and TSS (b) model evaluations between modeling techniques based on "All" explanatory variables. Each box-plot is built form three values (i.e. plant species). The boxes extend from the data's 1st to 3rd quartiles, box boundaries show the interquartile range and the horizontal bars in the box represent the median. RF= random forest, ANN = artificial neural networks, GBM = boosted regression trees, GAM = generalized additive models, GLM = generalized linear models, MARS = multivariate adaptive regression splines.

4.4.3. Comparison of models fitted with different explanatory

All of the models exhibited very good correctness under all three sets of the environmental variables (Table 2). Modeling with climate-only variables was significantly better than modeling with environment-only variables (Wilcoxon signed rank test; $P < 0.001$). Also, modeling using a set of climatic and environmental factors had a superior predictive ability than two other variable sets (Wilcoxon signed rank test from climate-only models; $P < 0.01$). According theses results, all analyses hereinafter consider only the models calibrated with both the climatic and topographic variables (ALL).

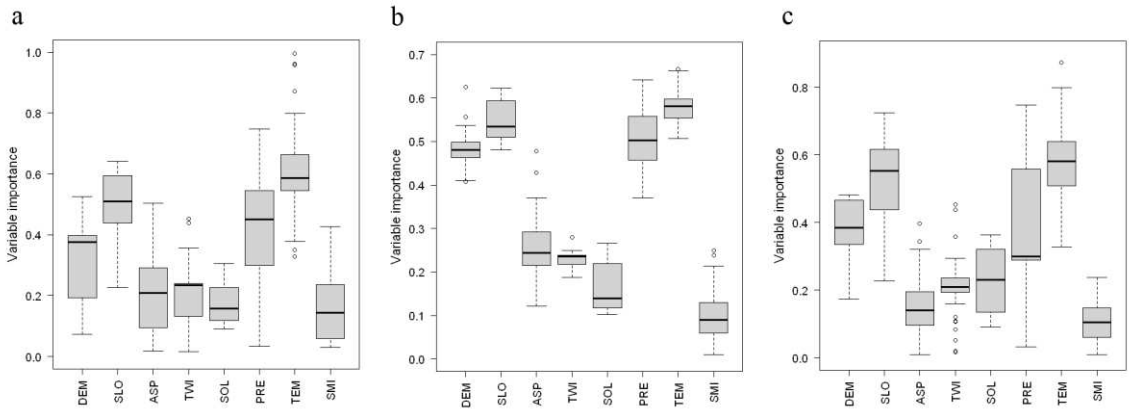


Fig. 2 Importance of each predictor used in calibrated models for three species; *F. rupicola* (a), *C. jacea* (b) and *A. millefolium* (c). A high value (like Temp) indicates an important influence of the predictor in the model. DEM= digital elevation model; SLO= slope in degrees; ASP= aspect in degrees; TWI= topographic wetness index; SOL= sum of solar radiation for the growing season (April–September); PRE= sum precipitation over the growing season; TEM= mean temperature for the growing season; SMI= soil moisture index.

The most important predictor for modeling the distribution of *F. rupicola* was the mean summer temperature (Fig. 2), followed by the sum summer precipitation, the slope and the DEM. The results show that the mean summer temperature followed by the sum summer precipitation, the slope are most important predictors for modeling the distribution of *C. jacea* and *A. millefolium*. Soil moisture index (SMI) was the least important predictor for these two plant species. The response curves of the best model in this study (i.e. GAM) in *F. rupicola* showed that the optimal value for the sum summer precipitation was between 300-400 mm (Fig. 3). Also, the suitability of habitat increased with decreasing elevation (< 430 m) and increasing slope (>5 degrees) and mean summer temperature (>14 °C).

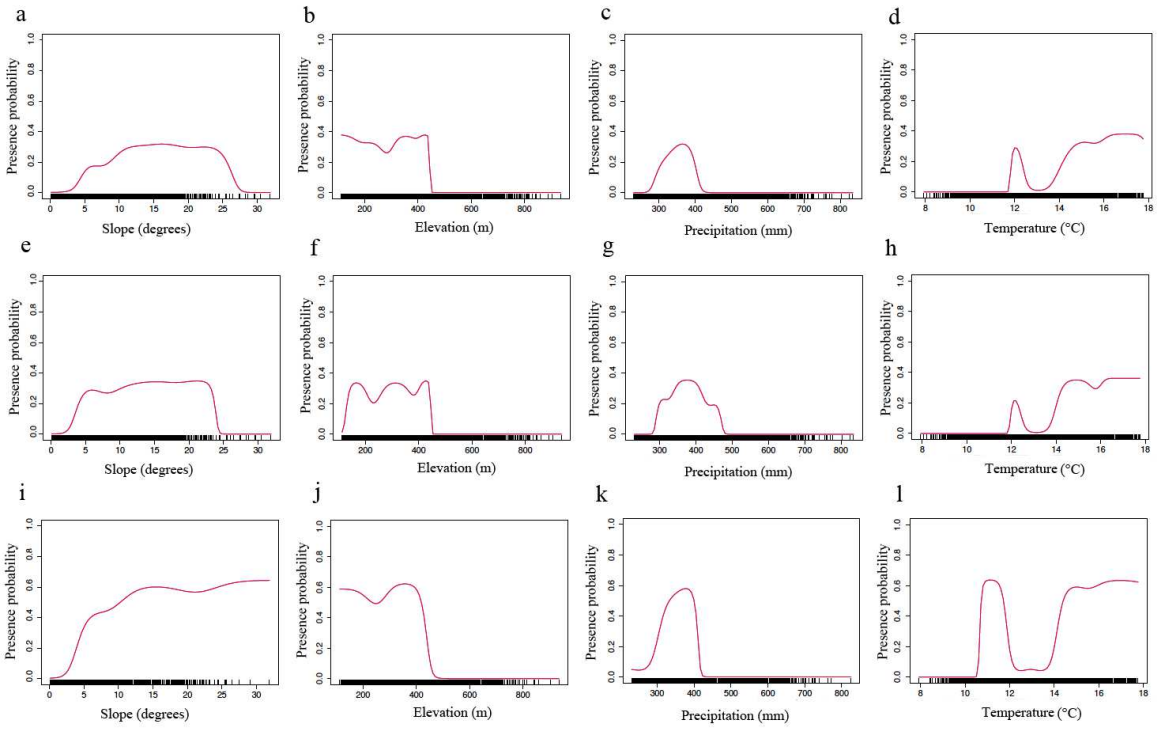


Fig. 3 Response curves for *A. millefolium* (a-d), *C. jacea* (e-h) and *F. rupicola* (i-l) based on GAM for the four most important predictors. Frequency distribution of the each predictor in study area is shown by black bands on x-axis.

4.4.4. Comparison of models across species

When investigating the distribution projections, all species obtained high evaluation scores (except ANN model for *F. rupicola* and *C. jacea*), with TSS values between 0.61 and 0.80 (Table 2). Such a value of TSS means that, on average, with the probability of 80-90% the model was properly able to predict the presence and absence of the species. *F. rupicola* is the species that obtained the highest evaluation score. *A. millefolium* is at average and *C. jacea* reached slightly lower evaluation scores than the other species.

Table 3 Statistical differences in the predictive performance of six different models for three plant species. Statistical tests of the differences among the predictive accuracies of different methods based on AUC scores were tested by Wilcoxon signed rank test (*P*-values). See Fig. 1 for the techniques abbreviations.

Models	RF	ANN	GBM	GAM	GLM	MARS
RF	-	< 0.001	0.008	< 0.001	0.001	< 0.001
ANN		-	< 0.001	< 0.001	< 0.001	< 0.001
GBM			-	< 0.001	0.011	0.007
GAM				-	0.016	0.091
GLM					-	0.253
MARS						-

4.5. Discussion

SDMs are valuable tools for the evaluation and protection of regions degrading and losing their biodiversity due to various factors (Bustamante and Seoane, 2004; Rodriguez et al., 2007). Robertson (2003) suggested that the prediction provided by each model may present different conceptions of the potential distribution and biology of the target species, though investigators should have a perfect perception of the restrictions and ambiguities embedded in species distribution modeling to produce suitable and precise models (e.g. Elith et al., 2002; Loiselle et al., 2003; Barry and Elith, 2006; Gibson et al., 2007). This study predicted the spatial location of three individual plant species at a 25 m spatial resolution and investigated the impact that different modeling techniques and different sets of explanatory variables (climate-only, topography-only and topo-climatic) have on model performance.

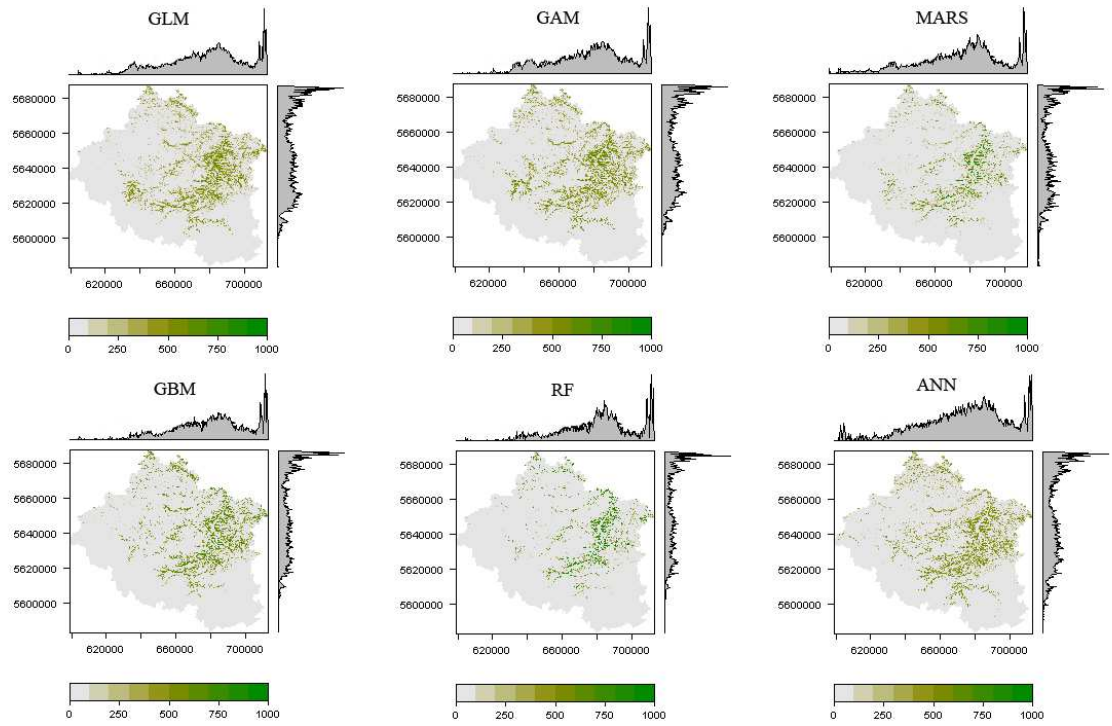


Fig. 4 Predicted environmental suitability maps for *F. rupicola* using six predictor algorithms. See Fig. 1 for the algorithms abbreviations.

Comparing the predictive power of the different SDMs, it was found that overall, the GAM model showed the best results (Fig. 1). This is consistent with the results of experiments performed by Leathwick et al. (2005) and Heikkinen et al. (2012) in the field of species distribution modeling. The performance of GBM and RF was acceptable but poorer than that of regression approaches when applied to predict suitable habitats. The ANN modeling technique received lower evaluation scores. The MARS and GLM models, very similar in performance to the GAM model, can then be considered as substitute mapping methods (Fig. 1 and Table 2). This result is consistent with the results obtained by Leathwick et al. (2005). They compared GAM and MARS techniques, including a number of new techniques applied to freshwater fish, and indicated that the two methods present similar outcomes. Additionally, previous studies also revealed that MARS is comparable to other regression techniques (i.e. GAM and GAM) in terms of function and capability (Leathwick et al., 2006; Guisan et al., 2007). In spite of comparable predictive operations, the quality of predicted distributions can differ owing to different emphasis on and modeled relationships with environmental variables (Elith and Graham, 2009; Ready et al., 2010). In the current study, for example, the MARS and GAM models had similar predictive performance (Fig. 1 and Table 3), but would have selected very different suitable habitat areas for*F. rupicola* and *C. jacea* (Fig. 4 and 5).

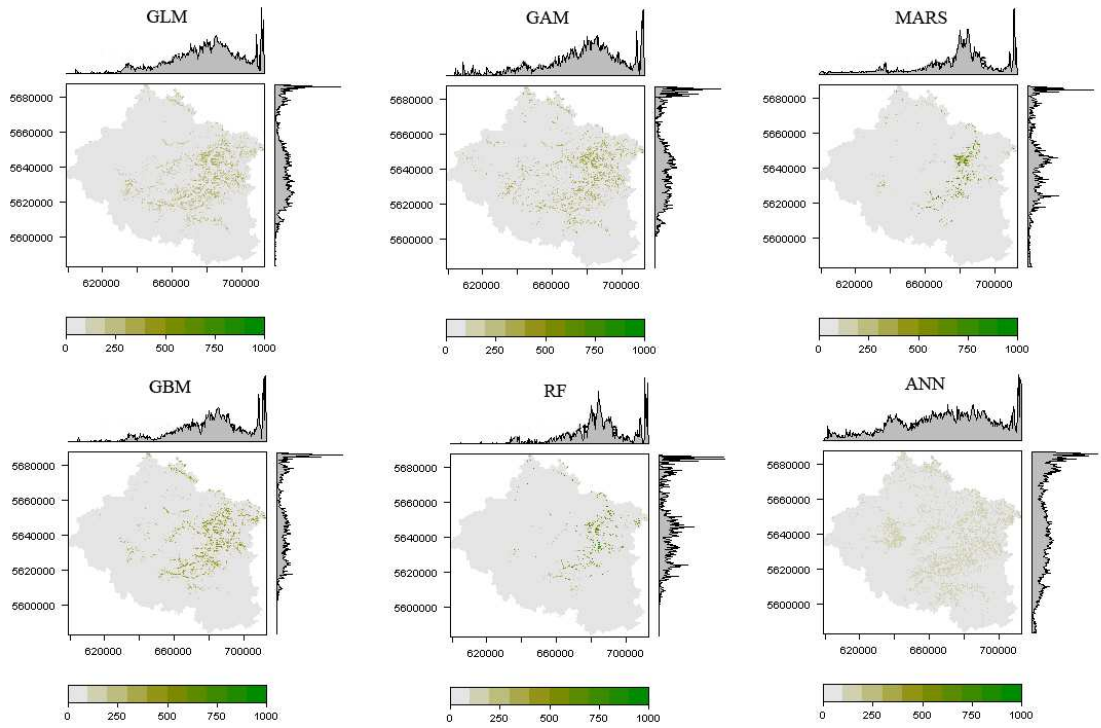


Fig. 5 Predicted environmental suitability maps for *C. jacea* using six predictor algorithms. See Fig. 1 for the algorithms abbreviations.

The explanatory variables included in this study were chosen so that they covered a broad spectrum of the possible ecological determinants of the distributions of the modeled plant species. Climate variables, particularly temperature and precipitation in the growing seasons, and topographic factors such as elevation and slope, appeared as significant determinants across all modeling techniques (Fig. 2). Such variables represent primary environmental regimes related to the physiological requirements of plants (Körner, 2003; Pearson et al., 2002).

Results of the current study show that the ALL scenarios (topo-climate variables) are the most important variables for predicting potential suitable habitats of target species. On the contrary, models based solely on TOPO variables showed lower evaluation scores. Likewise, simultaneous incorporation of topographic and climatic variables increased the models' prediction power significantly (Wilcoxon signed rank test comparing ALL and CLIM models; $P < 0.01$). For example, under the ALL scenario, mean TSS for *F. rupicola* showed 2.7% and 6% higher performance than CLIM and TOPO scenarios, respectively. Some earlier studies confirmed that topo-climate explanatory variables (ALL scenario) strongly predict habitat distribution of species (Engler et al., 2009; Gonzalez et al., 2010; Kissling et al., 2010).

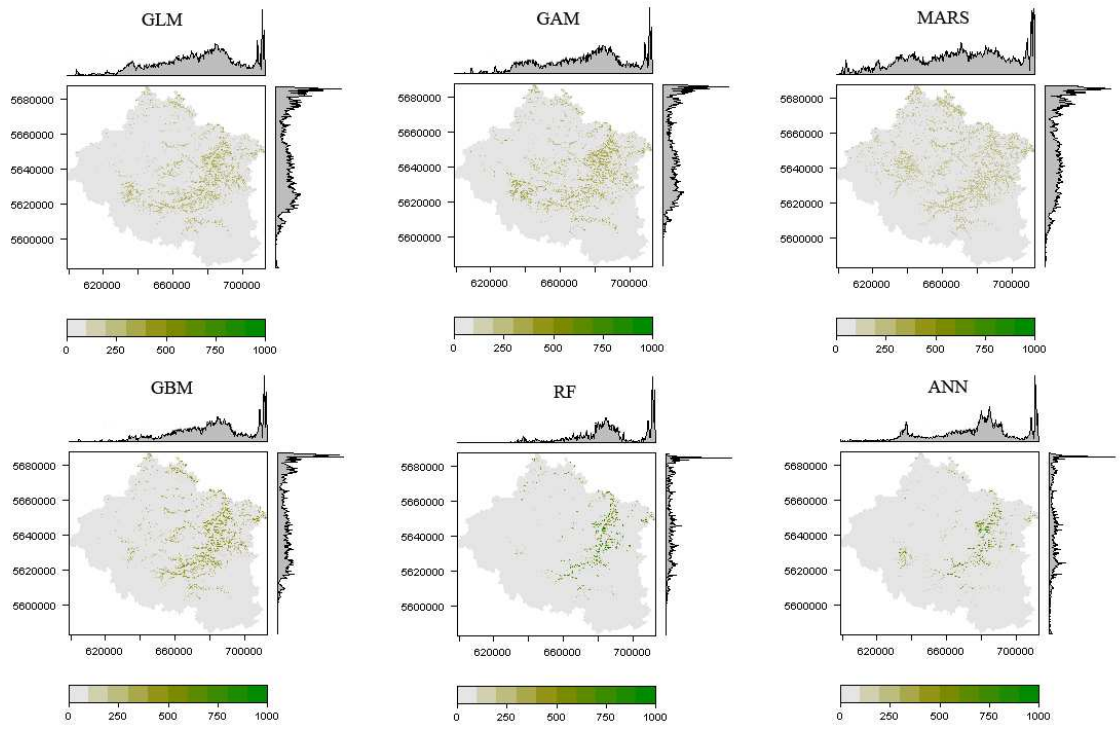


Fig. 6 Predicted environmental suitability maps for *A. millefolium* using six predictor algorithms. See Fig. 1 for the algorithms abbreviations.

Predictive species modeling can provide good information about the habitat of species and their method of interacting with their environment. Response curves are one of the important components in species distribution studies, because they show the tolerance range for

environmental changes by species. The most important predictor in the models was temperature with a higher suitability for mean summer temperatures >14°C for all plant species, corresponding approximately to elevations <430 m a.s.l. This result could not translate into a preference for lowland regions, because higher elevations are covered by forestland in the study area while target plant species occurred only in open lands.

Although most models used in this study performed the predictions well, there are a number of uncertainties in the projections of plant species distribution. First, all the models are sensitive to the qualities and quantities of the predictor and response variables. Although most of the factors used in this study demonstrated good predictive abilities for the projections of the species suitable habitats, these species were likely to be affected by other factors, the impacts of which were neglected in this study. However, to reduce uncertainties in this research, a combination of environmental and climate variables were utilized to display a better performance compared with those using only climate variables (Barbet-Massin et al., 2012).

Second, the validity of adequate information on species used to run niche models is contingent upon the potential biases in the availability of adequate information on the presence or absence of species. Several studies have shown that if absence data is collected along with presence data, niche models would be strengthened and the results could be closer to reality (Wiens et al., 2009). In the current study, the true absence data of the species was not available. Collecting presence and absence data of species during the growing season and performing replications within several successive years, like what was done in this study, can significantly enhance the quality of the observed data (Wiens et al., 2009).

All in all, the accuracy of the fitted models obtained from the GAM modeling technique in this study suggest that, at least at a regional level, useful and informative suitable habitat maps for species can be produced, and they can be used for a range of research and management applications.

Chapter 5

Potential Impacts of Climate and Landscape Fragmentation Changes on Plant Distributions: Coupling Multi-Temporal Satellite Imagery with GIS-based Cellular Automata Model

Hamidreza Keshtkar and Winfried Voigt

This manuscript has been published in 2016, Ecological informatics, **32**, 145-155.

5.1. Summary

Climate change and landscape fragmentation are considered to be the main treats to biodiversity. In this study, probable alteration of future species distribution was tested based on the association of landscape fragmentation and climate change scenarios compared to the classical approach that assumed an unchanged landscape. Also, projected range shifts including realistic dispersal scenarios were compared with classical models, in which no or full dispersal has been supposed.

A GIS-based cellular automata model, MigClim, was implemented to projection of future distribution over the 21st century for three plant species in a study area of the central Germany. For each species, simulations were run for four dispersal scenarios (full dispersal, no dispersal, realistic dispersal, and realistic dispersal with long-distance dispersal events), two landscape fragmentation (static and dynamic change) and two climate change (RCP4.5 and RCP8.5) scenarios. In this research, temporal satellite data were utilized to simulate landscape changes by the use of a hybrid (CA-Markov) model for the years 2020, 2040, 2060 and 2080.

A significant difference appears to be between the simulations of realistic dispersal limitations and those considering full or no dispersal for projected future distributions. Although simulations accounting for dispersal limitations produced, for our study area, results that were closer to no dispersal than to full dispersal. Additionally, our results revealed that change in landscape fragmentation is more effective than the climate change impacts on species distributions in this study.

Keywords: climate change, dispersal distance, landscape fragmentation, Cellular automata, Markov chain, plant species distribution.

5.2. Introduction

Due to human activities during the past century, we have been witnessing increased concentrations of greenhouse gases, especially carbon dioxide and methane; this issue has resulted in global warming. Climatic change may create new environments and forces the organisms to react. The

species generally go after suitable climatic conditions through displacement, and expanding or shrinking their dispersal range (Collingham and Huntley, 2000; Tingley et al., 2009; Huntley et al., 2010), or even may lead to extinction of native species (Thomas et al., 2004). These responses could have a significant impact on the future of biodiversity (Lambers, 2015), distributions (Parmesan and Yohe, 2003); phenological patterns (Fitter and Fitter, 2002) and carbon storage (Purves and Pacala, 2008). Therefore prediction of species range shifts under climate change and other physical processes is a crucial challenge for the management of natural resources.

Species distribution models (SDMs) have been widely used to predict the impact of future climate change on distribution of species plants (Hijmans and Graham, 2006; Elith et al., 2010; Dirnböck et al., 2011) that rely on the estimation of statistical species–environment relationships (Guisan and Zimmermann, 2000; Guisan and Thuiller, 2005; Zimmermann et al., 2010). Previous studies generally assumed that species are in equilibrium with their environment and hence, did not investigate adaptation and tolerance threshold of species towards climatic changes (Engler et al., 2009). But in the past decade, some researchers tried to predict the effects of climate change on distribution of plant species using dispersal models based on statistical relationships between species and environment (McConkey et al., 2012; Bateman et al., 2013). These models can predict suitable sites for establishment of the species based on future climate changes. However, it should be considered that the spread of species relies on the dispersal ability of species, presence of effective factors as well as barriers to dispersal. Neglecting dispersal process is usually seen as a defect and limitation in projects of SDM climatic change(Franklin, 2010). For predicting distribution of plant species, these models usually combine the current distribution of species with future climate scenarios in which dispersal biology is not regarded and full- or no-dispersal scenarios are used. In these two scenarios, ability of the species for migrating at different distances as well as the limiting factors of dispersal are ignored. Combining dispersal patterns of species with predictor models can clearly enhance the performance of these models in response to global changes (Meier et al., 2010).

Changes in landscape fragmentation (LF) can affect distribution of plant species through changing the quality and quantity of habitat suitability, as well as, via increasing, decreasing, or eliminating dispersal barriers. Natural barriers such as forests and mountains, and artificial barriers such as cities and villages can prevent or restrict plant species dispersal (Bateman et al., 2013). Although in some studies climate change scenarios have been used to examine future plant distributions (Elith et al., 2010; Dirnböck et al., 2011), little attention was paid to the synergetic effects of landscape fragmentation and climate change. This method of application assumes a negligible effect of LF change on future species distribution compared to climate change impact, which is expected to occur more rapidly and severely than LF change (Barbet-Massin et al., 2012), or alternatively, assumed that dynamic landscape variables are unavailable (Stanton et al., 2012).

To improve the prediction of migration potential of vegetation under future climate change scenarios, Engler et al. (2012) developed a dynamic model -MigClim- to couple dispersal processes and LF scenarios with SDMs in projecting distributional changes under climate change. The mentioned model has a cellular-automata structure which presents the study field as a regular network of stations and simulates the changes in species distribution. These changes are imposed by climatic time series and are under the influence of adjacent cells. MigClim lets that individual species respond to changes in their habitats, but the model ignores the dynamics LF changes over time.

To estimate the possible effects of these processes on dynamics of target plants, we applied a cellular automata-Markov chain model (CA-Markov) as an alternative to simulate landscape change scenarios during the 21st century. Markov-Chain model is one of the most widely used methods for quantifying the probability of landscape change from state A to state B (e.g. forest to built-up area) in discrete time stages. These probabilities then enter into the cellular automata (CA) model to predict spatial changes over a specific time period (Mitsova et al., 2011; Yang et al., 2012). CA-Markov model is based on the initial distribution and transition matrix; it assumes that the drivers, which have created the current situation for the region land-cover, will continue to operate as before in the future (Guan et al., 2011). In many studies, the combination of remote sensing and GIS are effectively used in CA-Markov model (Peterson et al., 2009). Although a few studies have examined the roles of climate change and habitat fragmentation on species distribution (Opdam and Wascher, 2004; Leimu et al, 2010; Oliver et al, 2015), to the best of our knowledge, this is the first study to combine the CA-Markov model with the models of species distribution to investigate species migration in the future.

The target species, spermatophyte plants, have logically dispersal abilities; thus, this research aimed to establish the realistic estimates of the abilities of these species to disperse within the study area. Here, the potential impacts of LF and climate change scenarios on the distributions of target plants were estimated and compared. To evaluate the possible effects of global changes on the range dynamics of the plants, MigClim model was employed to combine the projections of geographical habitat shifts under future climate changes with the mechanistic simulations of seed dispersal. The major objective of this paper was to integrate MigClim, SDM, CA, and Markov chain models so as to assess the effects of future LF and climate change scenarios on the geographic distributions of three open-land plant species. To this goal, the following procedures were followed: 1) Future LF changes were simulated through a temporal mapping of land-cover changes in 2000 and 2010 within central Germany; 2) Different simulations were run for the target plants under two climate change scenarios during the period of 1995–2100 besides investigating landscape changes and various dispersal limitations; and 3) the spatial patterns of the results were compared with the distributions under full and no dispersal scenarios and static land-cover.

Overall, the future potential distribution map of each species was simulated under two climate change scenarios (RCP4.5 and RCP8.5), four dispersal scenarios (no dispersal, full dispersal, and realistic dispersal distances with and without long-distance dispersal events), and two landscape patterns (static and dynamic) to investigate whether the plant species would likely gain or lose suitable environmental spaces.

5.3. Materials and methods

5.3.1. Study areas

The study area is located in Central Germany and covers 6900 km^2 (Fig. 1). Elevation ranges from 114 to 982 m.a.s.l, with higher elevations concentrated in the Grosser Beerberg Mountain located in the Thuringian Forest. The predominant climate is of the continental type with an average annual rainfall of 604 mm, and an average annual air temperature of 8.6 °C (based on monthly recording data of 18 stations, in Free State of Thuringia from 1960–1990). The soil parent material is mainly calcareous. The land-cover is a heterogeneous mixture of forest, grasslands, farmland, water bodies and built-up areas.

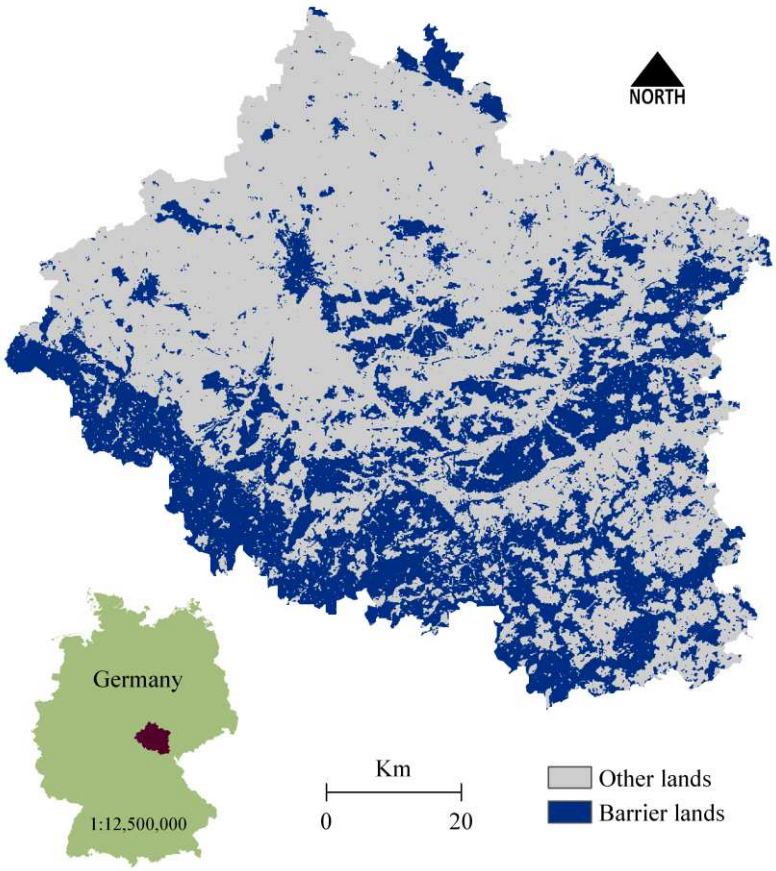


Figure 1 Location of study area and extracted barrier areas for year 1995. Other land contains farm land, grassland and water bodies, and barrier land contains forest and built-up areas.

5.3.2. Species and data preparation

In this study we chose three non-woody species that are native in grasslands of the study area: *Festuca rupicola* (*F. rupicola*), *Achillea millefolium millefolium* (*A. millefolium*), and *Centaurea jacea ssp. Angustifolia* (*C. jacea*). The selected species represent a balanced mix of occurrence frequency, with species being respectively very common (*F. rupicola*), relatively frequent (*A. millefolium*) and relatively rare (*C. jacea*). They are seed-bearing and short-lived perennial plants which grow in open-land area. Species-presence data were obtained from an unpublished data (Christiane Roscher) which were sampled during the summers of 1994-1996. Out of 201 available vegetation plots, we randomly selected occurrences separated by a minimum distance of 250 metres to minimize spatial autocorrelation effect. The process was separately performed for the observation points of each species. This yielded 144, 102, and 58 occurrences of *F. rupicola*, *A. millefolium*, and *C. jacea*, respectively.

5.3.3. Modeling framework

The modeling framework consisted of four steps (Fig. 2). For each species: 1) Land-cover change projection was based on spatially and temporally simulations of explicit landscape changes into the future using a hybrid model (CA-Markov) since the future landscape impacts on habitat extent and pattern were of our particular concerns; 2) Climate change scenarios were based on two socio-economic scenarios (RCP4.5 and RCP8.5) used with the SDM to project suitable habitat distributions under future climate scenarios; 3) Species distribution model was based on modeling habitat suitability maps for the target species using species location records and environmental predictor maps (including current climate variables); 4) Future distributions of plants were simulated over the 21st century using a GIS-based model, MigClim, that is capable of implementing various parameters such as dispersal distance, reproductive potential increase over time, and LF. The initial habitat patch map and time series of habitat suitability maps which constructed in the previous steps are the core data for this model.

5.3.3.1. Modeling landscape changes

In this section, we intended to monitor the trend of landscape change in the study area in order to simulate future changes in forests and built-up areas, which are considered as barriers and unsuitable locations to the observation of plant dispersal. To do so, remote sensing, GIS techniques, Markov chain and cellular-automata models were integrated to predict the forthcoming changes of landscape, comprising the following stages: 1) land-cover mapping of 1990, 2000 and 2010 using the classification of satellite images derived from object-based support vector machine (SVM) classification method, 2) computation of transition area matrix derived from a Markov process, indicating the number of pixels to be expected to change each land-cover class to another

class over a specified time interval (1990-2000 and 2000-2010); 3) getting transition suitability images by a multi-criteria evaluation (MCE) model based on the driving factors (These suitability images imply the suitability of each cell for a particular land-cover); and 4) land-cover change simulation using CA-Markov module. The input data were calibrated for simulating the selected landscape transitions by 2020, 2040, 2060 and 2080. In this study, although beside forest and built-up area, water bodies and roads are also considered as unsuitable locations to plant dispersal, however, we supposed that water bodies and roads have no change in future. Land-cover mapping (first step) was run in ENVI ZOOM (Version 4.8), and other steps were calculated in IDRISI-Selva software (https://clarklabs.org/). The detailed methodology steps are described in chapter 3.

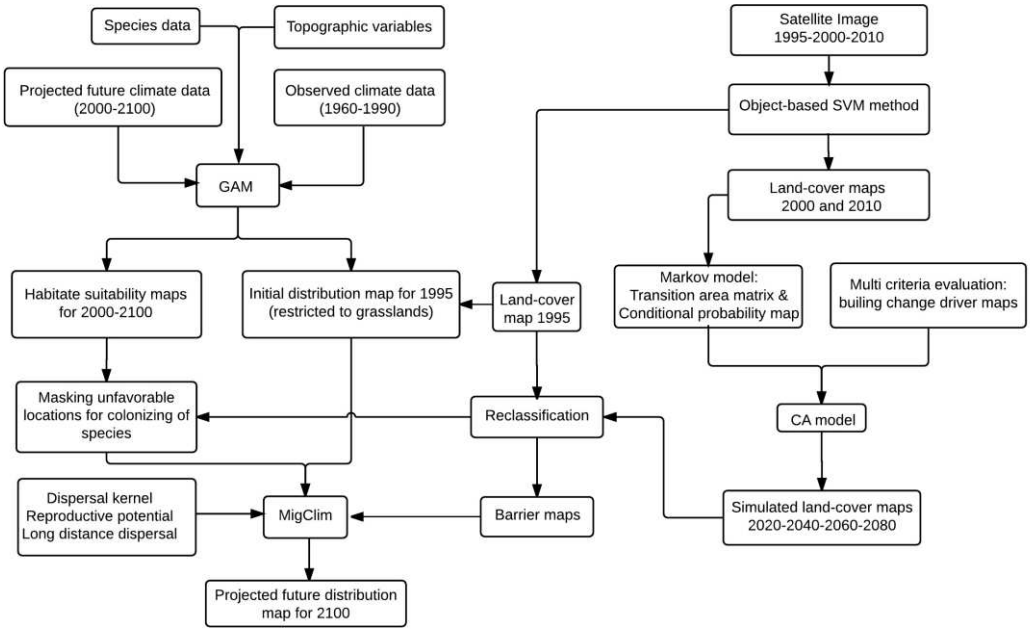


Figure 2 Conceptual framework showing the different databases used in the study.

5.3.3.2. Habitat suitability modeling

To model the species distributions, we used a generalized additive model (GAM; Hastie and Tibshirani, 1990); which is known to provide good predictions. This model relate the response to the independent variable in order to derive the probability of a pixel to host a given target species. GAM was fitted using a binomial error distribution and a logit link function. A bidirectional (forward-backward) stepwise procedure was used to select the most significant predictors, based on the Akaike information criterion (Akaike, 1998). This model was run in R software (R Development Core Team, 2009) using the biomod2 package (Thuiller et al., 2013).

As no independent data were available to evaluate the predictive power of the model, we used a repeated data-splitting procedure. To determining the optimum proportion of training and testing

data, we used the formula which provided by Huberty (1994). This heuristic formula is restricted to presence/absence models, and the ratio of testing data should be $[1 + (p - 1)^{½}]^{-1}$, where p is the number of predictors. Accordingly, a model was trained using a random data sample (70% of presences and pseudo-absences data) and evaluated on the remaining 30% with using both the widely reported true skill statistics (TSS, Allouche et al., 2006), and the area under the receiver operating characteristic curve (AUC; Hanley and McNeil, 1982). For the final calibration of each model the whole of data used to carry out spatial projections. While continuous predictions need conversion to a binary map, we used threshold classification according to the minimizing the absolute difference between sensitivity and specificity (Liu et al., 2005).

To minimize the effect of multicollinearity, explanatory variables were examined through Pearson correlation coefficient, and predictors that showed strong correlation with other variables (>|0.6|) were left out of the model before running. The model was run with the final selected topo-climatic predictors (summer precipitation (April to September), summer temperature, summer solar radiation, digital elevation model, slope and topographic wetness index) variables at a 25 m spatial resolution (Table 1), that cover the main physiological requirements (i.e. water, nutrients, and energy) of species (Pearson et al., 2002; Körner, 2003). All topographic predictors were derived from the digital elevation model (DEM). The topographic wetness index (TWI) is a popular measure to infer information about the local relative differences in moisture conditions (Grabs et al., 2009). Solar radiation is a direct ecological factor which can affect the habitat conditions by influencing soil moisture, soil temperature and near surface air temperature (Bennie et al., 2008). The long-term monthly means climatic data (i.e. average temperature (°C), and sum of precipitation (mm)) were derived from the German Meteorological Service for the period 1961-1990. Given that topography strongly affects temperature and precipitation (Fontaine et al., 2002; Hong et al., 2005) a co-kriging technique was executed to interpolate the long-term values over the whole study area using DEM as a co-variable.

Table 1 Topo-climatic variables used to model species distributions.

Variables	Unit	Details
Slope	Degrees	Slope inclination
Elevation	m	The elevation of a geographic locations
Topographic wetness index	m	It quantifies the role of topography for redistributing water in the landscape
Summer solar radiation	$kJ \times m^{-2} \times day^{-1}$	Sum of monthly average of daily global solar radiation from April to September
Mean summer temperature	°C	Average temperature from April to September
Sum summer precipitation	mm	Sum of precipitation from April to September

In this study, the current and future distributions were modeled using both dynamic and static variables (climate and topography) to improve the modeling accuracy and to predict a more realistic response to the global changes. To avoid over-predictions about the species occurrence, the future habitat suitability maps were filtered by masking forests, built-up areas, and water bodies (rivers, lakes and reservoirs) which were considered as unsuitable locations for colonizing

of species. Likewise, in order to precise estimation of the initial distribution of species –as source area– and to avoid misleading predictions of the plant colonization in the future, the initial distribution maps were restricted to only grassland areas. The grassland areas were extracted from land-cover map made of satellite images taken in 1995.

5.3.3.3. Climate change scenarios

In this research, for all the species, the current climatic conditions were used to map the initial distributions (i.e. in 1995) within their potential suitable habitats. Since changes in the climatic factors can make the current habitats climatically unsuitable, the possible shifts in habitat suitability of the species were created every five years from 2000 to 2100. To create a time series of dynamic suitable habitat, we used downscaled and calibrated climate data based on a multi-model ensemble mean for two socio-economic scenarios (RCP4.5 "medium emissions", and RCP8.5 "high emissions") from the Coupled Model Intercomparison Project Phase 5 (CMIP5) (Taylor et al., 2011), and 17 global climate models (See Appendix B). The future mean temperatures and precipitations were simulated in the web-based stochastic weather generator "MarkSim" based upon the method described by Jones and Thornton (2013). Each unique location in an area can be examined by MarkSim, proving to be thus often better than other climate predictors in terms of simulating climate variances in temperate climates. Moreover, it is able to closely model real climatic variances within special climate clusters by utilizing stochastic downscaling and weather typing (Boeckmann and Joyner, 2014).

5.3.3.4. MigClim model

MigClim is a GIS-based (hybrid) model that focuses on simulation the temporal and spatial dynamics of species. This model is a cellular automaton that linked to habitat suitability from SDMs, to generate range predictions that consider climate change scenarios and dispersal limitations.

This model requires the following data and parameters: 1) initial distribution map of plants; 2) habitat suitability maps to assess plant distribution shifts under climate change scenarios; 3) LF maps (i.e. pixels acting as barriers to dispersal and those showing permanent unfavorable locations); and 4) species' dispersal parameters such as dispersal kernel (the probability of colonization for a target cell based on its distance from a productive pixel), propagule production potential (probability of an occupied cell to produce seed as an incubation time function) and long-distance dispersal (Table 2). MigClim allows for comparing a regular dispersal (short-distance dispersal) and long-distance dispersal (LDD) using full-dispersal (FD) and no-dispersal (ND) scenarios. According to Engler and Guisan (2009), most of the seeds (e.g. 99%) are distributed within short distances with expected patterns and only a small number of them are affected by

LDD events. Unlike short-distance dispersal (SDD), LDD events can distribute seeds over the other side of a geographical barrier. In this study, the dispersal behavior of each species considered was specified according to the seed dispersal types defined in Vittoz and Engler (2007) and the required time to convert a newly-colonized cell into a cell source was set based on the information in Engler et al. (2009) and expert knowledge. Also, since a distance increase from a seed producing the source cell causes a decrease in the colonization probabilities of the target cells, a negative exponential seed dispersal kernel was presented to model regular seed dispersal. More detailed information in Engler and Guisan (2009).

Table 2 Parameters used in the MigClim model for each plant species. SDD= short-distance dispersal, LDD= long-distance dispersal.

	F. rupicola	*A. millefolium*	*C. jacea*	Reference
Maturity age	2 year	2 year	1 year	Aeschimann and Heitz (1996)
Max SDD	15 m	1500 m	1500 m	Vittoz and Engler (2007)
Max LDD	1000 m	5000 m	5000 m	Vittoz and Engler (2007)
LDD frequency	0.01	0.01	0.01	Engler et al. (2009)
Dispersal event freq. *	2 year	1 year	1 year	
Barriers	Forest and built-up areas			
Unsuitable habitats	Roads and water bodies (lake, river and reservoirs)			

* Dispersal event frequency indicates the time between two successive dispersal events.

To run MigClim over the study area, three interrelated sub-models were employed to simulate the overall dynamics: Firstly, an initial distribution map was produced for each plant using correlative species distribution models. This map displays the cells occupied by the target species at the beginning of the simulation. Secondly, habitat suitability modeling was established to indicate which cells are suitable for the species in future years. To this end, an initial distribution map was developed to refine projections of the species distributions under future climate change scenarios in each environmental change step (i.e. every five years). For each species, the future potential distribution maps were simulated under two IPCC (AR5) climate change scenarios (RCP4.5 and RCP8.5) on the basis of four dispersal scenarios (ND, FD, SDD, and LDD). In addition, each of the mentioned scenarios were implemented with two LF patterns (static landscape (SLS) and dynamic landscape (DLS)), except for the no-dispersal scenario (totally 14 scenarios). In this research, dispersal of the target species except for *F. rupicola* was simulated for each year. Since *F. rupicola* has a short dispersal distance (15 m), it cannot be modeled every year. Therefore, the dispersal distance of *F. rupicola* was multiplied by two and dispersal was simulated every two year. LDD event was generated with a probability of 0.01 in a random direction. All the simulations were performed over a period of 105 years (1996–2100) and repeated 25 times. In this research, forest and built-up areas were considered as the barrier pixels preventing dispersal, except for LDD events which are not affected by the barriers. These barrier areas were updated every 20 years (i.e. 2020, 2040, 2060, and 2080).

5.4. Results

5.4.1. Model performance and contribution of predictors

The evaluation values obtained for GAM model showed that all species under the present and future climatic conditions attained AUC values >0.9 and TSS Values >0.77. These modeling results were considered to be of an excellent standard (Swets, 1988). Among all predictors, mean summer temperature was the variable most frequently selected by the stepwise procedure for the current projecting of *F. rupicola* and *C. jacea* species (Fig. 3). While the most important predictor for *A. millefolium* was the sum of summer precipitation. TWI and solar radiation were the least important predictors across all species.

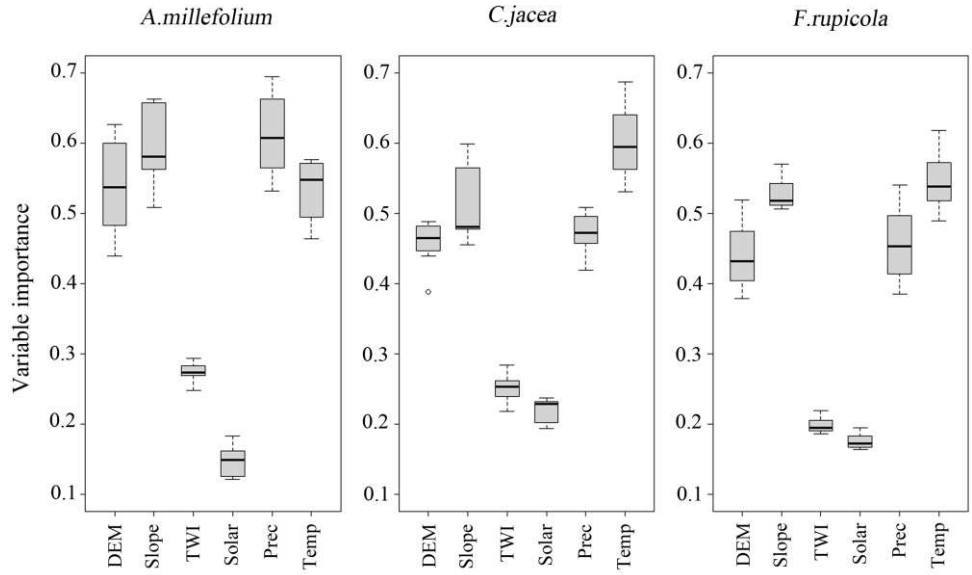


Figure 3 Importance of each predictor used in GAM model for three species: a high value (like Temp) indicates an important influence of the predictor in the model. DEM= digital elevation model; Slope= slope in degrees; TWI= topographic wetness index; Solar= sum of solar radiation for the growing season (April–September); Prec= sum precipitation over the growing season; Temp= mean temperature for the growing season.

5.4.2. Temporal landscape mapping

In order to validate the performance of the CA–Markov model, the Kappa coefficient (Kraemer, 1982) was computed. This statistical measure is based on a comparison of the actual changes and predicted changes. The actual land-cover map of the study area of 2010, which was obtained through remote sensing techniques, was used to assess the simulated results. The Kappa value was 0.87, which verifies the accuracy of this model. Our results indicate that 3.1% of the entire study area has been occupied as a built-up area in 1995, which will increase to 14.5% by 2080, while forest areas will show a decrease from 32.4% to 31.2% within the same period (Table 3). Generally, multi-temporal change analyses of barrier areas permit growth quantification over time.

Therefore, our spatial comparisons of the land-cover maps were able to represent enhancement of the barrier areas from 35.6% in 1995 to 45.7% in 2080. The extracted and simulated barrier maps for 1995 to 2080 are displayed in Figure 1 and Appendix D.

Table 3 Absolute quantities for land-cover classes (in km^2) for 1995-2080. Other land contains farm land, grassland, and water bodies.

	1995	2020	2040	2060	2080
Forest	2233.8	2208.7	2186.6	2168.2	2149.6
Built-up area	211.6	437.4	641.2	825.7	1003.3
Other land	4454.0	4253.3	4071.6	3905.5	3746.5

5.4.3. Projections of regional climate change

The mean summer temperatures and sum of summer precipitations were calculated in each grid cell using 17 global climate models under two emission scenarios for every five years. The estimated changes in the mentioned factors under each emission scenario are shown in Appendix C. The two different estimates of emissions are indicative of consistent warming trends by 2100 (i.e. 2.4°C under RCP4.5 and 5.3°C under RCP8.5). Under the high emissions scenario, precipitation sum during the growing season demonstrated a consistent decline from 2000 to 2100 (56.4 mm), while precipitation under the medium emissions scenario illustrated an oscillating motion and thus a total reduction of 17.3 mm was predicted.

5.4.4. Projected distributions under different scenarios

Comparison of the projected distributions of all the three species with their initial distributions for the year 2100 indicates that their distributions are always augmented under all the dispersal scenarios, except for no-dispersal scenario (Table 4).

Table 4 depicts that among the three plants, *F. rupicola* has the widest potentially suitable habitat (PSH, i.e. the area a species could occupy under full-dispersal scenario; Engler and Guisan, 2009) in a way to occupy 26% and 38% of the study area more than *A. millefolium* and *C. jacea* on the average, respectively. Furthermore, comparing the initial distributions of the study plant species with their projected distributions for 2100, *A. millefolium* and *C. jacea* were observed to lose distribution under ND scenario. *A. millefolium* and *C. jacea* were predicted to lose 30% and nearly 32% of their initial occupied areas, respectively, whereas *F. rupicola* distribution surface was seen to be constant over time. The comparisons revealed the differences between the dispersal scenarios (ND, FD, SDD, and LDD) in 2100 increment with growing initial PSH areas (except for the difference between SDD/LDD scenarios with ND scenario for *A. millefolium*) (Table 4).

The difference between the projections obtained using FD scenario demonstrated a significant expansion in the projected distributions of 2100. For example, as the occupied areas of *F. rupicola*

species were predicted to be 1.2%, 11.9%, and 13% under ND, SDD, and LDD scenarios on the average for 2100, respectively, its occupied area was found to be nearly 59.7% under FD scenario (Table 4). Also, it is worth mentioning that the results obtained from SDD and LDD scenarios indicated all the three plants will occupy less than 50% of PSH by 2100. These results actually portray an overestimated assessment under FD scenario.

Table 4 Estimated mean percentage of area for three species by 2100 as output by two climate change scenarios (RCP4.5, RCP8.5), four dispersal scenarios (ND, FD, SDD, LDD), and two landscape fragmentation change scenarios (SLS, DLS). Initial shows the percentage of occupied areas by each species in 1995. ND= no dispersal, FD= full dispersal, SDD= short distance dispersal, LDD= long-distance dispersal, SLS= static landscape, DLS= dynamic landscape.

	SLS			DLS		
	A. millefolium	*C. jacea*	*F. rupicola*	*A. millefolium*	*C. jacea*	*F. rupicola*
Initial	0.82	0.57	1.21	0.82	0.57	1.21
SDD						
RCP4.5	14.96	10.17	12.82	12.93	8.88	10.91
RCP8.5	15.35	10.23	12.93	13.30	8.94	11.03
LDD						
RCP4.5	15.52	10.72	13.97	13.42	9.30	11.91
RCP8.5	15.94	10.73	14.11	13.86	9.36	12.07
ND						
RCP4.5	0.57	0.39	1.21	0.57	0.39	1.21
RCP8.5	0.58	0.39	1.21	0.58	0.39	1.21
FD						
RCP4.5	34.63	23.09	64.40	29.20	19.56	54.97
RCP8.5	37.61	22.42	64.40	31.73	19.02	54.97

As it can be inferred from the results, the response of a given climate change scenario is not always similar to that of each dispersal distance scenario. For instance, the difference between SDD and LDD dispersal projections of 2100 is slightly larger under the high climate change scenario than the medium climate change scenario (except for *C. jacea* species under the SDD scenario) (Fig. 4). Moreover, the results were indicative of the existence of larger differences between ND and SDD/LDD under RCP8.5 climate change scenario, while the differences between FD and SDD/LDD were larger under RCP4.5 scenario (with the exception of *A. millefolium* showing a greater difference under the high-emission scenario in both cases). Except for *F. rupicola*, the projected future distributions for the plants were significantly different between two emission scenarios when using FD modeling approach ($p<0.001$). However, no significant difference was observed between the projections made under two climate change scenarios when considering LDD and SLS scenarios for *C. jacea* ($p>0.05$), while a significant difference existed between all the other projections made under LDD scenario ($p<0.001$) (Fig. 4).

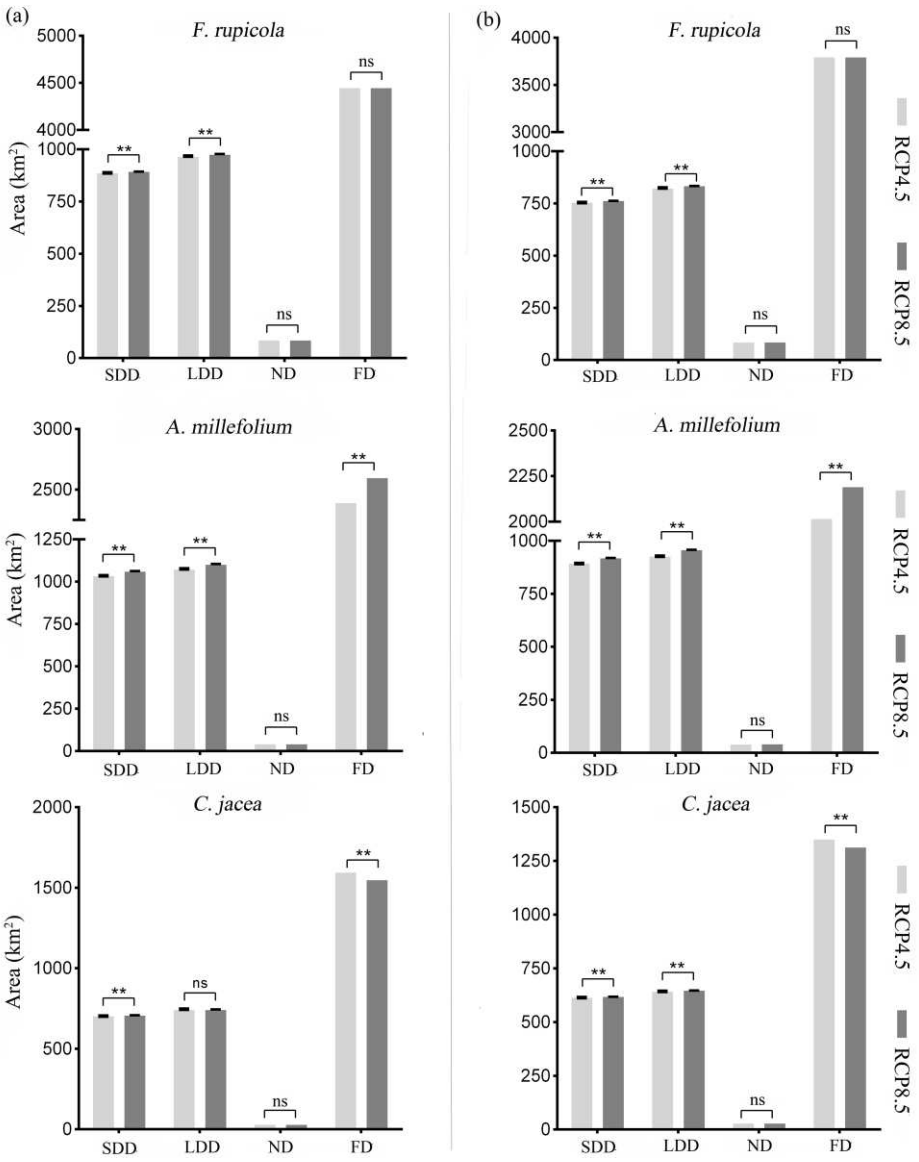


Figure 4 Colonised areas for three species in 2100 under the two climate changes (RCP4.5, RCP8.5) and dispersal scenarios (ND, FD, SDD, LDD). (a) Projected distributions in which species migrate under SLS scenario. (b) Projected distributions in which species migrate under DLS scenario. Data are expressed as means ± SEM (n=25). ns= non-significant ($p>0.05$), **= significant ($p<0.01$). Abbreviations as in table 4.

Introducing LF change scenarios in our simulations allowed for determining the descending rate of the projected distributions (Table 4) ranging from no (under ND scenario) up to 9.4% (for *F. rupicola* under FD scenario). Under SDD/LDD scenarios, *F. rupicola* projected area was seen to decline from 13.5% under SLS scenario to 11.5% under DLS scenario on the average for 2100. Also, for the other two plants under the mentioned scenarios, similar trends were observed, i.e. from 15.5% to 13.4% and 10.5% to 9.1% for *A. millefolium* and *C. jacea*, respectively (Table 4). These results demonstrated that the difference obtained under FD scenario for *F. rupicola* species is higher than those of the other species and the differences of the projected areas under SDD/LDD scenarios are greater for *A. millefolium* species when comparing landscape scenarios.

5.5. Discussion

5.5.1. Effects of dispersal limitations

In this study we produced 14 consensus forecasts for each grass species over the 21st century, one for each possible combination of two IPCC emission scenarios (RCP4.5 and RCP8.5), one for each dispersal scenarios, i.e. no-dispersal (species will persist only in their current suitable habitat inside grasslands), full-dispersal (species could disperse to all new potential suitable areas inside the study area), short distance distribution, long distance distribution and of two LF scenario, i.e. static and dynamic landscape. Table 4 shows *A. millefolium* and *C. jacea* species will lose a considerable part of their areas under ND scenarios by 2100. Thuiller et al. (2006), McKenney et al. (2007) and Engler et al. (2009) also obtained similar results. The reason should be sought in the hypothesis of this scenario which states that due to various constraints, the species cannot migrate to other habitats. Thus, they will have no chance of migration under this scenario and their range reductions in the coming decades will not be a surprise. Perhaps, that is why this scenario is referred to as the worst-case scenario for plant species distributions although the results of this scenario can be more realistic for the species that have lost their abilities to disperse under the influences of intrinsic and extrinsic factors or those having low speed and distribution powers (especially, when using coarse grid scales) (Bateman et al., 2013).

Besides, the results obtained by the limited projections of future plant distributions via realistic dispersal restrictions showed to be generally closer to ND than to FD scenario when compared with SDD/LDD scenarios (Fig. 4, 5). This result is inconsistent with the results obtained by Engler et al. (2009). The cause for this conflict can be found in the configuration differences of the two regions since their study was conducted in a mountainous area. Since the slope of climate changes is steeper in mountainous regions, when a prediction is made based on climate changes, a decrease in displacement of suitable habitats is the result. This is because shorter distances are probably required to follow suitable weather conditions as compared to flat areas (Engler et al., 2009). Therefore, it must have been unexpected to find much smaller differences between SDD/LDD and FD scenarios in mountainous areas when compared to our results.

This conclusion confirms that the projection of species distribution under FD scenario is generally overestimated (Midgley et al., 2006; Engler et al., 2009), especially in non-mountainous areas. For example, in this study, the difference in projected areas between FD and SDD/LDD scenarios were about 52%, 11%, and 19% on the average for *F. rupicola*, *C. jacea*, and *A. millefolium*, respectively, when simulations are made for high emission scenario (Fig. 4). Although FD scenario might be favorable for some species, such as invasive species and birds (Thuiller et al., 2005; Barbet-Massin et al., 2012), unrealistic results are mainly resulted for plant species (Bateman et al., 2013). That is why Midgley et al. (2006) have suggested that this scenario should be removed from plant species studies.

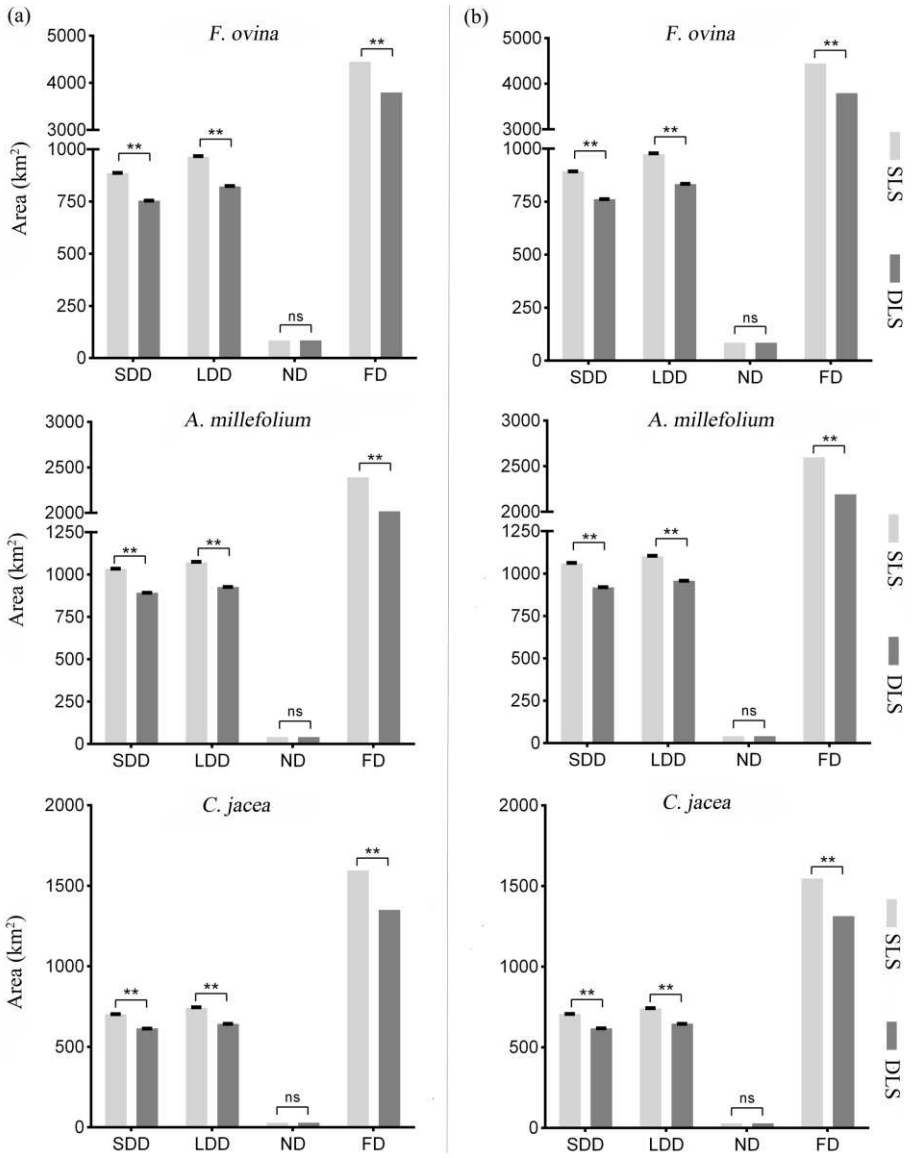


Figure 5 Area representation of plant distributions by 2100 under the four dispersal scenarios (ND, FD, SDD, LDD) and two landscape change events (SLS, DLS). (a) Projected distributions in which species migrate under RCP4.5 climate change scenario. (b) Projected distributions in which species migrate under RCP8.5 climate change scenario. Data are expressed as means ± SEM (n=25). ns = non-significant ($p>0.05$), **= significant ($p<0.01$). Abbreviations as in table 4.

Despite the fact that the use of SDD/LDD scenarios require information that is not generally available for many plants, these scenarios can significantly make closer to reality the projection of species distribution in the future even with guesswork and less accurate information at hand (Engler and Guisan, 2009; Bateman et al., 2013). For example, under the warmest climate change scenarios, the predicted area for *F. rupicola* under FD scenario was 53 times as large as its area under ND scenario for 2100, while SDD/LDD scenarios were only 12 times higher than ND scenario. When the difference was examined for the other two species of *A. millefolium* and *C. jacea*, reductions from 65 to 27 and 57 to 26 were observed, respectively (Fig. 4).

In our research, addition of LDD events to our simulations resulted significantly different from SDD simulations (as found by Imbach et al., 2013). Results from Engler et al. (2009) are in

contradiction with ours, they found no significant differences on species distributions between SDD and LDD dispersal rates. This is probably due to the smaller size of their study area and the nature of investigated species pool. The larger a study area is, the better an LDD event can show its effectiveness. In addition, the existence of unsuitable areas and barriers can affect its performance (Engler et al., 2009). Average increases of the predicted areas under LDD scenario were approximately 10.2%, 5.2%, and 4.1% for *F. rupicola*, *C. jacea*, and *A. millefolium* species, respectively, depending on the LF and climate change scenarios (Fig. 4 and 5).

5.5.2. Effects of climate change scenarios

The effect of climate change on plants has been considered as a permanent factor (Pearson and Dawson, 2003; Kelly and Goulden, 2008) influencing plant distribution in many studies (Pearson et al., 2002; Dirnböck et al., 2003; Loarie et al., 2008; Martin et al., 2013). In this investigation, it was shown that the predicted distribution areas for all the three species under RCP8.5 scenario will largely increase in the coming decades, which is probably due to the fact that at low altitudes, plants are more tolerant to heat and drought than in highlands. In addition, the plants in these areas are less sensitive to climate changes since they still have the chance to migrate to higher elevations and thus less likely face extinction (Engler et al., 2011) unless some factors limit their migrations to higher areas. Vittoz et al. (2009) also revealed that rises in temperature have been still suitable for lowland species since the past decades.

Our results demonstrated that the studied species have no greatly different responses to different climate change scenarios (Fig. 4, table 3) which might be due to little restrictions they have had for distribution and have thus been less affected by varied climate scenarios. Some studies have suggested that climate changes have greatest impacts on the plants in mountainous areas where they confront greater distribution restrictions due to the occurrences of less suitable habitats and increased competitions (Engler et al., 2009; Vittoz et al., 2009). Yet, the results obtained in this study cannot indicate the overall climate change impacts on the species distributions in our study areas because in addition to having direct effects on the habitat conditions, these changes (anthropogenic climatic changes)can exert an influence on the species distribution speeds and abilities by affecting some other factors such as phenological variability (Hampe, 2011; Nathan et al., 2011), seed dispersal by wind (Bullock et al., 2012), species interactions such as competition (Gilman et al., 2010), and the frequencies of disturbance factors such as fires (Hampe, 2011). Therefore, climate change certainly has a decisive role in plant growths and migrations (Kawakami et al., 2009; Bullock et al., 2012).

5.5.3. Effects of landscape change scenarios

Our results indicated that built-up patches will increase in area by the year 2080 (Table 3). The built-up areas are predicted to gain about 792 km^2. Whiles forest areas are forecasted to lose 84 km^2 in the same period. In total, results from CA-Markov model indicated an increase in barriers (forest and built-up areas). The barrier areas is expected to cover 45.7% of our study area in year 2080, which means a 28.4% increase in comparison with its current distribution. These results clearly showed the high degrees of habitat loss and LF in study area which can break habitat connectivity and create a landscape mosaic of suitable, less suitable, and unsuitable habitat patches (Wiens et al., 2009).

As expected, an increase in the barrier areas led to a decrease in suitable and colonised habitat areas for the studied species (Fig. 5, 6 and Appendix E). For instance, the PSH areas of *F. rupicola* are projected to lose 651 Km^2, which means a 9.4% reduction in comparison with its distribution under SLS scenario, which, in comparison to the two other plants, is the largest amount of reduction observed (5.7% and 3.5% declines for *A. millefolium* and *C. jacea*, respectively).

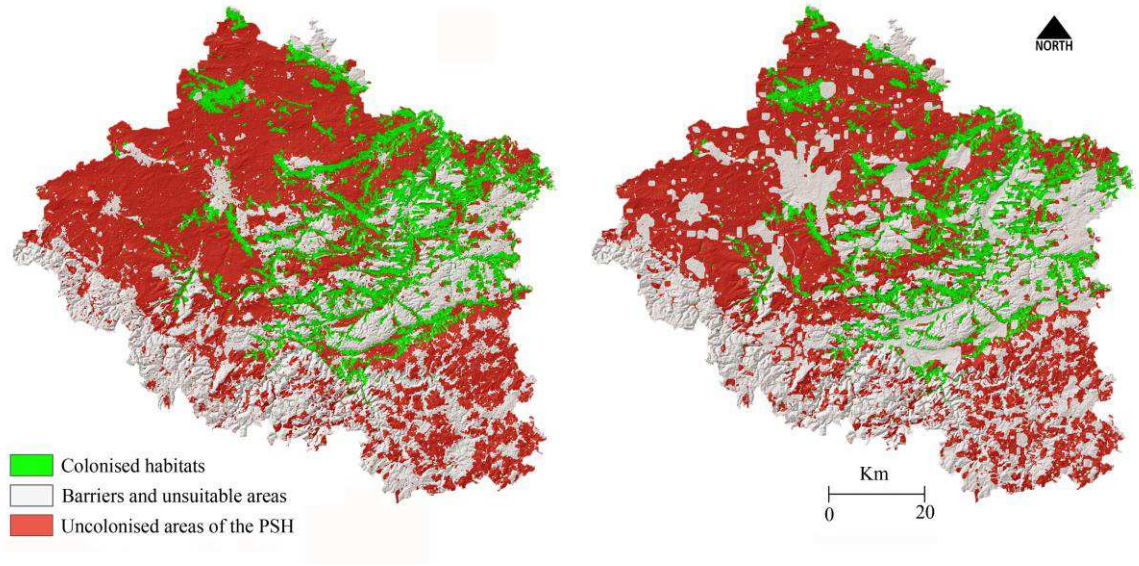


Figure 6 Predicted distribution maps of *F. rupicola* (green areas) for year 2100 with high-greenhouse-gas emission and long-dispersal (LDD) events under static (left) and dynamic (right) landscape scenarios. PSH= potentially suitable habitat.

A species needs to be able to distribute throughout suitable habitats to occupy parts of an environmental niche space (Pearson and Dawson, 2003; Jeschke and Strayer, 2008), while landscape patterns actually manage the existing corridors for its migration (Hannah, 2011). For this reason, some studies have investigated the effects of reducing habitat areas and increasing isolations on the distances and distribution qualities of seeds within and between the fragments of ecosystems (Rodríguez-Cabal et al., 2007; Kirika et al., 2008; Cordeiro et al., 2009; Lehouck et al., 2009; Uriarte et al., 2011) and proven the roles of barriers and fragmentation in the distribution

rates of species (Levey et al., 2008; Dewhirst and Lutscher, 2009; Brown et al., 2012). In consistency with the previous studies (Kinezaki et al., 2010; Meier et al., 2012), our projections indicated that landscape changes will have a strong role in reducing the distribution areas of plants in the coming decades, especially in local studies (Pearson and Dawson, 2003; Engler et al., 2011) like our study area. If human land-cover and fragmentation significantly limit the relationship between suitable habitats, application of ND scenario would seem to be appropriate in such circumstances (Lu et al., 2012).

5.5.4. Sources of uncertainty

Although the model used in this study has so well performed the predictions, there are a number of uncertainties in the projections of plant species distribution in the future, which are described as follows: First, lack of information on how the trend of climate change and its consequences will be followed in the coming years is considered as a source of uncertainty (Vittoz et al., 2009). Climate envelope techniques have been extensively used to evaluate climate change effects on the spatial patterns of biodiversity, while many of their uncertainty aspects have been debated (Heikkinen et al., 2006; Jeschke and Strayer, 2008). Vittoz et al. (2009) stated that the effects of climate change on landscape change should not be ignored. Some studies (e.g., Reichler and Kim, 2008; Pierce et al., 2009) showed that the multi-model ensemble average is the best determinant for predicting the global and regional climate. For this purpose, in this study the ensemble mean of the 17 climate models were calculated to reduce the uncertainties in future climate change projections.

Second, evolution of ecosystems and biotic interactions will undoubtedly undergo future climate changes (Davis et al., 1998; Brooker, 2006; Memmott et al., 2007; Suttle et al., 2007). Nevertheless, due to the lack of information on how our study species adapt or response to environmental changes, their evolutions were assumed to occur under constant conditions. Also, because insufficient information was available to evaluate the effects of biotic interactions on the present and future distributions of species, we considered the target species as independent entities. Therefore, competition as one of the most important biological factors was removed in this study. It is expected that competitions among species significantly reduce their migration speeds (Meier et al., 2012; Urban et al., 2012).

Third, predicting future landscape changes would be full of uncertainty due to unpredictable events (such as fires and floods), socioeconomic changes (Radeloff et al., 2012), possible changes in managerial attitudes, and potential uncertainty which coming from simulator models. In addition to the mentioned sources of uncertainty, we assumed that the categories of landscape will remain unchanged for 2080-2100, which possibly leading to overestimate in the forecasting.

Fourth, lack of knowledge about species distribution quantities and qualities has been another cause of uncertainty discussed in various studies (Engler et al., 2009; Bateman et al., 2013). The

question of how far the seeds can exactly migrate from their native plants or whether they will have the ability or chance of germination and production of new individuals will remain as a controversial issue. Also in our study the role and effect of aboveground and belowground herbivores were disregarded. The previous studies have shown that grazing reduces seed production in plants, especially in grasslands and can thus affect plant community composition, competition among species, and their ability to disperse (Collins and Gordon, 1985; Marage et al., 2006).

5.6. Conclusion

This study revealed how miscellaneous plant species distribution can be under the influence of different dispersal scenarios. The impact of different dispersal scenarios on distribution projections by 2100 clearly shows the overestimation made when using the full dispersal modeling approach. Generally, species distributions modeling under the influence of realistic dispersal limitation in the 21st century have produced results closer to ND rather than to FD scenario. Our results revealed that future barrier growth is more effective than the climate change impacts on species distribution, while affirming the particular importance of landscape structure and its effect on future plant distribution to be considered in fine-scale studies and the fact that more realistic results are obtained when incorporating future landscape changes in modeling (as stated by e.g. Gilman et al., 2010; Kissling et al., 2010). The projected future distributions under the two climate scenarios used in this study showed small differences. Nonetheless, significant differences were observed between SDD and LDD distribution scenarios for all three plants. Finally, our results highlighted the fact that the appropriate interpretation of the outputs of the models requires a careful selection of dispersal scenarios related to the mentioned questions when projecting them into new environments.

Chapter 6

General Discussion

6.1. General finding of this thesis

This thesis is compiled in five principal chapters. Chapter 1 reviews the fundamentals of this research that were supposed to deal within this project (e.g. problem statement, research questions, research objectives and research approach). Chapter 2 different algorithms to the classification of satellite images were examined to determine the best classifier approach with respect to characteristics of the study area and features of the used satellite images. The best method was then used to obtain the land-cover maps corresponding to different years. In Chapter 3, the classified maps in the previous chapter were imported into CA-Markov model. This model uses the changes observed in the past few years to simulate possible future changes in future land-cover over time and space. These simulated maps show the future changes in dispersal barriers and suitable and unsuitable habitats. In Chapter 4 several SDM models were evaluated to determine the model and the set of predictors most suitable to forecast the distribution of studied species. In chapter 5 we investigate the potential impacts of climate and landscape fragmentation changes on distribution of three individual plant species.

Can regression-based models as well as machine-learning models be used for prediction suitable habitat maps?

In Chapter 4, we compared the performance of some of the most common methods of presence-absence distribution models. Specifically, two categories of methods, machine-learning algorithms and regression-based methods, were compared. It was found that regression-based methods (GAM, GLM, and MARS), showed a better performance than machine-learning algorithms (RF, ANN, and GBM), whereas the opposite was found by Bucklin et al. (2015).

One of the machine-learning methods, GBM, produced a prediction map that performed as well as the regression models (mean AUC= 0.87). A visual comparison between the predicted maps revealed that the map produced by GBM is very similar to maps obtained from the regression models (Chapter 4, Figures 4, 5, and 6). Among all predictor techniques used in this

study, GAM and MARS performed better. This finding verifies previous results of species dispersal modeling (Leathwick et al., 2005; Heikkinen et al., 2012).

We also investigated the impact of different set of explanatory variable (climate-only, topography-only and topo-climatic) on model performance. Selection of ecological factors to apply as predictors is one of the big challenges in species distribution modeling (SDM) (Araujo and Guisan, 2006). Selection of predictors with direct effect on species distribution is the best solution for this problem (Austin, 2002, 2007). In many cases, due to unavailability of adequate information, limitation of accessible resources, or incomplete ecological knowledge, it is not generally possible to examine a large number of variables in a study (Bucklin et al., 2015).

Given the simplicity of climate data provision and collection, in some studies, researchers only use climatic factors as the driver of species distributions (Pearson and Dawson, 2003). For example, Martinez-Meyer (2005) stated that the combination of precipitation and temperature as the two main factors among climate parameters could express the extent of physiological tolerance for living organisms.

Some researchers criticize studies in which climatic factors are only used as predictors, because they think there is not enough evidence for climatic range determination in species distributions (Bahn and McGill, 2007; Beale et al., 2008), and climatic data alone cannot show the environmental conditions required for the establishment of species (Araujo and Peterson, 2012; Bucklin et al., 2015). In response to those criticisms, along with other modelers (Austin and Van Niel, 2011) we include additional, non-climate predictors (topography variables) beside climate variables in SDMs.

Topography variable are commonly believed to influence species ranges, and have been often been included in SDMs (e.g. Marmion et al., 2009; Tsoar et al., 2007). Our results have shown that using climatic and topographic parameters improved the accuracy of predictions. This result is inconsistent with the results obtained by Tsoar et al. (2007). Many previous studies tried to predict suitable sites for establishment of the species based on similar indices (Box et al., 1993; Carpenter et al., 1993; Gonzalez et al., 2010; Kissling et al., 2010).

What is the potential impact of climate change on plant species?

Climate change is one of the factors that cause significant changes in plant distribution and landscape change. So development and the use of models that can incorporate climatic change scenarios are of great important, especially in local scale studies. The major achievement of this thesis has been to assessment of potential climate change impacts on three plant species distribution in future (Chapter 5).

Our results show that the predicted distribution areas for all the three species under RCP8.5 scenario will largely increase in the coming decades. Response curves of predictor variables in chapter 4 (Fig. 3) clearly show that target plants prefer low altitude, low precipitation and high temperature in the study area, which corroborate our results in chapter 5. Also, studies revealed that rises in temperature have been still suitable for lowland species since the past decades (Vittoz et al., 2009).

Our results demonstrated that the studied species have no greatly different responses to different climate change scenarios (Chapter 5: Fig. 4, Table 3) which might be due to little restrictions they have had for distribution and have thus been less affected by varied climate scenarios.

How much does the inclusion of dispersal limitation events affect projections?

Migration processes should also be paid attention to in order to study changes in the ranges of biological diversities of species caused by climate change. For this very reason, dispersal limitation have been mentioned as one of the most important sources of uncertainty in the process of studying species distributions as influenced by climate change (Araújo and Guisan, 2006). SDMs have the capacity to investigate distribution of species on various scales (from fine to large), although they do not study the capability and/or the power of species dispersal (Guisan and Zimmermann, 2000; Engler et al., 2009). Therefore, incorporating migration processes into SDMs is one of the basic factors in improving predictions under varying conditions (Guisan and Thuiller, 2005; Zurell et al., 2009; Meier, 2011). Thus, the importance of dispersal limitations in projections of future species distribution under climate change has been given special attention in this thesis (Chapter 5).

When assessing species dispersal in our study area (local or regional assessment), considering no-dispersal provided satisfying results that were close to those obtained from the simulations

that accounted for dispersal limitations. Although the obtained results must be tested in other study regions and on other plant species, yet it seems logical to think that these results can be extrapolated to other regions with similar features and sizes.

Nevertheless, as mentioned above, this obtained result is only valid for projections made on local to regional scales. Moreover, it is only valid for non-mountainous regions, because results for mountainous regions are expected to be very different. In other words, simulations of limited dispersal can be closer to full-dispersal than to no-dispersal (e.g., see Engler et al., 2009).

One of the parameters that was particularly paid attention to in this research was long distance dispersal (LDD) events that allow movement and transfer of seeds over long distances often through unusual dispersal agents and methods. Despite the fact that many studies have been carried out on seed dispersal, the importance of long distance dispersal in migration and dispersal of plants is still a controversial issue (Pearson, 2006).

Although various methods were developed to estimate LDD events, and considerable progress was made in modeling LDD (e.g., Teckenberg, 2003; Soons et al., 2004; Nathan et al., 2006), yet it is still difficult to combine them accurately into species distribution and dispersal. In fact, some have acknowledged that accurate prediction of LDD events may be impossible due to their unpredictability (Clark et al., 2003). Nevertheless, it seems illogical to ignore this type of dispersal and omit it from studies on distribution of plant species (Engler, 2009).

When the importance of LLD events is discussed, using certain models, which make possible studying various scenarios and states, allow researchers to scrutinize species distribution process more closely. For example, in this specific case, we can test whether incorporating this type of dispersal can have significant effects on projections. Moreover, repeating simulations with LDD events may reveal LDD-related uncertainty both temporally and spatially. This type of uncertainties must be included in studies in order to increase the ability of models in making projections and somehow give researchers more assurance. This issue becomes significantly important when the results of these models are used to enforce climate change reduction policies.

I applied different dispersal scenarios in an SDM-based projection framework (MigClim), in which migration depends on dispersal distance and landscape fragmentation. MigClim has already been used to investigate the relative impacts of dispersal distance in a local study in

Western Swiss Alps (Engler et al., 2009). Engler (2009) stated that the MigClim model has considerable strengths. For example, he states that this model can be linked with different modeling techniques, and its parameters are flexible. Engler (2009) also suggested that this model can examine the potential distribution of a species steadily over time. Also, there is the possibility for a researcher to not only show the potential distribution map of the desired species at the end of the simulation, but also examine the process of the distribution in specific intervals.

To what extent does landscape fragmentation prevent the movement of species toward suitable habitats?

One of major methodological limitation of SDMs is that they do not take limitation of migration into account. Nowadays, with the increase in landscape fragmentation, the significance of this limitation is more than ever before, since there have been more barriers in the way of species movement pattern that can reduce the speed and rate of dispersal (Engler, 2009).

Most of modelers ignore the change of landscape fragmentation over time. In this study we simulated future landscape changes and projected future plants distribution under two scenarios (static and dynamic). A cellular automata-Markovian spatial modeling process was used (chapter 3) to simulate future landscape changes (dynamic scenario) based on biophysical and proximate drivers. Our results showed that the barrier areas is expected to cover 45.7% of our study area in year 2080, which means a 28.4% increase in comparison with its current distribution.

As expected, an increase in the barrier areas led to a decrease in suitable and colonised habitat areas for the studied species. For instance, the potential suitable habitat of *F. rupicola* is projected to lose 651 Km^2 when projections were carried out under dynamic landscape scenario, which means a 9.4% reduction in comparison with its distribution under static scenario. Our projections indicated that expected change in landscape fragmentation will have a strong role in reducing the distribution areas of plants in the coming decades, especially in local studies (Pearson and Dawson, 2003; Engler et al., 2011). In total, our results revealed that change in landscape fragmentation is more effective than the climate change impacts on species distributions in this study.

References

Adam, E., Mutanga, O., Abdel-Rahman, E.M., & Ismail, R. (2014). Estimating standing biomass in papyrus (Cyperus papyrus L.) swamp: exploratory of in situ hyperspectral indices and random forest regression. *International Journal of Remote Sensing, 35*, 693-714

Aeschimann, D., & Heitz, C. (1996). *Index synonymique de la flore de Suisse et territoires limitrophes (ISFS)*. Centre du réseau Suisse de floristique

Ahmed, B., & Ahmed, R. (2012). Modeling Urban Land Cover Growth Dynamics Using Multi- Temporal Satellite Images: A Case Study of Dhaka, Bangladesh. *ISPRS International Journal of Geo-Information, 1*, 3

Akaike, H. (1998). A New Look at the Statistical Model Identification. In E. Parzen, K. Tanabe, & G. Kitagawa (Eds.), *Selected Papers of Hirotugu Akaike* (pp. 215-222): Springer New York

Allouche, O., Tsoar, A., & Kadmon, R. (2006). Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology, 43*, 1223-1232

Anderson, J.R. (1976). *A Land Use and Land Cover Classification System for Use with Remote Sensor Data*. U.S. Government Printing Office

Angell, D.L., & McClaran, M.P. (2001). Long-term influences of livestock management and a non-native grass on grass dynamics in the Desert Grassland. *Journal of Arid Environments, 49*, 507-520

Araújo, M.B., & Guisan, A. (2006). Five (or so) challenges for species distribution modelling. *Journal of Biogeography, 33*, 1677-1688

Araújo, M.B., & Pearson, R.G. (2005). Equilibrium of species' distributions with climate. *Ecography, 28*, 693-695

Araya, Y.H., & Cabral, P. (2010). Analysis and Modeling of Urban Land Cover Change in Setúbal and Sesimbra, Portugal. *Remote Sensing, 2*, 1549

Austin, M.P. (2002). Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecological Modelling, 157*, 101-118

B. Lind, S. Stein, A. Kärcher, & Klein, M. (2009). Where have All the Flowers Gone? Grünland im Umbruch. In. Bonn: German Federal Agency for Nature Conservation (BfN)

Bagan, H., & Yamagata, Y. (2012). Landsat analysis of urban growth: How Tokyo became the world's largest megacity during the last 40years. *Remote Sensing of Environment, 127*, 210-222

Balmford, A., Moore, J.L., Brooks, T., Burgess, N., Hansen, L.A., Williams, P., & Rahbek, C. (2001). Conservation conflicts across Africa. *Science, 291*, 2616-2619

Barbet-Massin, M., Thuiller, W., & Jiguet, F. (2012). The fate of European breeding birds under climate, land-use and dispersal scenarios. *Global Change Biology, 18*, 881-890

Barry, S., & Elith, J. (2006). Error and uncertainty in habitat models. *Journal of Applied Ecology, 43*, 413-423

Basse, R.M., Omrani, H., Charif, O., Gerber, P., & Bódis, K. (2014). Land use changes modelling using advanced methods: Cellular automata and artificial neural networks. The spatial and explicit representation of land cover dynamics at the cross-border region scale. *Applied Geography, 53*, 160-171

Bateman, B.L., Murphy, H.T., Reside, A.E., Mokany, K., & VanDerWal, J. (2013). Appropriateness of full-, partial- and no-dispersal scenarios in climate change impact modelling. *Diversity and Distributions, 19*, 1224-1234

Bell, S. (1999). *Landscape: Pattern, Perception, and Process*. E & FN Spon

Ben-Hur, A., & Weston, J. (2010). A User's Guide to Support Vector Machines. In O. Carugo, & F. Eisenhaber (Eds.), *Data Mining Techniques for the Life Sciences* (pp. 223-239): Humana Press

Bennie, J., Huntley, B., Wiltshire, A., Hill, M.O., & Baxter, R. (2008). Slope, aspect and climate: Spatially explicit and implicit models of topographic microclimate in chalk grassland. *Ecological Modelling, 216*, 47-59

Benz, U.C., Hofmann, P., Willhauck, G., Lingenfelder, I., & Heynen, M. (2004). Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of Photogrammetry and Remote Sensing, 58*, 239-258

Blaschke, T. (2010). Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing, 65*, 2-16

Bock, M., Xofis, P., Mitchley, J., Rossner, G., & Wissen, M. (2005). Object-oriented methods for habitat mapping at multiple scales – Case studies from Northern Germany and Wye Downs, UK. *Journal for Nature Conservation, 13*, 75-89

Boeckmann, M., & Joyner, T.A. (2014). Old health risks in new places? An ecological niche model for I. ricinus tick distribution in Europe under a changing climate. *Health & Place, 30*, 70-77

Box, E.O., Crumpacker, D.W. & Hardin, E.D. (1993) A climatic model for location of plant species in Florida, USA. *Journal of Biogeography, 20*, 629–644

Braun, E., van den Berg, P.L., & van der Meer, J. (2012). *National Policy Responses to Urban Challenges in*

Europe. Ashgate Publishing Limited
Breiman, L. (2001). Random Forests. *Machine Learning, 45*, 5-32
Breiman, L., Friedman, J., Stone, C.J., & Olshen, R.A. (1984). *Classification and Regression Trees*. Taylor & Francis
Brooker, R.W. (2006). Plant–plant interactions and environmental change. *New Phytologist, 171*, 271-284
Brown, R.G., James, A.F., Pitchford, J.W., & Plank, M.J. (2012). Habitat fragmentation: Simple models for local persistence and the spread of invasive species. *Journal of Theoretical Biology, 310*, 231-238
Buchmann, C.M., Schurr, F.M., Nathan, R., & Jeltsch, F. (2013). Habitat loss and fragmentation affecting mammal and bird communities—The role of interspecific competition and individual space use. *Ecological Informatics, 14*, 90-98
Bucklin, D.N., Basille, M., Benscoter, A.M., Brandt, L.A., Mazzotti, F.J., Romañach, S.S., Speroterra, C., & Watling, J.I. (2015). Comparing species distribution models constructed with different subsets of environmental predictors. *Diversity and Distributions, 21*, 23-35
Bullock, J.M. (2012). Plant dispersal and the velocity of climate change. In J. Clobert, M. Baguette, T.G. Benton, & J.M. Bullock (Eds.), *Dispersal Ecology and Evolution* (pp. 366-377): OUP Oxford
Bullock, J.M., White, S.M., Prudhomme, C., Tansey, C., Perea, R., & Hooftman, D.A.P. (2012). Modelling spread of British wind-dispersed plants under future wind speeds in a changing climate. *Journal of Ecology, 100*, 104-115
Bustamante, J., & Seoane, J. (2004). Predicting the distribution of four species of raptors (Aves: Accipitridae) in southern Spain: statistical models work better than existing maps. *Journal of Biogeography, 31*, 295-306
Carpenter, G., Gillison, A.N. & Winter, J. (1993) DOMAIN — A flexible modeling procedure for mapping potential distributions of plants and animals. *Biodiversity and Conservation, 2,* 667–680
Castella, J.-C., Pheng Kam, S., Dinh Quang, D., Verburg, P.H., & Thai Hoanh, C. (2007). Combining top-down and bottom-up modelling approaches of land use/cover change to support public policies: Application to sustainable management of natural resources in northern Vietnam. *Land Use Policy, 24*, 531-545
Castilla, G., & Hay, G.J. (2008). Image objects and geographic objects. In T. Blaschke, S. Lang, & G. Hay (Eds.), *Object-Based Image Analysis* (pp. 91-110): Springer Berlin Heidelberg
CBD (2010). Fourth National Report under the Convention on Biological Diversity (CBD)—Germany. In
Chen, J., Gong, P., He, C., Pu, R., & Shi, P. (2003). Land-use/land-cover change detection using improved change-vector analysis. *Photogrammetric Engineering & Remote Sensing, 69*, 369-379
Clark, J.S., Lewis, M., McLachlan, J.S., & HilleRisLambers, J. (2003). ESTIMATING POPULATION SPREAD: WHAT CAN WE FORECAST AND HOW WELL? *Ecology, 84*, 1979-1988
Collingham, Y.C., & Huntley, B. (2000). Impacts of Habitat Fragmentation and Patch Size upon Migration Rates. *Ecological Applications, 10*, 131-144
Collins, S.L., & Gordon, E.U. (1985). Seed Predation, Seed Dispersal, and Disturbance in Grasslands: A Comment. *The American Naturalist, 125*, 866-872
Conchedda, G., Durieux, L., & Mayaux, P. (2008). An object-based method for mapping and change analysis in mangrove ecosystems. *ISPRS Journal of Photogrammetry and Remote Sensing, 63*, 578-589
Congalton, R.G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment, 37*, 35-46
Congalton, R.G., & Green, K. (2002). *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*. CRC Press
Connell, J.H. (1961). The Influence of Interspecific Competition and Other Factors on the Distribution of the Barnacle Chthamalus Stellatus. *Ecology, 42*, 710-723
Coppin, P., Jonckheere, I., Nackaerts, K., Muys, B., & Lambin, E. (2004). Review ArticleDigital change detection methods in ecosystem monitoring: a review. *International Journal of Remote Sensing, 25*, 1565-1596
Cordeiro, N.J., Ndangalasi, H.J., McEntee, J.P., & Howe, H.F. (2009). Disperser limitation and recruitment of an endemic African tree in a fragmented landscape. *Ecology, 90*, 1030-1041
Dai, X., & Khorram, S. (1999). Remotely sensed change detection based on artificial neural networks. *photogrammetric engineering and remote sensing, 65*, 1187-1194
Darwin, C. (1859). *On the Origin of Species by Means of Natural Selection, Or, The Preservation of Favoured Races in the Struggle for Life*. J. Murray
Davin, E.L., & de Noblet-Ducoudré, N. (2010). Climatic Impact of Global-Scale Deforestation: Radiative versus Nonradiative Processes. *Journal of Climate, 23*, 97-112
Davis, A.J., Jenkinson, L.S., Lawton, J.H., Shorrocks, B., & Wood, S. (1998). Making mistakes when predicting shifts in species range in response to global warming. *Nature, 391*, 783-786
de Leeuw, J., Jia, H., Yang, L., Liu, X., Schmidt, K., & Skidmore, A.K. (2006). Comparing accuracy assessments

to infer superiority of image classification methods. *International Journal of Remote Sensing, 27*, 223-232
Deng, X., Huang, J., Rozelle, S., & Uchida, E. (2008). Growth, population and industrialization, and urban land expansion of China. *Journal of Urban Economics, 63*, 96-115
Dewhirst, S., & Lutscher, F. (2009). Dispersal in heterogeneous habitats: thresholds, spatial scales, and approximate rates of spread. *Ecology, 90*, 1338-1345
Dingle Robertson, L., & King, D.J. (2011). Comparison of pixel- and object-based classification in land cover change mapping. *International Journal of Remote Sensing, 32*, 1505-1529
Dirnböck, T., Dullinger, S., & Grabherr, G. (2003). A regional impact assessment of climate and land-use change on alpine vegetation. *Journal of Biogeography, 30*, 401-417
Dirnböck, T., Essl, F., & Rabitsch, W. (2011). Disproportional risk for habitat loss of high-altitude endemic species under climate change. *Global Change Biology, 17*, 990-996
Division, G.O.o.M.G.E., & Settel, A. (1946). *A year of Potsdam: the German economy since the surrender*. Lithographed by the Adjutant General, OMGUS
Dullinger, S., Gattringer, A., Thuiller, W., Moser, D., Zimmermann, N.E., Guisan, A., Willner, W., Plutzar, C., Leitner, M., Mang, T., Caccianiga, M., Dirnbock, T., Ertl, S., Fischer, A., Lenoir, J., Svenning, J.-C., Psomas, A., Schmatz, D.R., Silc, U., Vittoz, P., & Hulber, K. (2012). Extinction debt of high-mountain plants under twenty-first-century climate change. *Nature Clim. Change, 2*, 619-622
Duro, D.C., Franklin, S.E., & Dubé, M.G. (2012). A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. *Remote Sensing of Environment, 118*, 259-272
Duro, D.C., Franklin, S.E., & Dubé, M.G. (2012). Multi-scale object-based image analysis and feature selection of multi-sensor earth observation imagery using random forests. *International Journal of Remote Sensing, 33*, 4502-4526
Eastman, J.R. (2006). *IDRISI Andes*. Worcester, MA, USA: Clark University
El-Hallaq, M.A., & Habboub, M.O. (2015). Using Cellular Automata-Markov Analysis and Multi Criteria Evaluation for Predicting the Shape of the Dead Sea. *Advances in Remote Sensing, Vol.04No.01*, 13
Elith, J., Burgman, M.A., & Regan, H.M. (2002). Mapping epistemic uncertainties and vague concepts in predictions of species distribution. *Ecological Modelling, 157*, 313-329
Elith, J., & Graham, C.H. (2009). Do they? How do they? WHY do they differ? On finding reasons for differing performances of species distribution models. *Ecography, 32*, 66-77
Elith, J., Kearney, M., & Phillips, S. (2010). The art of modelling range-shifting species. *Methods in Ecology and Evolution, 1*, 330-342
Engler, R. (2009). Unfolding nature in silico – Overcoming practical and methodological limitations of species distribution models. Biology. University of Lausanne, p 376.
Engler, R., & Guisan, A. (2009). MigClim: Predicting plant distribution and dispersal in a changing climate. *Diversity and Distributions, 15*, 590-601
Engler, R., Hordijk, W., & Guisan, A. (2012). The MIGCLIM R package – seamless integration of dispersal constraints into projections of species distribution models. *Ecography, 35*, 872-878
Engler, R., Randin, C.F., Thuiller, W., Dullinger, S., Zimmermann, N.E., AraÚJo, M.B., Pearman, P.B., Le Lay, G., Piedallu, C., Albert, C.H., Choler, P., Coldea, G., De Lamo, X., DirnbÖCk, T., GÉGout, J.-C., GÓMez-GarcÍA, D., Grytnes, J.-A., Heegaard, E., HØIstad, F., NoguÉS-Bravo, D., Normand, S., PuŞCaŞ, M., SebastiÀ, M.-T., Stanisci, A., Theurillat, J.-P., Trivedi, M.R., Vittoz, P., & Guisan, A. (2011). 21st century climate change threatens mountain flora unequally across Europe. *Global Change Biology, 17*, 2330-2341
Engler, R., Randin, C.F., Vittoz, P., Czáka, T., Beniston, M., Zimmermann, N.E., & Guisan, A. (2009). Predicting future distributions of mountain plants under climate change: does dispersal capacity matter? *Ecography, 32*, 34-45
EU-COM (2009). Composite Report on the Conservation Status of Habitat Types and Species as required under Article 17 of Habitats Directive. In (p. 358). Brussles: Report from the Commission to the Council and the European Parliament
FAO (2011). State of the World's Forests. In (p. 179). Rome: Forestry Department
Fauvel, M., Chanussot, J., & Benediktsson, J.A. (2006). Evaluation of Kernels for Multiclass Classification of Hyperspectral Remote Sensing Data. In, *Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on* (pp. II-II)
Fischer, J., & Lindenmayer, D.B. (2007). Landscape modification and habitat fragmentation: a synthesis. *Global Ecology and Biogeography, 16*, 265-280
Fitter, A.H., & Fitter, R.S.R. (2002). Rapid Changes in Flowering Time in British Plants. *Science, 296*, 1689-1691
Foley, J.A., DeFries, R., Asner, G.P., Barford, C., Bonan, G., Carpenter, S.R., Chapin, F.S., Coe, M.T., Daily,

G.C., Gibbs, H.K., Helkowski, J.H., Holloway, T., Howard, E.A., Kucharik, C.J., Monfreda, C., Patz, J.A., Prentice, I.C., Ramankutty, N., & Snyder, P.K. (2005). Global Consequences of Land Use. *Science, 309*, 570-574

Fontaine, T.A., Cruickshank, T.S., Arnold, J.G., & Hotchkiss, R.H. (2002). Development of a snowfall–snowmelt routine for mountainous terrain for the soil water assessment tool (SWAT). *Journal of Hydrology, 262*, 209-223

Foody, G.M., & Mathur, A. (2004). Toward intelligent training of supervised image classifications: directing training data acquisition for SVM classification. *Remote Sensing of Environment, 93*, 107-117

Franklin, J. (2009). *Mapping Species Distributions: Spatial Inference and Prediction*. Cambridge: Cambridge University Press

Franklin, J. (2010). Moving beyond static species distribution models in support of conservation biogeography. *Diversity and Distributions, 16*, 321-330

Friedman, J.H., Hastie, T., & Tibshirani, R. (2000). Additve logistic regression: a statistical view of boosting. *Ann. Stat*, 28, 337–407

Fung, T., So, L.L.H., Chen, Y., Shi, P., & Wang, J. (2008). Analysis of green space in Chongqing and Nanjing, cities of China with ASTER images using object-oriented image classification and landscape metric analysis. *International Journal of Remote Sensing, 29*, 7159-7180

Funkenberg, T., Binh, T.T., Moder, F., & Dech, S. (2014). The Ha Tien Plain – wetland monitoring using remote-sensing techniques. *International Journal of Remote Sensing, 35*, 2893-2909

Garzón, M.B., Blazek, R., Neteler, M., Dios, R.S.d., Ollero, H.S., & Furlanello, C. (2006). Predicting habitat suitability with machine learning models: The potential area of Pinus sylvestris L. in the Iberian Peninsula. *Ecological Modelling, 197*, 383-393

GBO3 (2010). Secretariat of the Convention on Biological Diversity. In. Global Biodiversity Outlook 3—Executive Summary, Montreal.

Genuer, R., Poggi, J.M., & Tuleau-Malot, C. (2010). Variable selection using random forests. *Pattern Recognition Letters, 31*, 2225-2236

Geomatica (2013). Atmospheric correction (with ATCOR). In

(Germany), O.o.M.G. (1946). *A year of Potsdam: the German economy since the surrender*. Lithographed by the Adjutant General, OMGUS

Ghimire, B., Rogan, J., & Miller, J. (2010). Contextual land-cover classification: incorporating spatial dependence in land-cover classification models using random forests and the Getis statistic. *Remote Sensing Letters, 1*, 45-54

Gibson, L., Barrett, B., & Burbidge, A. (2007). Dealing with uncertain absences in habitat modelling: a case study of a rare ground-dwelling parrot. *Diversity and Distributions, 13*, 704-713

Gilman, S.E., Urban, M.C., Tewksbury, J., Gilchrist, G.W., & Holt, R.D. (2010). A framework for community interactions under climate change. *Trends in Ecology & Evolution, 25*, 325-331

Gislason, P.O., Benediktsson, J.A., & Sveinsson, J.R. (2006). Random forests for land cover classification. *Pattern Recognition Letters, 27*, 294-300

Gong, P. (1993). Change detection using principal component analysis and fuzzy set theory. *Canadian Journal of Remote Sensing, 19*, 22-29

González, C., Wang, O., Strutz, S.E., González-Salazar, C., Sánchez-Cordero, V., & Sarkar, S. (2010). Climate Change and Risk of Leishmaniasis in North America: Predictions from Ecological Niche Models of Vector and Reservoir Species. *PLoS Neglected Tropical Diseases, 4*, e585

Grabs, T., Seibert, J., Bishop, K., & Laudon, H. (2009). Modeling spatial patterns of saturated areas: A comparison of the topographic wetness index and a dynamic distributed model. *Journal of Hydrology, 373*, 15-23

Guan, D., Li, H., Inohae, T., Su, W., Nagaie, T., & Hokao, K. (2011). Modeling urban land use change by the integration of cellular automaton and Markov model. *Ecological Modelling, 222*, 3761-3772

Guisan, A., & Thuiller, W. (2005). Predicting species distribution: offering more than simple habitat models. *Ecology Letters, 8*, 993-1009

Guisan, A., & Zimmermann, N.E. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling, 135*, 147-186

Guisan, A., Zimmermann, N.E., Elith, J., Graham, C.H., Phillips, S., & Peterson, A.T. (2007). WHAT MATTERS FOR PREDICTING THE OCCURRENCES OF TREES: TECHNIQUES, DATA, OR SPECIES' CHARACTERISTICS? *Ecological Monographs, 77*, 615-630

Hadjimitsis, D.G., Papadavid, G., Agapiou, A., Themistocleous, K., Hadjimitsis, M.G., Retalis, A., Michaelides, S., Chrysoulakis, N., Toulios, L., & Clayton, C.R.I. (2010). Atmospheric correction for satellite remotely sensed data intended for agricultural applications: impact on vegetation indices. *Nat. Hazards Earth Syst. Sci., 10*, 89-

95
Hafeez, K., Zhang, Y., & Malak, N. (2002). Determining key capabilities of a firm using analytic hierarchy process. *International Journal of Production Economics, 76*, 39-51
Hampe, A. (2011). Plants on the move: The role of seed dispersal and initial population establishment for climate-driven range expansions. *Acta Oecologica, 37*, 666-673
Hanley, J.A., & McNeil, B.J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology, 143*, 29-36
Hannah, L.E.E. (2011). Climate Change, Connectivity, and Conservation Success. *Conservation Biology, 25*, 1139-1142
Hansen, M., & DeFries, R. (2004). Detecting Long-term Global Forest Change Using Continuous Fields of Tree-Cover Maps from 8-km Advanced Very High Resolution Radiometer (AVHRR) Data for the Years 1982–99. *Ecosystems, 7*, 695-716
Hasler, N., Werth, D., & Avissar, R. (2009). Effects of tropical deforestation on global hydroclimate: A multimodel ensemble analysis. *Journal of Climate, 22*, 1124-1141
Hastie, T.J., & Tibshirani, R.J. (1990). *Generalized Additive Models.* . New York: Chapman and Hall
Haykin, S. (1994). *Neural Networks: A Comprehensive Foundation*. Macmillan College Publishing Company
He, J., Liu, Y., Yu, Y., Tang, W., Xiang, W., & Liu, D. (2013). A counterfactual scenario simulation approach for assessing the impact of farmland preservation policies on urban sprawl and food security in a major grain-producing area of China. *Applied Geography, 37*, 127-138
Heikkinen, R.K., Luoto, M., Araújo, M.B., Virkkala, R., Thuiller, W., & Sykes, M.T. (2006). Methods and uncertainties in bioclimatic envelope modelling under climate change. *Progress in Physical Geography, 30*, 751-777
Heikkinen, R.K., Marmion, M., & Luoto, M. (2012). Does the interpolation accuracy of species distribution models come at the expense of transferability? *Ecography, 35*, 276-288
Hijmans, R.J., & Graham, C.H. (2006). The ability of climate envelope models to predict the effect of climate change on species distributions. *Global Change Biology, 12*, 2272-2281
Hong, Y.T., Hong, B., Lin, Q.H., Shibata, Y., Hirota, M., Zhu, Y.X., Leng, X.T., Wang, Y., Wang, H., & Yi, L. (2005). Inverse phase oscillations between the East Asian and Indian Ocean summer monsoons during the last 12000 years and paleo-El Nino. *Earth and Planetary Science Letters, 231*, 337–346
Hostert, P., Röder, A., & Hill, J. (2003). Coupling spectral unmixing and trend analysis for monitoring of long-term vegetation dynamics in Mediterranean rangelands. *Remote Sensing of Environment, 87*, 183-197
Houghton, R.A. (1994). The Worldwide Extent of Land-use Change: In the last few centuries, and particularly in the last several decades, effects of land-use change have become global. *BioScience, 44*, 305-313
Howarth, P.J., & Boasson, E. (1983). Landsat digital enhancements for change detection in urban environments. *Remote Sensing of Environment, 13*, 149-160
Hu, Z., & Lo, C.P. (2007). Modeling urban growth in Atlanta using logistic regression. *Computers, Environment and Urban Systems, 31*, 667-688
Hua, W.J., & Chen, H.S. (2013). Impacts of regional-scale land use/land cover change on diurnal temperature range. *Advances in Climate Change Research, 4*, 166-172
Huang, C., Davis, L.S., & Townshend, J.R.G. (2002). An assessment of support vector machines for land cover classification. *International Journal of Remote Sensing, 23*, 725-749
Huberty, C.J. (1994). *Applied Discriminant Analysis.* New York: Wiley
Huntley, B., Barnard, P., Altwegg, R., Chambers, L., Coetzee, B.W.T., Gibson, L., Hockey, P.A.R., Hole, D.G., Midgley, G.F., Underhill, L.G., & Willis, S.G. (2010). Beyond bioclimatic envelopes: dynamic species' range and abundance modelling in the context of climatic change. *Ecography, 33*, 621-626
Hutchinson, G.E. (1957). Population studies - animal ecology and demography - concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology, 22*, 422-427
Im, J., & Jensen, J.R. (2005). A change detection model based on neighborhood correlation image analysis and decision tree classification. *Remote Sensing of Environment, 99*, 326-340
Imbach, P.A., Locatelli, B., Molina, L.G., Ciais, P., & Leadley, P.W. (2013). Climate change and plant dispersal along corridors in fragmented landscapes of Mesoamerica. *Ecology and Evolution, 3*, 2917-2932
Isabelle, B., Damien, G., & Wilfried, T. (2014). FATE-HD: a spatially and temporally explicit integrated model for predicting vegetation structure and diversity at regional scale. *Global Change Biology, 20*, 2368-2378
Jensen, J.R. (2005). *Introductory Digital Image Processing: A Remote Sensing Perspective*. Prentice Hall
Jensen, J.R. (2013). *Remote Sensing of the Environment: Pearson New International Edition: An Earth Resource Perspective*. Pearson Education Limited
Jeschke, J.M., & Strayer, D.L. (2008). Usefulness of Bioclimatic Models for Studying Climate Change and

Invasive Species. *Annals of the New York Academy of Sciences, 1134*, 1-24
Jetz, W., Wilcove, D.S., & Dobson, A.P. (2007). Projected Impacts of Climate and Land-Use Change on the Global Diversity of Birds. *PLoS Biology, 5*, e157
Johnson, R.D., & Kasischke, E. (1998). Change vector analysis: a technique for the multispectral monitoring of land cover and condition. *International Journal of Remote Sensing, 19*, 411-426
Jones, P.G., & Thornton, P.K. (2013). Generating downscaled weather data from a suite of climate models for agricultural modelling applications. *Agricultural Systems, 114*, 1-5
Kampouraki, M., Wood, G.A., & Brewer, T.R. (2008). Opportunities and limitations of object based image analysis for detecting urban impervious and vegetated surfaces using true-colour aerial photography. In T. Blaschke, S. Lang, & G. Hay (Eds.), *Object-Based Image Analysis* (pp. 555-569): Springer Berlin Heidelberg
Kamusoko, C., Aniya, M., Adi, B., & Manjoro, M. (2009). Rural sustainability under threat in Zimbabwe - Simulation of future land use/cover changes in the Bindura district based on the Markov-cellular automata model. *Applied Geography, 29*, 435-447
Kang, J.H., Lee, S.W., Cho, K.H., Ki, S.J., Cha, S.M., & Kim, J.H. (2010). Linking land-use type and stream water quality using spatial data of fecal indicator bacteria and heavy metals in the Yeongsan river basin. *Water Research, 44*, 4143-4157
Kawakami, K., Mizusawa, L., & Higuchi, H. (2009). Re-established mutualism in a seed-dispersal system consisting of native and introduced birds and plants on the Bonin Islands, Japan. *Ecological Research, 24*, 741-748
Kelly, A.E., & Goulden, M.L. (2008). Rapid shifts in plant distribution with recent climate change. *Proceedings of the National Academy of Sciences, 105*, 11823-11826
Kennedy, R.E., Cohen, W.B., & Schroeder, T.A. (2007). Trajectory-based change detection for automated characterization of forest disturbance dynamics. *Remote Sensing of Environment, 110*, 370-386
Keshtkar, H.R., Azarnivand, H., Arzani, H., Alavipanah, S.K., & Mellati, F. (2013). Land Cover Classification Using IRS-1D Data and a Decision Tree Classifier *Desert, 17*, 137-146
Kinezaki, N., Kawasaki, K., & Shigesada, N. (2010). The effect of the spatial configuration of habitat fragmentation on invasive spread. *Theoretical Population Biology, 78*, 298-308
Kinlan, B.P., & Gaines, S.D. (2003). PROPAGULE DISPERSAL IN MARINE AND TERRESTRIAL ENVIRONMENTS: A COMMUNITY PERSPECTIVE. *Ecology, 84*, 2007-2020
Kirika, J.M., Bleher, B., Böhning-Gaese, K., Chira, R., & Farwig, N. (2008). Fragmentation and local disturbance of forests reduce frugivore diversity and fruit removal in Ficus thonningii trees. *Basic and Applied Ecology, 9*, 663-672
Kissling, W.D., Field, R., Korntheuer, H., Heyder, U., & Böhning-Gaese, K. (2010). Woody plants and the prediction of climate-change impacts on bird diversity. *Philosophical Transactions of the Royal Society B: Biological Sciences, 365*, 2035-2045
Körner, C. (2003). *Alpine plant life - Functional Plant Ecology of High Mountain Ecosystems*. (2nd ed.). Heidelberg: Springer
Kovats, R.S., Valentini, R., Bouwer, L.M., Georgopoulou, E., Jacob, D., Martin, E., Rounsevell, M., & Soussana, J.F. (2014). Europe. In V.R. Barros, C.B. Field, D.J. Dokken, M.D. Mastrandrea, K.J. Mach, T.E. Bilir, M. Chatterjee, K.L. Ebi, Y.O. Estrada, R.C. Genova, B. Girma, E.S. Kissel, A.N. Levy, S. MacCracken, P.R. Mastrandrea, & L.L. White (Eds.), *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part B: Regional Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel of Climate Change* (pp. 1267-1326). Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press
Kraemer, H.C. (1982). *Kappa coefficient. In S. Kotz and N. L. Johnson (Eds.), Encyclopedia of Statistical Sciences.* . New York: John Wiley & Sons
Kuemmerle, T., Chaskovskyy, O., Knorn, J., Radeloff, V.C., Kruhlov, I., Keeton, W.S., & Hostert, P. (2009). Forest cover change and illegal logging in the Ukrainian Carpathians in the transition period from 1988 to 2007. *Remote Sensing of Environment, 113*, 1194-1207
Kuenzer, C., Leinenkugel, P., Vollmuth, M., & Dech, S. (2014). Comparing global land-cover products – implications for geoscience applications: an investigation for the trans-boundary Mekong Basin. *International Journal of Remote Sensing, 35*, 2752-2779
Lambers, J.H.R. (2015). Extinction risks from climate change. *Science, 348*, 501-502
Lambin, E., Smelser, N., & Baltes, P. (2001). Remote sensing and GIS analysis. In (pp. 13150-13155): Pergamon: Oxford, UK
Lambin, E.F. (1997). Modelling and monitoring land-cover change processes in tropical regions. *Progress in Physical Geography, 21*, 375-393

Lambin, E.F., & Geist, H.J. (2003). Regional Differences in Tropical Deforestation. *Environment: Science and Policy for Sustainable Development, 45*, 22-36

Lambin, E.F., & Meyfroidt, P. (2011). Global land use change, economic globalization, and the looming land scarcity. *Proceedings of the National Academy of Sciences, 108*, 3465-3472

Lawrence, P.J., Feddema, J.J., Bonan, G.B., Meehl, G.A., O'Neill, B.C., Oleson, K.W., Levis, S., Lawrence, D.M., Kluzek, E., Lindsay, K., & Thornton, P.E. (2012). Simulating the Biogeochemical and Biogeophysical Impacts of Transient Land Cover Change and Wood Harvest in the Community Climate System Model (CCSM4) from 1850 to 2100. *Journal of Climate, 25*, 3071-3095

Le Houérou, H.N. (1996). Climate change, drought and desertification. *Journal of Arid Environments, 34*, 133-185

Leathwick, J.R., Elith, J., & Hastie, T. (2006). Comparative performance of generalized additive models and multivariate adaptive regression splines for statistical modelling of species distributions. *Ecological Modelling, 199*, 188-196

Leathwick, J.R., Rowe, D., Richardson, J., Elith, J., & Hastie, T. (2005). Using multivariate adaptive regression splines to predict the distributions of New Zealand's freshwater diadromous fish. *Freshwater Biology, 50*, 2034-2052

Lehouck, V., Spanhove, T., Colson, L., Adringa-Davis, A., Cordeiro, N.J., & Lens, L. (2009). Habitat disturbance reduces seed dispersal of a forest interior tree in a fragmented African cloud forest. *Oikos, 118*, 1023-1034

Leimu, R., Vergeer, P., Angeloni, F., & Ouborg, N.J. (2010). Habitat fragmentation, climate change, and inbreeding in plants. *Ann N Y Acad Sci, 1195*, 84-98

Leinenkugel, P., Kuenzer, C., Oppelt, N., & Dech, S. (2013). Characterisation of land surface phenology and land cover based on moderate resolution satellite data in cloud prone areas — A novel product for the Mekong Basin. *Remote Sensing of Environment, 136*, 180-198

Levey, D.J., Tewksbury, J.J., & Bolker, B.M. (2008). Modelling long-distance seed dispersal in heterogeneous landscapes. *Journal of Ecology, 96*, 599-608

Li, S., Gu, S., Tan, X., & Zhang, Q. (2009). Water quality in the upper Han River basin, China: The impacts of land use/land cover in riparian buffer zone. *Journal of Hazardous Materials, 165*, 317-324

Licciardi, G., Pacifici, F., Tuia, D., Prasad, S., West, T., Giacco, F., Thiel, C., Inglada, J., Christophe, E., Chanussot, J., & Gamba, P. (2009). Decision Fusion for the Classification of Hyperspectral Data: Outcome of the 2008 GRS-S Data Fusion Contest. *Geoscience and Remote Sensing, IEEE Transactions on, 47*, 3857-3865

Lind, B., Stein, S., Kärcher, A., & Klein, M. (2009). Where have All the Flowers Gone? Grünland im Umbruch. In. Bonn: German Federal Agency for Nature Conservation (BfN)

Liu, C., Berry, P.M., Dawson, T.P., & Pearson, R.G. (2005). Selecting thresholds of occurrence in the prediction of species distributions. *Ecography, 28*, 385-393

Liu, D., & Xia, F. (2010). Assessing object-based classification: advantages and limitations. *Remote Sensing Letters, 1*, 187-194

Liu, X., Li, X., Yeh, A.-O., He, J., & Tao, J. (2007). Discovery of transition rules for geographical cellular automata by using ant colony optimization. *Science in China Series D: Earth Sciences, 50*, 1578-1588

Loarie, S.R., Carter, B.E., Hayhoe, K., McMahon, S., Moe, R., Knight, C.A., & Ackerly, D.D. (2008). Climate Change and the Future of California's Endemic Flora. *PLoS ONE, 3*, e2502

Loeb, C. (2006). Planning reunification: the planning history of the fall of the Berlin Wall. *Planning Perspectives, 21*, 67-87

Loiselle, B.A., Howell, C.A., Graham, C.H., Goerck, J.M., Brooks, T., Smith, K.G., & Williams, P.H. (2003). Avoiding Pitfalls of Using Species Distribution Models in Conservation Planning

Evitando Dificultades Resultantes del Uso de Modelos de Distribución de Especies en Planeación de Conservación. *Conservation Biology, 17*, 1591-1600

Long, J.B., & Giri, C. (2011). Mapping the Philippines' Mangrove Forests Using Landsat Imagery. *Sensors, 11*, 2972

Loveland, T.R., & Belward, A.S. (1997). The IGBP-DIS global 1km land cover data set, DISCover: First results. *International Journal of Remote Sensing, 18*, 3289-3295

Lu, D., Hetrick, S., & Moran, E. (2011). Impervious surface mapping with Quickbird imagery. *International Journal of Remote Sensing, 32*, 2519-2533

Lu, D., Mausel, P., Brondizio, E., & Moran, E. (2004). Change detection techniques. *International Journal of Remote Sensing, 25*, 2365-2401

Lu, D., & Weng, Q. (2007). A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing, 28*, 823-870

Lu, N., Jia, C.-X., Lloyd, H., & Sun, Y.-H. (2012). Species-specific habitat fragmentation assessment, considering

the ecological niche requirements and dispersal capability. *Biological Conservation, 152*, 102-109
Macedo, M.N., Coe, M.T., DeFries, R., Uriarte, M., Brando, P.M., Neill, C., & Walker, W.S. (2013). Land-use-driven stream warming in southeastern Amazonia. *Philos Trans R Soc Lond B Biol Sci, 368*, 20120153
Maestre, F.T., Callaway, R.M., Valladares, F., & Lortie, C.J. (2009). Refining the stress-gradient hypothesis for competition and facilitation in plant communities. *Journal of Ecology, 97*, 199-205
Malcolm, J.R., Markham, A., Neilson, R.P., & Garaci, M. (2002). Estimated migration rates under scenarios of global climate change. *Journal of Biogeography, 29*, 835-849
Malczewski, J. (1999). *GIS and Multicriteria Decision Analysis*. New York: John Wiley & Sons
Manandhar, R., Odeh, I., & Ancev, T. (2009). Improving the Accuracy of Land Use and Land Cover Classification of Landsat Data Using Post-Classification Enhancement. *Remote Sensing, 1*, 330
Marage, D., Rameau, J.-C., & Garraud, L. (2006). Soil seed banks and vegetation succession in the Southern Alps: effects of historical and ecological factors. *Canadian Journal of Botany, 84*, 99-111
Marmion, M., Parviainen, M., Luoto, M., Heikkinen, R.K., & Thuiller, W. (2009). Evaluation of consensus methods in predictive species distribution modelling. *Diversity and Distributions, 15*, 59-69
Martin, Y., Van Dyck, H., Dendoncker, N., & Titeux, N. (2013). Testing instead of assuming the importance of land use change scenarios to model species distributions under climate change. *Global Ecology and Biogeography, 22*, 1204-1216
Martinez-Meyer, E. (2005). Climate Change and Biodiversity: Some Considerations in Forecasting Shifts in Species' Potential Distributions. *2005, 2*
McConkey, K.R., Prasad, S., Corlett, R.T., Campos-Arceiz, A., Brodie, J.F., Rogers, H., & Santamaria, L. (2012). Seed dispersal in changing landscapes. *Biological Conservation, 146*, 1-13
McKenney, D.W., Pedlar, J.H., Lawrence, K., Campbell, K., & Hutchinson, M.F. (2007). Potential Impacts of Climate Change on the Distribution of North American Trees. *BioScience, 57*, 939-948
Meier, E.S., Kienast, F., Pearman, P.B., Svenning, J.-C., Thuiller, W., Araújo, M.B., Guisan, A., & Zimmermann, N.E. (2010). Biotic and abiotic variables show little redundancy in explaining tree species distributions. *Ecography, 33*, 1038-1048
Meier, E.S., Lischke, H., Schmatz, D.R., & Zimmermann, N.E. (2012). Climate, competition and connectivity affect future migration and ranges of European trees. *Global Ecology and Biogeography, 21*, 164-178
Memmott, J., Craze, P.G., Waser, N.M., & Price, M.V. (2007). Global warming and the disruption of plant–pollinator interactions. *Ecology Letters, 10*, 710-717
Meynard, C.N., & Quinn, J.F. (2007). Predicting species distributions: a critical comparison of the most common statistical models using artificial species. *J. Biogeogr*, 34, 1455–1469
Midgley, G.F., Hughes, G.O., Thuiller, W., & Rebelo, A.G. (2006). Migration rate limitations on climate change-induced range shifts in Cape Proteaceae. *Diversity and Distributions, 12*, 555-562
Mitsova, D., Shuster, W., & Wang, X. (2011). A cellular automata model of land cover change to integrate urban growth with open space conservation. *Landscape and Urban Planning, 99*, 141-153
Möller, M., Lymburner, L., & Volk, M. (2007). The comparison index: A tool for assessing the accuracy of image segmentation. *International Journal of Applied Earth Observation and Geoinformation, 9*, 311-321
Morris, C.C., & Simon, T.P. (2012). *Nutrient Indicator Models for Determining Biologically Relevant Levels: A Case Study Based on the Corn Belt and Northern Great Plain Nutrient Ecoregion*. Springer
Muñoz, J., & Felicísimo, Á.M. (2004). Comparison of statistical methods commonly used in predictive modelling. *Journal of Vegetation Science, 15*, 285-292
Murphy, H.T., VanDerWal, J., & Lovett-Doust, J. (2010). Signatures of range expansion and erosion in eastern North American trees. *Ecology Letters, 13*, 1233-1244
Myint, S.W., Gober, P., Brazel, A., Grossman-Clarke, S., & Weng, Q. (2011). Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery. *Remote Sensing of Environment, 115*, 1145-1161
Nathan, R., & Muller-Landau, H.C. (2000). Spatial patterns of seed dispersal, their determinants and consequences for recruitment. *Trend in Ecology and Evolution, 15*, 278-285
Nathan, R. (2006). Long-Distance Dispersal of Plants. *Science, 313*, 786-788
Nathan, R., Katul, G., Bohrer, G., Kuparinen, A., Soons, M., Thompson, S., Trakhtenbrot, A., & Horn, H. (2011). Mechanistic models of seed dispersal by wind. *Theoretical Ecology, 4*, 113-132
Neill, W.V.J., & Schwedler, H.U. (2001). *Urban Planning and Cultural Inclusion: Lessons from Belfast and Berlin*. Palgrave Macmillan
Oliver, T.H., Marshall, H.H., Morecroft, M.D., Brereton, T., Prudhomme, C., & Huntingford, C. (2015). Interacting effects of climate change and habitat fragmentation on drought-sensitive butterflies. *Nature Clim. Change, 5*, 941-945

Opdam, P., & Wascher, D. (2004). Climate change meets habitat fragmentation: linking landscape and biogeographical scale levels in research and conservation. *Biological Conservation, 117*, 285-297

Oppel, S., Meirinho, A., Ramírez, I., Gardner, B., O'Connell, A.F., Miller, P.I., & Louzao, M. (2012). Comparison of five modelling techniques to predict the spatial distribution and abundance of seabirds. *Biological Conservation, 156*, 94-104

O'Sullivan, D., & Unwin, D.J. (2003). *Geographic Information Analysis*. Wiley

Otukei, J.R., & Blaschke, T. (2010). Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms. *International Journal of Applied Earth Observation and Geoinformation, 12*, S27-S31

Pal, M., & Mather, P.M. (2005). Support vector machines for classification in remote sensing. *International Journal of Remote Sensing, 26*, 1007-1011

Park, Y.S., Kwon, Y.S., Hwang, S.J., & Park, S. (2014). Characterizing effects of landscape and morphometric factors on water quality of reservoirs using a self-organizing map. *Environmental Modelling and Software, 55*, 214-221

Parmesan, C., & Yohe, G. (2003). A globally coherent fingerprint of climate change impacts across natural systems. *Nature, 421*, 37-42

Parviainen, M., Luoto, M., Ryttäri, T., & Heikkinen, R.K. (2008). Modelling the occurrence of threatened plant species in taiga landscapes: methodological and ecological perspectives. *Journal of Biogeography, 35*, 1888-1905

Pearson, R.G., & 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*, 361-371

Pearson, R.G., & Dawson, T.P. (2005). Long-distance plant dispersal and habitat fragmentation: identifying conservation targets for spatial landscape planning under climate change. *Biological Conservation, 123*, 389-401

Pearson, R.G., Dawson, T.P., Berry, P.M., & Harrison, P.A. (2002). SPECIES: A Spatial Evaluation of Climate Impact on the Envelope of Species. *Ecological Modelling, 154*, 289-300

Peterson, L.K., Bergen, K.M., Brown, D.G., Vashchuk, L., & Blam, Y. (2009). Forested land-cover patterns and trends over changing forest management eras in the Siberian Baikal region. *Forest Ecology and Management, 257*, 911-922

Petropoulos, G.P., Kalaitzidis, C., & Prasad Vadrevu, K. (2012). Support vector machines and object-based classification for obtaining land-use/cover cartography from Hyperion hyperspectral imagery. *Computers & Geosciences, 41*, 99-107

Petropoulos, G.P., Kontoes, C.C., & Keramitsoglou, I. (2012). Land cover mapping with emphasis to burnt area delineation using co-orbital ALI and Landsat TM imagery. *International Journal of Applied Earth Observation and Geoinformation, 18*, 344-355

Petropoulos, G.P., Vadrevu, K.P., Xanthopoulos, G., Karantounias, G., & Scholze, M. (2010). A Comparison of Spectral Angle Mapper and Artificial Neural Network Classifiers Combined with Landsat TM Imagery Analysis for Obtaining Burnt Area Mapping. *Sensors (Basel, Switzerland), 10*, 1967-1985

Pierce, D.W., Barnett, T.P., Santer, B.D., & Gleckler, P.J. (2009). Selecting global climate models for regional climate change studies. *Proceedings of the National Academy of Sciences, 106*, 8441-8446

Pijanowski, B.C., Brown, D.G., Shellito, B.A., & Manik, G.A. (2002). Using neural networks and GIS to forecast land use changes: a Land Transformation Model. *Computers, Environment and Urban Systems, 26*, 553-575

Pimm, S.L., & Raven, P. (2000). Biodiversity. Extinction by numbers. *Nature, 403*, 843-845

Pitelka, L.F., & Plant Migration Workshop, G. (1997). Plant Migration and Climate Change: A more realistic portrait of plant migration is essential to predicting biological responses to global warming in a world drastically altered by human activity. *American Scientist, 85*, 464-473

Poelmans, L., & Van Rompaey, A. (2010). Complexity and performance of urban expansion models. *Computers, Environment and Urban Systems, 34*, 17-27

Pompe, S., Hanspach, J., Badeck, F., Klotz, S., Thuiller, W., & Kühn, I. (2008). Climate and land use change impacts on plant distributions in Germany. *Biology Letters, 4*, 564-567

Pontius, G.R. (2000). quantification error versus location error in comparison of categorical maps. *photogrammetric engineering and remote sensing, 66*, 1011-1016

Poschlod, P., Bakker, J.P., & Kahmen, S. (2005). Changing land use and its impact on biodiversity. *Basic and Applied Ecology, 6*, 93-98

Potapov, P.V., Turubanova, S.A., Hansen, M.C., Adusei, B., Broich, M., Altstatt, A., Mane, L., & Justice, C.O. (2012). Quantifying forest cover loss in Democratic Republic of the Congo, 2000–2010, with Landsat ETM + data. *Remote Sensing of Environment, 122*, 106-116

Prakash, A., & Gupta, R. (1998). Land-use mapping and change detection in a coal mining area-a case study in the Jharia coalfield, India. *International Journal of Remote Sensing, 19*, 391-410

Prasad, A., Iverson, L., & Liaw, A. (2006). Newer Classification and Regression Tree Techniques: Bagging and Random Forests for Ecological Prediction. *Ecosystems, 9*, 181-199

Pu, R., Landry, S., & Yu, Q. (2011). Object-based urban detailed land cover classification with high spatial resolution IKONOS imagery. *International Journal of Remote Sensing, 32*, 3285-3308

Purves, D., & Pacala, S. (2008). Predictive Models of Forest Dynamics. *Science, 320*, 1452-1453

Qian, Y., Zhou, W., Yan, J., Li, W., & Han, L. (2014). Comparing Machine Learning Classifiers for Object-Based Land Cover Classification Using Very High Resolution Imagery. *Remote Sensing, 7*, 153

Quinlan, J.R. (1987). Simplifying decision trees. *Int. J. Man-Mach. Stud., 27*, 221-234

Quinlan, J.R. (2014). *C4.5: Programs for Machine Learning*. Elsevier Science

Radeloff, V.C., Nelson, E., Plantinga, A.J., Lewis, D.J., Helmers, D., Lawler, J.J., Withey, J.C., Beaudry, F., Martinuzzi, S., Butsic, V., Lonsdorf, E., White, D., & Polasky, S. (2012). Economic-based projections of future land use in the conterminous United States under alternative policy scenarios. *Ecological Applications, 22*, 1036-1049

Rai, S.C., & Sharma, E. (1998). Comparative assessment of runoff characteristics under different land use patterns within a Himalayan watershed. *Hydrological Processes, 12*, 2235-2248

Ramankutty, N., & Foley, J.A. (1998). Characterizing patterns of global land use: An analysis of global croplands data. *Global Biogeochemical Cycles, 12*, 667-685

Ready, J., Kaschner, K., South, A.B., Eastwood, P.D., Rees, T., Rius, J., Agbayani, E., Kullander, S., & Froese, R. (2010). Predicting the distributions of marine organisms at the global scale. *Ecological Modelling, 221*, 467-478

Reichler, T., & Kim, J. (2008). How Well Do Coupled Models Simulate Today's Climate? *Bulletin of the American Meteorological Society, 89*, 303-311

Ridd, M.K., & Liu, J. (1998). A comparison of four algorithms for change detection in an urban environment. *Remote Sensing of Environment, 63*, 95-100

Riecken, U., Finck, P., Raths, U., Schröder, E., & Ssymank, A. (2008). Die Gefährdung der Biotoptypen in Deutschland Aktueller Stand nach Vorlage der 2. *Fassung der Roten. Liste Natursch. Biol. Vielf.,, 2008*, 189–194

Robertson, M.P., Peter, C.I., Villet, M.H., & Ripley, B.S. (2003). Comparing models for predicting species' potential distributions: a case study using correlative and mechanistic predictive modelling techniques. *Ecological Modelling, 164*, 153-167

Robinson, D.J., Redding, N.J., & Crisp, D.J. (2002). Implementation of a fast algorithm for segmenting SAR imagery. In. Australia: Defense Science and Technology Organization

Robson, B.J. (2014). State of the art in modelling of phosphorus in aquatic systems: Review, criticisms and commentary. *Environmental Modelling & Software, 61*, 339-359

Röder, A., Hill, J., Duguy, B., Alloza, J.A., & Vallejo, R. (2008). Using long time series of Landsat data to monitor fire events and post-fire dynamics and identify driving factors. A case study in the Ayora region (eastern Spain). *Remote Sensing of Environment, 112*, 259-273

Rodríguez, J.P., Brotons, L., Bustamante, J., & Seoane, J. (2007). The application of predictive modelling of species distribution to biodiversity conservation. *Diversity and Distributions, 13*, 243-251

Rodríguez-Cabal, M.A., Aizen, M.A., & Novaro, A.J. (2007). Habitat fragmentation disrupts a plant-disperser mutualism in the temperate forest of South America. *Biological Conservation, 139*, 195-202

Rodriguez-Galiano, V.F., & Chica-Rivas, M. (2012). Evaluation of different machine learning methods for land cover mapping of a Mediterranean area using multi-seasonal Landsat images and Digital Terrain Models. *International Journal of Digital Earth, 7*, 492-509

Rodriguez-Galiano, V.F., Ghimire, B., Rogan, J., Chica-Olmo, M., & Rigol-Sanchez, J.P. (2012). An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing, 67*, 93-104

Rogan, J., & Chen, D. (2004). Remote sensing technology for mapping and monitoring land-cover and land-use change. *Progress in planning, 61*, 301-325

Rogan, J., Miller, J., Stow, D., Franklin, J., Levien, L., & Fischer, C. (2003). Land-cover change monitoring with classification trees using Landsat TM and ancillary data. *Photogrammetric Engineering & Remote Sensing, 69*, 793-804

Sala, O.E., Chapin, F.S., 3rd, Armesto, J.J., Berlow, E., Bloomfield, J., Dirzo, R., Huber-Sanwald, E., Huenneke, L.F., Jackson, R.B., Kinzig, A., Leemans, R., Lodge, D.M., Mooney, H.A., Oesterheld, M., Poff, N.L., Sykes, M.T., Walker, B.H., Walker, M., & Wall, D.H. (2000). Global biodiversity scenarios for the year 2100. *Science, 287*, 1770-1774

Salzberg, S. (1994). C4.5: Programs for Machine Learning by J. Ross Quinlan. Morgan Kaufmann Publishers, Inc., 1993. *Machine Learning, 16*, 235-240

Sang, L., Zhang, C., Yang, J., Zhu, D., & Yun, W. (2011). Simulation of land use spatial pattern of towns and villages based on CA–Markov model. *Mathematical and Computer Modelling, 54*, 938-943

Segurado, P. & Araújo, M.B. (2004). An evaluation of methods for modelling species distributions. *Journal of Biogeography, 31*, 1555–1568

Sidle, R.C., Ziegler, A.D., Negishi, J.N., Nik, A.R., Siew, R., & Turkelboom, F. (2006). Erosion processes in steep terrain - Truths, myths, and uncertainties related to forest management in Southeast Asia. *Forest Ecology and Management, 224*, 199-225

Siehoff, S., Lennartz, G., Heilburg, I.C., Roß-Nickoll, M., Ratte, H.T., & Preuss, T.G. (2011). Process-based modeling of grassland dynamics built on ecological indicator values for land use. *Ecological Modelling, 222*, 3854-3868

Singh, A. (1989). Review article digital change detection techniques using remotely-sensed data. *International Journal of Remote Sensing, 10*, 989-1003

Snyder, P.K., Delire, C., & Foley, J.A. (2004). Evaluating the influence of different vegetation biomes on the global climate. *Climate Dynamics, 23*, 279-302

Soons, M.B., & Ozinga, W.A. (2005). How important is long-distance seed dispersal for the regional survival of plant species? *Diversity and Distributions, 11*, 165-172

Stanton, J.C., Pearson, R.G., Horning, N., Ersts, P., & Reşit Akçakaya, H. (2012). Combining static and dynamic variables in species distribution models under climate change. *Methods in Ecology and Evolution, 3*, 349-357

Stuckens, J., Coppin, P.R., & Bauer, M.E. (2000). Integrating Contextual Information with per-Pixel Classification for Improved Land Cover Classification. *Remote Sensing of Environment, 71*, 282-296

Subedi, P., Subedi, K., & Thapa, B. (2013). Application of a Hybrid Cellular Automaton C Markov (CA-Markov) Model in Land-Use Change Prediction: A Case Study of Saddle Creek Drainage Basin, Florida. *Applied Ecology and Environmental Sciences, 1*, 126-132

Sun, H., Forsythe, W., & Waters, N. (2007). Modeling Urban Land Use Change and Urban Sprawl: Calgary, Alberta, Canada. *Networks and Spatial Economics, 7*, 353-376

Suttle, K.B., Thomsen, M.A., & Power, M.E. (2007). Species Interactions Reverse Grassland Responses to Changing Climate. *Science, 315*, 640-642

Swets, J. (1988). Measuring the accuracy of diagnostic systems. *Science, 240*, 1285-1293

Tackenberg, O., Poschlod, P., & Bonn, S. (2003). Assessment of Wind Dispersal Potential in Plant Species. *Ecological Monographs, 73*, 191-205

Tarkesh, M., & Jetschke, G. (2012). Comparison of six correlative models in predictive vegetation mapping on a local scale. *Environmental and Ecological Statistics, 19*, 437-457

Taylor, K.E., Stouffer, R.J., & Meehl, G.A. (2011). An Overview of CMIP5 and the Experiment Design. *Bulletin of the American Meteorological Society, 93*, 485-498

Team, R.D.C. (2009). R: A Language and Environment for Statistical Computing. . In. Vienna, Austria: R Foundation for Statistical Computing

Therneau, T., & Atkinson, E. (1997). An introduction to recursive partitioning using the RPART routines. Rochester, MN: Mayo Clinic.

Thomas, C.D., Cameron, A., Green, R.E., Bakkenes, M., Beaumont, L.J., Collingham, Y.C., Erasmus, B.F.N., de Siqueira, M.F., Grainger, A., Hannah, L., Hughes, L., Huntley, B., van Jaarsveld, A.S., Midgley, G.F., Miles, L., Ortega-Huerta, M.A., Townsend Peterson, A., Phillips, O.L., & Williams, S.E. (2004). Extinction risk from climate change. *Nature, 427*, 145-148

Thuiller, W. (2004). Patterns and uncertainties of species' range shifts under climate change. *Global Change Biology, 10*, 2020-2027

Thuiller, W., Georges, D., & Engler, R. (2013). biomod2: Ensemble platform for species distribution modelling. In

Thuiller, W., Lavorel, S., Sykes, M.T., & Araújo, M.B. (2006). Using niche-based modelling to assess the impact of climate change on tree functional diversity in Europe. *Diversity and Distributions, 12*, 49-60

Thuiller, W., Richardson, D.M., PyŠEk, P., Midgley, G.F., Hughes, G.O., & Rouget, M. (2005). Niche-based modelling as a tool for predicting the risk of alien plant invasions at a global scale. *Global Change Biology, 11*, 2234-2250

Tian, F., Yang, L., Lv, F., & Zhou, P. (2009). Predicting liquid chromatographic retention times of peptides from the Drosophila melanogaster proteome by machine learning approaches. *Anal Chim Acta, 644*, 10-16

Tingley, M.W., Monahan, W.B., Beissinger, S.R., & Moritz, C. (2009). Birds track their Grinnellian niche through a century of climate change. *Proceedings of the National Academy of Sciences, 106*, 19637-19643

Tölle, A. (2010). Urban identity policies in Berlin: From critical reconstruction to reconstructing the Wall. *Cities, 27*, 348-357

Townshend, J., Justice, C., Li, W., Gurney, C., & McManus, J. (1991). Global land cover classification by remote sensing: present capabilities and future possibilities. *Remote Sensing of Environment, 35*, 243-255

Tsoar, A., Allouche, O., Steinitz, O., Rotem, D., & Kadmon, R. (2007). A comparative evaluation of presence-only methods for modelling species distribution. *Diversity and Distributions, 13*, 397-405

Urban, M.C., Tewksbury, J.J., & Sheldon, K.S. (2012). On a collision course: competition and dispersal differences create no-analogue communities and cause extinctions during climate change. *Proceedings of the Royal Society of London B: Biological Sciences*

Uriarte, M., Anciães, M., da Silva, M.T.B., Rubim, P., Johnson, E., & Bruna, E.M. (2011). Disentangling the drivers of reduced long-distance seed dispersal by birds in an experimentally fragmented landscape. *Ecology, 92*, 924-937

USGS (2013). Landsat processing details. In

Van Coillie, F.M.B., Verbeke, L.P.C., & De Wulf, R.R. (2007). Feature selection by genetic algorithms in object-based classification of IKONOS imagery for forest mapping in Flanders, Belgium. *Remote Sensing of Environment, 110*, 476-487

Vapnik, V.N. (1995). *The Nature of Statistical Learning Theory*. Springer

Vitousek, P.M. (1994). Beyond Global Warming: Ecology and Global Change. *Ecology, 75*, 1861-1876

Vittoz, P., & Engler, R. (2007). Seed dispersal distances: a typological system for data analyses and models. *Botanica Helvetica, 117*, 109-124

Vittoz, P., Randin, C., Dutoit, A., Bonnet, F., & Hegg, O. (2009). Low impact of climate change on subalpine grasslands in the Swiss Northern Alps. *Global Change Biology, 15*, 209-220

Walsh, S.J., McCleary, A.L., Mena, C.F., Shao, Y., Tuttle, J.P., González, A., & Atkinson, R. (2008). QuickBird and Hyperion data analysis of an invasive plant species in the Galapagos Islands of Ecuador: Implications for control and land use management. *Remote Sensing of Environment, 112*, 1927-1941

Waske, B., van der Linden, S., Oldenburg, C., Jakimow, B., Rabe, A., & Hostert, P. (2012). imageRF – A user-oriented implementation for remote sensing image analysis with Random Forests. *Environmental Modelling & Software, 35*, 192-193

Wen, L., Dong , S., Li, Y., Li, X., Shi, J., Wang, Y., Liu, D., & Ma, Y. (2013). Effect of Degradation Intensity on Grassland Ecosystem Services in the Alpine Region of Qinghai-Tibetan Plateau, China. *PLoS ONE, 8*, e58432

White, R., & Engelen, G. (1994). Cellular Dynamics and GIS: Modelling Spatial Complexity. *Journal of Geographical Systems*

Wiens, J.A., Stralberg, D., Jongsomjit, D., Howell, C.A., & Snyder, M.A. (2009). Niches, models, and climate change: Assessing the assumptions and uncertainties. *Proceedings of the National Academy of Sciences of the United States of America, 106*, 19729-19736

Williams, J.N., Seo, C., Thornes, J., Nelson, J.K., Erwin, S., O'Brien, J.M., & Schwartz, M.W. (2009). Using species distribution models to predict new occurrences for rare plants. *Diversity and Distributions*, 15, 565–576

Willson, M.F., & Traveset, A. (2000). The ecology of seed dispersal. In M. Fenner (Ed.), *Seeds: the ecology of regeneration in plant communities* (pp. 85–110). Wallingford: CABI Publishing

Wilson, J.P., & Gallant, J.C. (2000). *Terrain Analysis: Principles and Applications*. Wiley

Yang, X., Zheng, X.-Q., & Chen, R. (2014). A land use change model: Integrating landscape pattern indexes and Markov-CA. *Ecological Modelling, 283*, 1-7

Yang, X., Zheng, X.-Q., & Lv, L.-N. (2012). A spatiotemporal model of land use change based on ant colony optimization, Markov chain and cellular automata. *Ecological Modelling, 233*, 11-19

Yee, T.W., & Mitchell, N.D. (1991). Generalized Additive Models in plant ecology. *Journal of Vegetation Science*, 2, 587–602

Ying, X., Zeng, G.-M., Chen, G.-Q., Tang, L., Wang, K.-L., & Huang, D.-Y. (2007). Combining AHP with GIS in synthetic evaluation of eco-environment quality—A case study of Hunan Province, China. *Ecological Modelling, 209*, 97-109

Yu, Q., Gong, P., CLINTON, N., BIGING, G., KELLY, M., & SCHIROKAUER, D. (2006). Objectbased detailed vegetation classification with airborne high spatial resolution remote sensing imagery. *photogrammetric engineering and remote sensing*, 799–811

Yuan, D., Elvidge, C.D., & Lunetta, R.S. (1998). Survey of multispectral methods for land cover change analysis

Yuan, F., Sawaya, K.E., Loeffelholz, B.C., & Bauer, M.E. (2005). Land cover classification and change analysis of the Twin Cities (Minnesota) Metropolitan Area by multitemporal Landsat remote sensing. *Remote Sensing of Environment, 98*, 317-328

Zhang, Q., & Zhang, X. (2012). Impacts of predictor variables and species models on simulating Tamarix

ramosissima distribution in Tarim Basin, northwestern China. *Journal of Plant Ecology*

Zhang, Y., Zhang, J., Zhang, X., Wu, H., & Guo, M. (2015). Land Cover Classification from Polarimetric SAR Data Based on Image Segmentation and Decision Trees. *Canadian Journal of Remote Sensing, 41*, 40-50

Zimmermann, N.E., Edwards, T.C., Graham, C.H., Pearman, P.B., & Svenning, J.-C. (2010). New trends in species distribution modelling. *Ecography, 33*, 985-989

Zimmermann, N.E., Edwards, T.C., Moisen, G.G., Frescino, T.S., & Blackard, J.A. (2007). Remote sensing-based predictors improve distribution models of rare, early successional and broadleaf tree species in Utah. *Journal of Applied Ecology, 44*, 1057-1067

Zimmermann, N.E., & Kienast, F. (1999). Predictive mapping of alpine grasslands in Switzerland: Species versus community approach. *Journal of Vegetation Science, 10*, 469-482

Zurell, D., Jeltsch, F., Dormann, C.F., & Schröder, B. (2009). Static species distribution models in dynamically changing systems: how good can predictions really be? *Ecography, 32*, 733-744

Appendix

Appendix A The theoretical basics of cellular automata and Markov chain models are explained in this section.

1. Cellular Automata

Cellular automaton (CA) is a mathematical model used for system simulations and calculations. Cellular automata are discrete simple systems, which can display complicated behavior and calculations using simple local rules. The "local" part implies that in determining the new value of each cell, adjacent cells (the neighborhood) are influential while distant cells are not. Each cell has a set of states, and the cell momentarily switches between the states depending on its states and those of its neighbors. States of all cells are updated simultaneously. The overall system behavior is determined based on the combined effects of all of the local transition rules. Therefore, the system's state evolves throughout discrete time steps (Liu, 2009; Wolfram, 1984).
Based on the above definitions, the CA model consists of the following five basic components:

Cell

The CA model is defined as cellular units within a raster network. The cells may exist in either a two- or three dimensional spatial tessellations. Even though CA models often have a two-dimensional structure, other structures, such as one dimensional ones, are also employed for solving linear problems. The cellular space can also be designed as a three-dimensional grid (lattice) (Nagel et al., 1997).

State

A state is defined as a system attribute. At any given time, each cell has a specific state which can change to another state during a unit of time. A cell's state can reflect different land uses such as forest or urban uses (Portugali et al., 1995).

Neighborhood

A neighborhood is composed of a cell and other cells situated at a radial distance from the cell of concern. The cell changes state based on its mutual relations with its neighbor as a result of transition rules. Size of a neighborhood varies for different CA model. In a one-dimensional CA, each cell has three neighboring cells. The most well-known neighborhood in two-dimensional CA is the Moore neighborhood, which consists of nine cells arranged in the 3×3 order, and the Von neighborhood consists of five cells (Rezazadeh et al., 2010). The von Neumann and Moore neighborhoods are shown in Fig. 1.

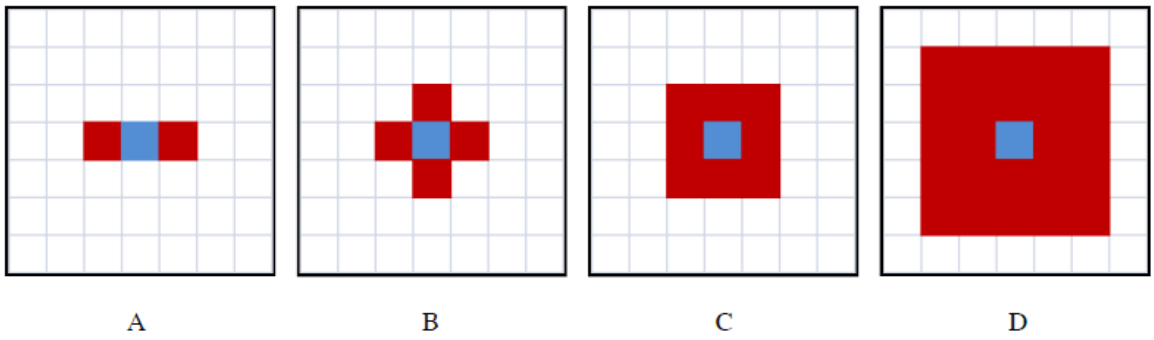


Fig. 1 CA neighborhoods. A: Neighborhood in one-dimensional CA. B: von Neumann Neighborhood. C: 3×3 Moore Neighborhood. D: 5×5 Moore Neighborhood (Rezazadeh et al., 2010).

Transition rules

The transition rule determines the change of a cell's state according to its current state and its adjacent cells. This element is one of the main parts of CA, because these rules reflect the system modeling and are necessary for success of modeling (White, 1998).

Time

A CA model is defined in a time dimension; as mentioned earlier, cells can change from one state to another during a unit of time. In other words, all the cells within a network are updated simultaneously, and they either change according to the rules of the model or remain unchanged until the next time unit (White, 1994).

A general example

In a one-dimensional CA, cells form a circle, with each cell having two neighbors and two states. In the T time step, the cell's state depends on its own state and state of its neighbors in the T-1 time step. For example, if the target cell's state is shown by A in the T-1 time step and states of its two neighbors are shown by B, and if the transition rules are designed to transform the target cell to B, then the rule can be written as follows: BAB=B

Each cell can generate 8 different combinations with its neighboring cells: BBB, BBA, BAB, ABB, BAA, AAB, ABA, and AAA.

Imagine we have the following transition rules: (AAA→ B), (BBB→ B), (BBA→ A), (BAB→ A), (ABB→ B), (BAA→ A), (AAB→ A), (ABA→ A). Transition rules determine the target cell's state in each of the composition of the next step. In this case, with AAAABAAAA as the state, this line will be written as BBBAAABBB in the next step (the first and last cells are linked as the 9 cells form a circle). This is because the middle cell B follows the ABA $\rightarrow$A rule. The cell on the right center follows the BAA $\rightarrow$A rule, whereas the left cell follows the AAB $\rightarrow$A transition rule.

Other cells follow the AAA $\rightarrow$B rule for state transition. Using the same rules, the following combination is obtained in the third phase: BBAABABBB.

Game of Life

The "Game of Life" is another example which can be mentioned regarding CA models. This model was designed by John Conway (Gardner, 1972). In this game, cells exist in a two-dimensional network, and they can assume two states: "Dead" or "Alive". Moreover, in this game, the state of the target cell is determined based on the state of its Moore Neighborhood – i.e. its eight neighboring cells.

The rules of transformation in a CA model are a set of *If-then* rules. For examples, a set of three rules are used in the "Game of Life". In this game, cells can remain alive, die, or be born in the next time unit based on the rules. The first rule dictates that if a cell is neighboring two or three alive cells, then it will remain alive. The second rule holds that if a cell is placed next to less than two or more than three alive cells, then it will die because of seclusion or overpopulation in the next step. The third rule dictates that if a dead cell has exactly three alive cells in its immediate surroundings it will become alive. Fig. 2 depicts an example of simulations by this model.

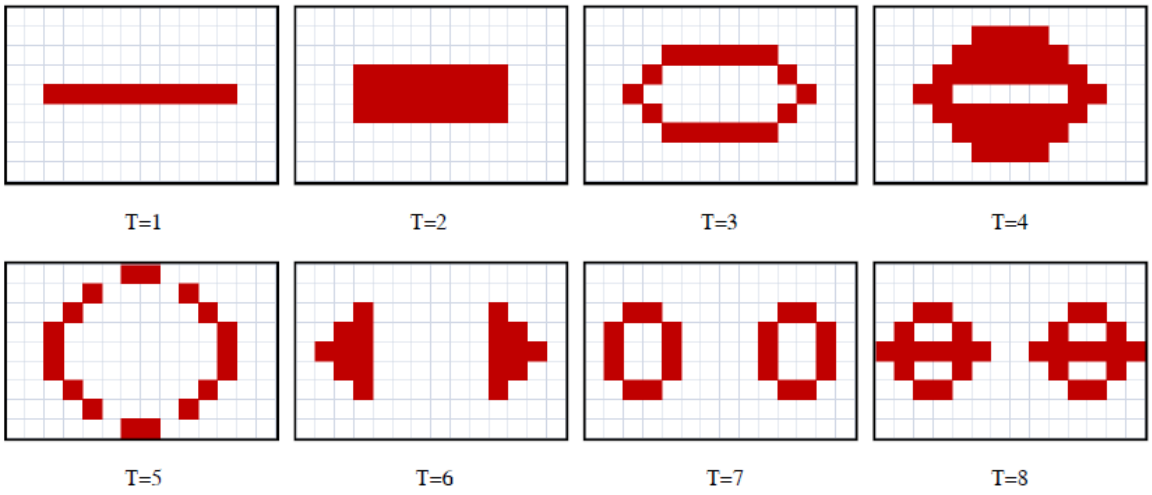


Fig. 2 A simple simulation based on "Game of Life" (brown cells are alive, and white cells are dead; T also shows the time step)

2. Markov Chain

A random process is a set of random variables shown by {X(n), n€N}, in which the possible values of $X_{(n)}$ form the process state. In practice, the N subscript set refers to time, and values of $X_{(n)}$ normally show sizes or observations of a system at time n. Markov chain is a random process in which random variables transfer from one state to another. A Markov property indicates that the subsequent state of a variable only depends on its current state and is independent of previous events (Higgins and Keller-McNulty, 1995).

Assume {X(n)} is a random process and {X(n)} shows different states of the random process. Therefore, the Markov property applies to the process, if:

$$P[X(n+1)= S(n+1) \mid X(n)= S(n), X(n-1)= S(n-1),\ldots,X(0)= S(0)]= P[X(n+1)= S(n+1) \mid X(n) = S(n)] \quad (1)$$

The random process to which the Markov property applies is called a Markov process, and processes with discrete states form a Markov chain. Fig. 3 depicts a Markov property. As seen, each cell only depends on a precedent cell and is independent of other cells (Papoulis, 2002). The probability of a one-stage transition is defined as follows based on conditional probabilities:

$$P(i \rightarrow j)=P[X=j \mid X-1=i) \quad (2)$$

The probability of a one-stage transition has a key role in the theory and application of Markov chains and it is equal to the probability of transition for state i to state j. Table 1 presents an example of the transition probability matrix for three states. In this matrix, elements of the transition matrix correspond to probabilities of transition from state i to j. Dimensions of this matrix are defined based on possible states of a random variable. In addition, the row and column indices correspond to the current (i) and subsequent states (j), respectively (Papoulis, 2002). In Table 1, the element on the third row and second column indicates that probability of transition from state 3 to the next state is 0.36, if the next state is 2. Evidently, elements on a row sum up to one.

Table 2 An example of transition probability matrix for three states (Papoulis, 2002)

State	A	B	C
A	0.65	0.28	0.07
B	0.15	0.67	0.18
C	0.12	0.36	0.52

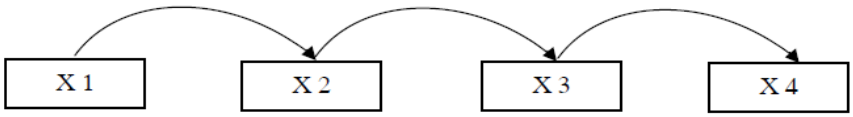


Fig. 3 A schematic view of a Markov property

In Fig. 4, there is a sequence of events with the Markov property. Cells x, x-1, and N are in states S_k, S_1 and S_q, respectively. Occurrence of the event S_k also depends on events S_1 and S_q (Z_i denotes the state of cell x=1,…, N). Probability of a state transition (pr) from state i to j is expressed as follows (Li and Zhang, 2008):

$$Pr(Z_i=S_k \mid Z_{i-1}=S_1, Z_n=S_q) \quad (3)$$

Relation (3) can be written as follows:

$$Pr(Z_i=S_k \mid Z_{i-1}=S_1, Z_n=S_q)=\frac{pr(Z_{i-1}=S_1, Z_1=S_k, Z_n=S_q)}{pr(Z_{i-1}=S_1, Z_n=S_q)} \quad (4)$$

Considering the Markov chain, conditional probabilities of equation (4) can be expressed as follows:

$$Pr(Z_i=S_k \mid Z_{x-1}=S_1, Z_n=S_q)=\frac{P_{kq}^{(N-x)}Plk}{P_{lq}^{(N-x+1)}} \quad (5)$$

Where, $P_{kq}^{(N-x)}$ shows probability of transition from state k to q with a (N-x)-cell interval, and $P_{lq}^{(N-x+1)}$ is probability of transition from state l to q with a (N-x+1)-cell interval (Li and Zhang, 2008; Nikoogoftar et al., 2015). By assuming an initial value as the observed space and a one-stage transition matrix, the movement paths of the Markov chain can be simulated.

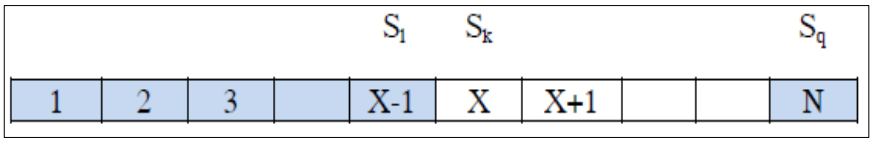


Fig. 4 A sequence of events for the subsequent Markov chain (values of the blue cells are known and values of the white cells are estimated by the Markov method) (Nikoogoftar et al., 2015).

References

Gardner, M. (1972). The fantastic combinations of John Conway's new solitaire game 'Life'. *Scientific American*, 233, 41.

Higgins, J., & Keller-McNulty, S. (1995). Concepts in Probability and Stochastic Modeling, Duxbury Press; first edition.

Li, W., & Zhang, C. (2008). A single-chain-based multidimensional Markov chain model for subsurface characterization. *Environmental and Ecological Statistics*, 15, 157–174.

Liu, Y. (2009). Modelling Urban Development with Geographical Information Systems and Cellular Automata. New York: Taylor & Francis, 25-51

Nagel, K., Rasmussen, S., & Barrett, C. (1997). Network traffic as a self-organized critical phenomenon. In Self-organization of complex structures: from individual to collective dynamics. New York: CRC Press: 414

Nikoogoftar, H., Bahroodi, A., Tokhmchi, B., Norouzi, H., & Mehrgini, B. (2015). Modeling with Markov Chains methodology, case study: Lithofacies in a hydrocarbon reservoir, SW Iran. *Geosciences*, 94, 97-102

Papoulis, A. (2002) Probability, random variables and stochastic, McGraw-Hill Europe, 97 p.

Portugali, J., & Benenson, I. (1995). Artificial planning experience by means of a heuristic cell-space model: simulating international migration in the urban process. *Environment and Planning*, 27, 963

Rezazadeh, R., & Mirahmadi, M. (2010). Cellular automation model, a new method to simulate urban development. Fannavari-e-Amoozesh, 4th time, 4, 48-55. [In Persian]

White, R., & Engelen, G. (1994). Cellular dynamics and GIS: modelling spatial complexity. *Geographical Systems*, 1, 53-273

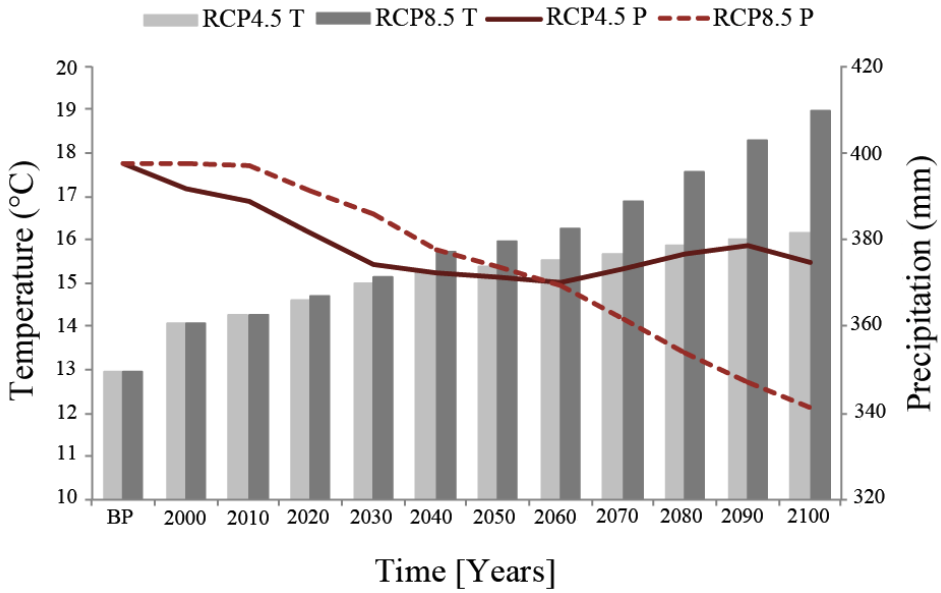
White, R. (1998). Cities and cellular automata. Discrete Dynamics in Nature and Society, 2, 32

Wolfram, S. (1984). Cellular automata as models of complexity. *Nature*, 311, 61-67

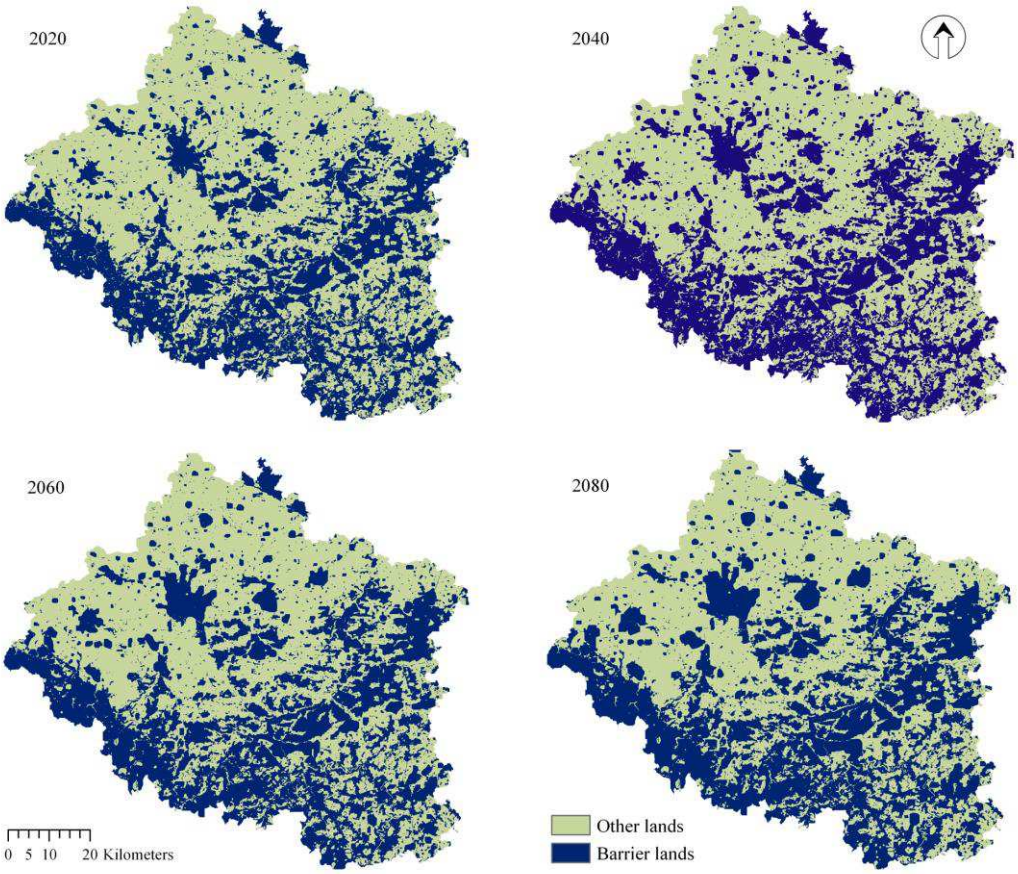
Appendix B List of IPCC (AR5) GCMs used in this study with a brief indication of their origin, and resolution for each climate change scenario.

Modeling Center	Institution ID	Model	Resolution (lon by lat)
Beijing Climate Center, China	BCC	BCC-CSM1-1	2.815 × 2.815
Beijing Climate Center, China	BCC	BCC-CSM1-1-M	1.12 × 1.12
Commonwealth Scientific and Industrial Research, Australia	CSIRO	CSIRO-Mk3.6	1.875 × 1.875
The First Institute of Oceanography, SOA, China	FIO	FIO-ESM	2.8 × 2.8
NOAA Geophysical Fluid Dynamics Laboratory	NOAA GFDL	GFDL-CM3	2.5 × 2.0
NOAA Geophysical Fluid Dynamics Laboratory	NOAA GFDL	GFDL-ESM2G	2.5 × 2.0
NOAA Geophysical Fluid Dynamics Laboratory	NOAA GFDL	GFDL-ESM2M	2.5 × 2.0
NASA Goddard Institute for Space Studies, USA	NASA GISS	GISS-E2-H	2.5 × 2.0
NASA Goddard Institute for Space Studies, USA	NASA GISS	GISS-E2-R	2.5 × 2.0
Met Office Hadley Centre, UK	MOHC	HadGEM2-ES	1.875 × 1.25
Institut Pierre-Simon Laplace, France	IPSL	IPSL-CM5A-LR	3.75 × 1.875
Institut Pierre-Simon Laplace, France	IPSL	IPSL-CM5A-MR	2.5 × 1.25
Atmosphere and Ocean Research Institute, Japan	MIROC	MIROC5	1.40 × 1.40
Japan Agency for Marine–Earth Science and Technology, Japan	MIROC	MIROC-ESM	2.815 × 2.815
Japan Agency for Marine–Earth Science and Technology, Japan	MIROC	MIROC-ESM-CHEM	2.815 × 2.815
Meteorological Research Institute, Japan	MRI	MRI-CGCM3	1.125 × 1.125
Norwegian Climate Centre, Norway	NCC	NorESM1-M	2.5 × 1.875

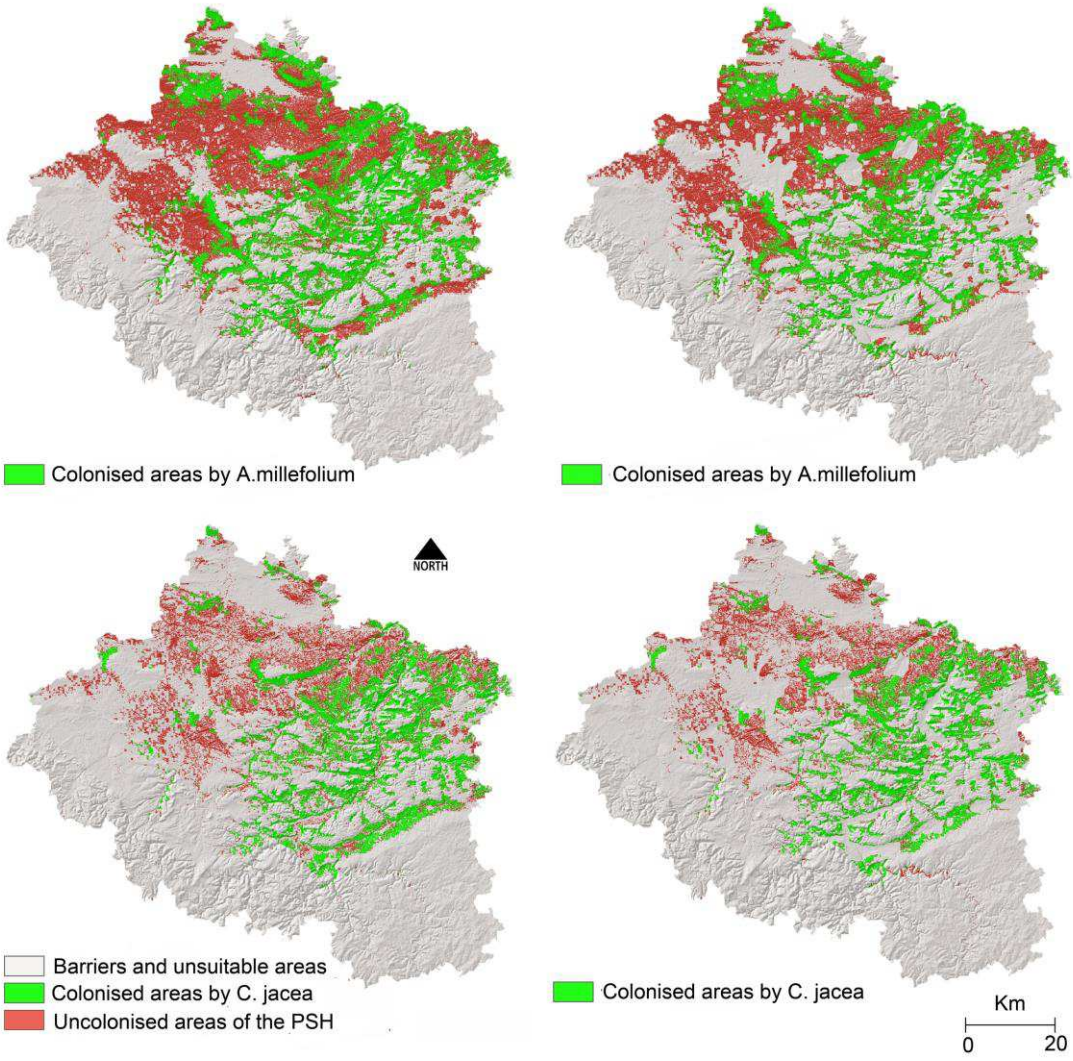
Appendix C Projected changes of precipitation and temperature under RCP4.5 and RCP8.5 scenarios. RCP4.5T and RCP8.5T show the summer mean temperature under medium and high emission scenarios, respectively. RCP4.5P and RCP8.5T are the summer total precipitations correspond respectively to medium and high emission scenarios. BP was obtained by averaging the baseline period (1960-1990).



Appendix D Time series of simulated barrier maps for 2020-2080. Other land contains farm land, grassland and water bodies, and barrier land contains forest and built-up areas.



Appendix E Examples of distribution maps obtained for year 2100 with RCP8.5 scenario and LDD events for *A. millefolium* and *C. jacea* based on static landscape (SLS) scenario (left) and dynamic landscape (DLS) scenario (right). PSH= potentially suitable habitat.



Acknowledgments

There are a great many people who have helped support me throughout the time taken to complete this PhD. I firstly want to thank my supervisors Winfried Voigt and Steffen Halle for giving me the opportunity to complete this PhD, and for providing all of the support and patience I required to make this such an enjoyable and rewarding experience. I am also grateful to Dr. Gottfried Jetsk who discussed the ecological side of the topic with me, and Dr. Christiane Roscher for providing species data.

I am grateful to the University of Jena and Iranian Ministry of Science, Research and Technology (MSRT) for financial support. I would like to thank the Thüringer Landesanstalt für Umwelt und Geologie, Jena, Germany for providing digital data, and also the United States Geological Survey (USGS) and European Space Agency (ESA) for preparing free archive of Landsat Earth-observing satellites images. I am also grateful to the Center of Geospatial at the University of Stanford.

I am deeply grateful to all my friends and colleagues from the institute of Ecology for their helpful comments, critiques and discussions about species distribution modeling and many other topics. I would like to thanks Dr. Robin Engler and Dr. Pascal Vittoz for helpful advices regarding plant distribution modeling.

I also want to thank both of my parents and the rest of my family, for taking an interest in what must appear to be a rather obscure choice of career. And to my Goli and my little son Alireza, I thank you for all your love and support over what has been an often tumultuous period in my life. Your constant encouragement to pursue my ambitions has done more to inspire me to succeed than I could honestly have thought possible.

Declaration

The material presented in this thesis is the result of original research, conducted between September 2011 and October 2015, under the supervision of Prof. Stefan Halle and Dr. Winfried Voigt. It has not been previously submitted, in part or whole, to any university of institution for any degree, diploma, or other qualification. Any material not of my own is acknowledged in the text.

Hamidreza Keshtkar Jena, 12.05.2016 _____