ISTANBUL TECHNICAL UNIVERSITY ★ INSTITUTE OF INFORMATICS

DISTRIBUTION OF NORTH AMERICAN TREE SPECIES UNDER CLIMATIC CHANGE: AN ECOLOGICAL NICHE MODELING STUDY USING ARTIFICIAL NEURAL NETWORKS

M.Sc. Thesis by

Hasan Serhan AKIN

Department : **Informatics**

Programme: **Computational Science & Engineering**

JUNE 2007

ISTANBUL TECHNICAL UNIVERSITY ★ INSTITUTE OF INFORMATICS

DISTRIBUTION OF NORTH AMERICAN TREE SPECIES UNDER CLIMATIC CHANGE: AN ECOLOGICAL NICHE MODELING STUDY USING ARTIFICIAL NEURAL NETWORKS

M.Sc. Thesis by

Hasan Serhan AKIN

702041009

Date of Submission : 7 May 2007

Date of defence examination: 11 June 2007

Supervisor(Chairman): Prof. Dr. Hasan Nüzhet Dalfes

Members of the Examining Committee: Assist. Prof. Dr. Neslihan Serap Şengör

Assist. Prof. Dr. Alper Tunga Akarsubaşı

JUNE 2007

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ BİLİŞİM ENSTİTÜSÜ

İKLİMSEL DEĞİŞİKLİKLERİN KUZEY AMERİKA AĞAÇ TÜRLERİNİN DAĞILIMI ÜZERİNE ETKİLERİNİN TAHMİNİ: YAPAY SİNİR AĞLARI YAKLAŞIMIYLA EKOLOJİK NİŞ MODELLEME ÇALIŞMASI

YÜKSEK LİSANS TEZİ

Hasan Serhan AKIN

702041009

Tezin Enstitüye Verildiği Tarih : 7 Mayıs 2007

Tezin Savunulduğu Tarih : 11 Haziran 2007

Tez Danışmanı : Prof. Dr. Hasan Nüzhet Dalfes

Diğer Jüri Üyeleri : Yrd. Doç. Dr. Neslihan Serap Şengör

Yrd. Doç. Dr. Alper Tunga Akarsubaşı

HAZİRAN 2007

ACKNOWLEDGEMENTS

Completing a thesis is, first and foremost, an academic exercise. It would not have been possible without a kind and supportive academic advisor. I would hereby like to thank to my supervisor, Prof. Nüzhet Dalfes, who has always guided me through the steps of my scientific research and shared his academic experiences with me. On the other hand, he made easier for me to work on ecology and climate sciences which I was not familiar with. It was my luck to work with him.

I also would like to thank to Asst. Prof. Zehra Çataltepe who gave me a data mining vision and assisted me at machine-learning scope of my thesis. Her past experiences at neural networks researches were priceless for me.

Also, thanks to Asst. Prof. Neslihan Serap Şengör and Asst. Prof. Alper Tunga Akarsubaşı for their kind and helpful advises.

Thanks to Informatics Institute stuff for satisfying all my thirst for knowledge and my friends who shared their knowledge with me.

Finally, my warmest thanks to my family for teaching me the value of a good education and for always encouraging my academic pursuits.

June 2007 Hasan Serhan AKIN

CONTENTS

ABBREVIATIONS **vi**
LIST OF TABLES **vii**
LIST OF FIGURES **viii**
LIST OF SYMBOLS **iix**
SUMMARY **x**
ÖZET **xii**

1. INTRODUCTION **1**

2. CLIMATE CHANGE **3**

2.1. Climate Models and Emissions Scenarios 6
2.2. Effects of Climate Change on Species and Ecosystems 8

3. ECOLOGICAL NICHE MODELING **9**

3.1 Generalized Linear and Additive Models 10
3.2 Classification and Regression Trees 11
3.3. Artificial Neural Networks 13
3.3.1 Brief History 13
3.3.2 Basic Concepts 14
3.3.3 Back-propagation Algorithm 17
3.3.4 Artificial Neural Networks in Ecological Modeling Literature 18
3.4. Other Techniques 19

4. METHODOLOGY **20**

4.1. Data 23
4.1.1 Climatology Data 23
4.1.2 Species' Distribution Data 24
4.1.3 Climate Scenarios Data 26
4.2. Data Preprocessing 27
4.3 Training Artificial Neural Network 28
4.3.1 The Network Architecture 28

4.3.2 Back-propagation Parameters 30

5. RESULTS 34

5.1 Performance Measures 34
5.1.1 ROC Analysis 34
5.1.2 Cohen's Kappa Statistics 35
5.2 Model Performance and Threshold Selection 35
5.3 Model Comparison 37
5.4 Future Predictions 39

6. DISCUSSION AND CONCLUSION 40

REFERENCES 42
APPENDIX A 50
APPENDIX B 67
APPENDIX C 70
BIOGRAPHY 77

ABBREVIATIONS

ANNs : Artificial Neural Networks
AOGCM : Atmosphere-Ocean Global Climate Model
AUC : Area under the Curve
CART : Classification and Regression Tree
CRU : Climate Research Unit
FN : False Negative
FP : False Positive
GCM : Global Climate Model
GLM : Generalized Linear Model
GAM : Generalized Additive Model
IPCC : Intergovernmental Panel on Climate Change
MLP : Multilayer Perceptron
ROC : Receiver Operating Characteristic
TN : True Negative
TP : True Positive

LIST OF TABLES

Table 3.1: Some of link functions used in GLMs 11

Table 4.1: Relation between environmental variables and scale domains 22

Table 4.2: Optimal nH, η and α values obtained from the experimental search 32

Table 4.3: Training and validation errors during 100,000 epochs 32

Table 5.1: Confusion Matrix 34

Table 5.2: Model assessment by using Kappa measure 35

Table 5.3: Kappa statistics of each species 36

Table 5.4: Species' prediction success on test data with respect to selected cut-point threshold 36

Table 5.5: AUC and κ measures of GLM, CART and ANNs models on six pilot species 38

LIST OF FIGURES

Figure 2.1: Climate Change during the last 65 million years..... 3

Figure 2.2: Temperature Changes during Ice Age**.**..... 4

Figure 2.3: Medieval Warm Period and Little Ice Age..... 4

Figure 2.4: The instrumental record of global average temperatures since the last half of 19th century. 5

Figure 2.5: The time evolution of the globally averaged temperature change relative to the years (1961 to 1990) of the SRES simulation B2. 7

Figure 2.6: The time evolution of the globally averaged temperature change relative to the years (1961 to 1990) of the SRES simulations 7

Figure 3.1: A binary decision tree splits a domain into sub-domains by using recursive test until sub-domains are pure enough. 12

Figure 3.2: Nonlinear model of a neuron..... 15

Figure 3.3: Sigmoid (logistic) function with λ=1..... 15

Figure 3.4: Neural network with different flow types. (*i*) a back-propagation network, (*ii*) a Hopfield network 16

Figure 4.1: Resolution error of *Asimina triloba*..... 20

Figure 4.2: Model outline 22

Figure 4.3: Six pilot species used in model 25

Figure 4.4: Distribution Maps of each species..... 26

Figure 4.5: Network architecture with 36 input nodes and one output node. 29

Figure 4.6: Stochastic Back-propagation Algorithm 30

Figure 4.7: Training and validation errors during 100,000 epochs for *Prosopis juliflora* 33

Figure 5.1: Maps of current range of each species predicted by the model. 36

Figure 5.2: ROC curves of each 6 species with three modelling technique. 38

LIST OF SYMBOLS

φ	: Activation function
η	: Learning rate
α	: Momentum
δ_j	: Error on a node
w_{ij}	: Weight
Δw	: Weight change
$J(w)$	: Total error
n_I	: Number of input units
n_O	: Number of output units
n_H	: Number of hidden units
t_k	: Target value
κ	: Kappa value

DISTRIBUTION OF NORTH AMERICAN TREE SPECIES UNDER CLIMATIC CHANGE: AN ECOLOGICAL NICHE MODELING STUDY USING ARTIFICIAL NEURAL NETWORKS

SUMMARY

Since 19th century human activities have been significantly increasing the concentration of greenhouse gases in the atmosphere which changes earth's radiative balance and warms earth surface. Climate Scenarios assessed by the Inter-governmental Panel on Climate Change (IPCC) estimates 1.4 to 5.8°C warming at mean surface temperatures, 0.99 to 0.88 m. rise at sea-levels, precipitation increase in high-latitude and equatorial areas, and precipitation decrease in the subtropics with increase in heavy precipitation events. These dramatic changes will affect directly on ecosystems and species. While some species will be able to adapt to new conditions and even expand, some will shrink both population and distribution. To determine habitat suitability in new conditions arises as a critical problem that must be revealed by ecologists and conservation biologist especially regarding endangered species. In this study, a popular machine learning method, artificial neural networks, is employed to model spatial distribution of tree species in North America. A feed-forward neural network with back-propagation learning algorithm is used to identify relationships between environmental conditions and species distribution. Three main climatic variables from CRU 1.0 Dataset at 0.5˚ latitude/longitude resolution containing 1961-1990 climatology are used: mean temperature, diurnal temperature range and precipitation. Six tree species from digital representation of "Atlas of United States Trees" are chosen as having different geographic distributions: a large perennial shrub native to western North America *Amelanchier alnifolia*, a large shrub native to eastern North America *Asimina triloba*, a medium-sized deciduous tree native to northern North America *Betula papyrifera*, a perennial deciduous thorny shrub indigenous to arid Central America *Prosopis juliflora*, a small, deciduous tree native to eastern coast of United States *Ptelea trifoliate* and a perennial tree originating in the hammocks of the American Tropics *Zanthoxylum fagara*. For each six pilot species, the feed-forward network was constructed with one hidden layer. The inputs of the network are three climate variables for each month, 3×12=36 nodes, and output is one node as binary value of presence/absence of selected species on a point of the 0.5˚ global grid. Available data is split to three parts as 70% for training, 20% for validation to find optimal weights and 10% for testing performance of the network. To determine optimal number of hidden units, learning rate and momentum parameters, an exhaustive search is carried out for each species; and parameters combination ensuring minimum validation error during 10,000 epochs is chosen as best. Training with optimal parameters are repeated for 100,000 epochs and at every 100 epoch, quadratic error on the validation data is measured and the weights of the network generating minimum validation error are chosen as optimal

weights. Final network is tested on the test data and a performance analysis is carried out on it. In addition, Generalized Linear Modelling and Classification-Regression Trees techniques are applied on to the same data set. Then, their performance compared with the performance of the ANNs technique and comparison shows that ANNs technique has the highest performance. Finally, to select a cut-point threshold value for binary classification, a threshold value in range of [0, 1] achieving the highest Kappa measure is chosen for each species. Then, the network is used to predict presence/absence of species with respect to twelve combinations of projected climate conditions of 2050 and 2100 years with the best and worst SRES emissions scenarios (A1FI and B1) and three common models (CGCM2, DOE PCM, HadCM3). Predictions generated shows that *Prosopis juliflora* native to arid climate will expand towards regions currently with temperate or humid climate and *Betula papyrifera* will shift towards the North.

İKLİMSEL DEĞİŞİKLİKLERİN KUZEY AMERİKA AĞAÇ TÜRLERİNİN DAĞILIMI ÜZERİNE ETKİLERİNİN TAHMİNİ: YAPAY SİNİR AĞLARI YAKLAŞIMIYLA EKOLOJİK NİŞ MODELLEME ÇALIŞMASI

ÖZET

19. yüzyıldan beri süregelen insan faaliyetleri atmosferdeki sera gazı miktarı arttırmakta ve bu da dünyanın ısıl dengesini değiştirerek yüzey sıcaklarının artmasına neden olmaktadır. Hükümetlerarası İklim Değişikliği Panel'inde (IPCC) değerlerdirilen iklim modelleri önümüzdeki dönemde, ortalama yüzey sıcaklıklarında 1.4 ile 5.8°C artış, deniz seviyesinde 0.99 ile 0.88 m yükselme, ekvator çevresi ile yüksek enlemlerdeki bölgelerde yağış artışı ve subtropik bölgelerde ise ortalama yağışta azalmayla birlikte kısa süreli ve yoğun yağış olaylarının gerçekleşmesini beklemektedir. Bu ciddi değişiklikler hiç şüphesiz doğrudan türleri ve daha geniş perspektifte ekosistemleri etkileyecektir. Bu değişiklikler neticesinde bazı türler yeni şartlara iyi adapte olabilecek ve hatta bazıları istalacı tür haline gelecek, bazı türler ise yeni şartlara uyum sağlayamadığı için yaşam alanları daralacak ve de yok olma tehdidiyle yüzyüze gelecekler. Bu açıdan, iklim değişiklikleriyle birlikte ortaya çıkacak yeni şartlar altında türlerin ne tür tepkiler verebileceğini öngörebilmek ekoloji alanında çalışan bilim insanları açısından kritik bir önem arzetmektedir. Bu çalışmada popular bir makine öğrenmesi yöntemi olan yapay sinir ağları tekniği Kuzey Amerika'daki ağaç türlerinin uzaysal dağılımlarının modellenmesi amacıyla kullanılmıştır. Bu maksatla, ileri beslemeli bir sinir ağı yapısı, geriye yayınım öğrenme algoritması yardımıyla, seçilmiş bir türün coğrafi dağılımı ile çevre koşulları arasındaki ilişkileri ortaya çıkarmakta kullanılmıştır. İklim verisi olarak, 1961-1990 yılları ortalamaları üzerinden derlenmiş üç temel iklim değişkeni; ortalama sıcaklık, günlük sıcaklık değişimi ve yağış değişkenleri alınmıştır. "Birleşik Devletler'deki Ağaçlar Atlası" içinden seçilen altı farklı ağaç türü model için kullanılmıştır: Batı Kuzey Amerika'ya özgü *Amelanchier alnifolia*, Kuzey Amerika'nın doğusunda görülen *Asimina triloba*, Kuzey Amerika'nın kuzeyine özgü *Betula papyrifera*, kurak Orta Amerika'ya özgü iri bir çalı türü olan *Prosopis juliflora*, Kuzey Amerika'nın güneydoğusuna özgü *Ptelea trifoliate*, ve Tropik Amerika'ya özgü *Zanthoxylum fagara*. Her altı tür için ileri beslemeli ağ mimarisi, bir girdi, bir çıktı ve de bir gizli katman ile oluşturulmuştur. Ağın girdileri grid bir noktaya ait 3 iklim değişkeninin tüm aylar için sahip olduğu ortalama değerlere karşılık gelen 36 adet düğümden oluşturulmuş ve ağın çıktı katmanı da türün herhangi bir grid noktadaki varlığı ya da yokluğunu temsil eden ikilik değere karşık gelen bir düğüm ile ifade edilmiştir. Mevcut veri kümesindeki örnekler rastgele bir biçimde %70, %20 ve %10'luk üç ayrı kümeye ayrılımış bunlardan ilk küme ağın eğitilmesi, ikinci küme ağın ayarlanması (validation) ve sonuncu ise ağın başarımını ölçmek maksadıyla kullanılmıştır. Ağın en uygun şekilde eğitilmesini sağlayacak parametrelerin buluna-bilmesi için ayrıntılı bir araştırma yürütülmüştür. Altı farklı türün herbiri için yapılan bu deneysel araştırmada; saklı düğümlerin sayısı, öğrenme

oranı ve momentum parametrelerinin farklı değerlerinden oluşan 105 kombinasyon ayrı ayrı 10,000 devir boyunca eğitilmiş ve içlerinden minumum doğrulama (validation) hatası üreten kombinasyon 'en iyi' olarak seçilmiştir. Elde edilen 'en iyi' parametreler daha sonra ağın 100,000 devir boyunca eğitilmesinde kullanılmıştır. Bu eğitim süreçlerinde yine her tür için minumum doğrulama hatasının oluştuğu devirdeki ağırlıklar en iyi olarak kabul edilmiş ve ağın bu hali başarım testleri ve türlerin dağılım tahmini işlemlerinde kullanılmıştır. Buna ek olarak elde edilen model sonuçları, Genelleştirilmiş Doğrusal Modelleme (Generalized Linear Modelling) ve Karar Ağaçları teknikleriyle elde edilen sonuçlar ile karşılaştırılmıştır. Sonuçlar yapay sinir ağlarının bu yöntemlere kıyasla daha başarılı sonuçlar ürettiğini ortaya koymaktadır. Son olarak, yeni iklim koşullarında her tür için, oluşturulacak dağılımların, türün varlığı ya da yokluğu bağlamında sınıflandırabilmesi amacıyla [0,1] aralığında en yüksek Kappa ölçüsünü üreten değer, eşik değer olarak seçilmiştir. Daha sonra ağ, en iyi ve en kötü sera gazı salınım senaryolarına göre üç iklim modeliyle oluşturulmuş 2050 ve 2100 yıllarına ait iklim girdilerini kullanılarak, bu yıllara ait dağılım tahminleri üretmekte kullanılmıştır. Sonuçlar göstermektedir ki kurak iklim koşullarında yaşayabilen *Prosopis juliflora*, bugün daha ılıman sayılan bölgelere doğru yayılma gösterirken, *Betula papyrifera* türünün dağılımı kuzeye doğru ötelenecektir.

1. INTRODUCTION

Observations on atmosphere (e.g. **Keeling et al. 1995; Andres et al., 1999**) expose that greenhouse gases concentration such as CO_2 and CH_4 are increasing. At the same time, temperature, precipitation, sea levels, polar glacial are changing and also, extraordinary climatic events such as heavy precipitation or heat waves are occurring more frequently. On the other hand, **IPCC (2002)** states that climate is a major factor on species which will directly be affected by a potential climate change. In new conditions, some species will be able to survive and even become invasive species or some will fail to survive and extinct. For this reason, prediction of species availability and spatial distributions becomes very important for conservation biology, ecology, and invasive-species management fields.

In this study, an artificial neural network based model is used to explore future distributions of tree species' by correlating spatial distributions with climate variables. Six pilot species is selected from different parts of the North American region to scan different climatic conditions and different intensities of climate change. To model the relationship between climate parameters and species availability, a classification task is carried out on the pilot species in the study region. Because of difficulties of dynamic models specified in section 3, a static model is used with time-independent equilibrium predictions of species occurrence as a function of climate variables. To achieve classification, model uses current presence or absence (binary) information of species at all grid points of study region and for each species learns relations between climate and species. Then, trained model is used to explore future distribution changes in response to expected climate change until the end of 21st century.

Because of inter-discipliner characteristic of this study, some introductory material about concepts and methods referred in the research should be supplied. For this reason, section 2 and section 3 include definitions and brief literature about climate change and ecological niche modelling with common statistical techniques respectively. In Section 4, model details are introduced and all steps followed in the

development of the model are explained. In Section 5, performance of model and future projections are represented.

2. CLIMATE CHANGE

Climate change refers to any change in climate over time as a result of natural processes or human activities **(IPCC, 2001)**. In the environmental policy usage, climate change usually refers to changes in modern climate that is increase of average surface temperature and variation of precipitation, which is also known as *global warming*.

The Earth experienced numerous climate changes in the past. This fact has been brought to light by means of evidences coming from several sources such as analysis of pollen and beetles in freshwater and land sediments, glacial geology records and historical documents. For example, paleoclimatologic observations expose a sudden global climate change at the end of Paleocene 55.5 to 54.8 million years ago with the most rapid and extreme heating up in the geologic history of earth ever been seen. Figure 2.1 illustrates this temperature change with oxygen isotope measurements of sea water as continental ice sheets at Vostok, Antartica (**Rohde, 2005** based on **Zachos et al., 2001**). **Kennett and Stott (1991)** found a sharp decrease in the amount of heavy carbon in 55-million-year-old marine fossils which led to the discovery of mass extinction of deep-sea benthic foraminifera, a large group of amoeboid protists.

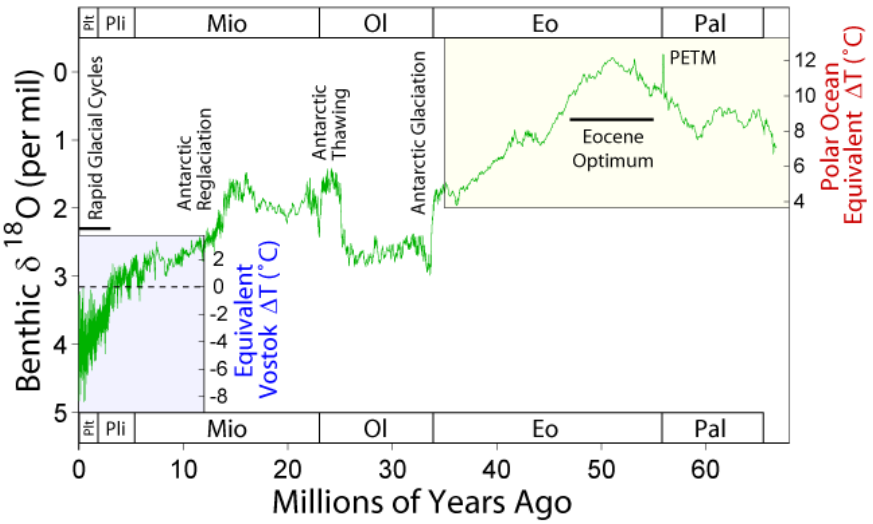


Figure 2.1: Climate Change during the last 65 million years.

Another example among the dramatic climate changes is Dansgaard-Oeschger events during the last ice age between 110,000 and 23,000 years ago which were rapid climate fluctuations. Figure 2.2 illustrates temperature changes in which the first two curves shows local changes in temperature at two sites in Antarctica as derived from deuterium isotopic measurements on ice cores (**Rohde, 2005** based on **EPICA Community Members, 2004**; **Petit et al., 1999**). Greenland ice core records GRIP and GISP2 (**Alley et al., 1995**) demonstrate that quick warming decades usually followed by long cool periods. For instance, the Greenland warmed by roughly 8˚C over 40 years nearly 11,500 years ago. **(Alley 2000; Steward 2005)**.

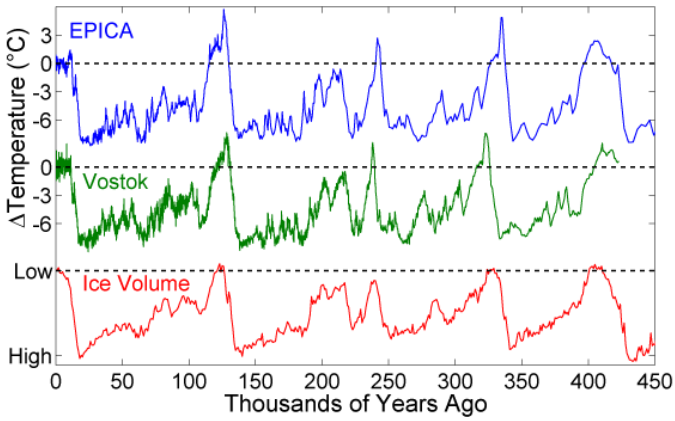


Figure 2.2: Temperature Changes during Ice Age**.**

Climate changes did not occur only during prehistoric periods. For example, Medieval Warm Period was a time of unusually warm weather around 800-1300 AD, during the European Medieval period and the following Little Ice Age to the 16th to the mid-19th centuries with slight warming intervals. (Figure 2.3 by **Rohde, 2005**)

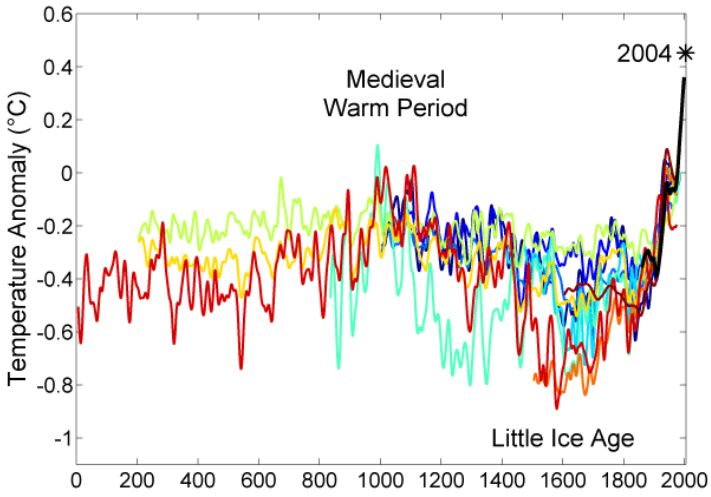


Figure 2.3: Medieval Warm Period and Little Ice Age derived from separate historical sources

However, it is claimed that these cold and warm periods could show regional changes. (**Bradley and Jones, 1993**; **Hughes and Diaz, 1994; Crowley and Lowery, 2000**) Evidence from mountain glacial exposes that those changes were not a globally-synchronous increased glaciations but could be regional climate changes which are parallel to historically documented climate changes in Europe. (**Bradley, 1999)**.

All climate changes occurred on the Earth were results of natural processes such as change of greenhouse gases concentration, plate tectonics, solar or orbital variations and volcanic emissions. Variations within Earth's climate such as glaciations or ocean variability led to climate changes as well. However, growth in industry, agriculture, and transportation since the Industrial Revolution has caused to be emitted additional quantity of greenhouse gases, especially CO_2, augmenting the thermal blanket and leading to a human induced climate change. According to **Folland et al. (2001), Levinson (2005)** global surface mean temperatures has risen around 0.6°C since the last quarter of 19th century. Figure 2.4 illustrates this fact compiled by the Climatic Research Unit of the University of East Anglia and the Hadley Centre of the UK Meteorological Office (**Rohde, 2005**).

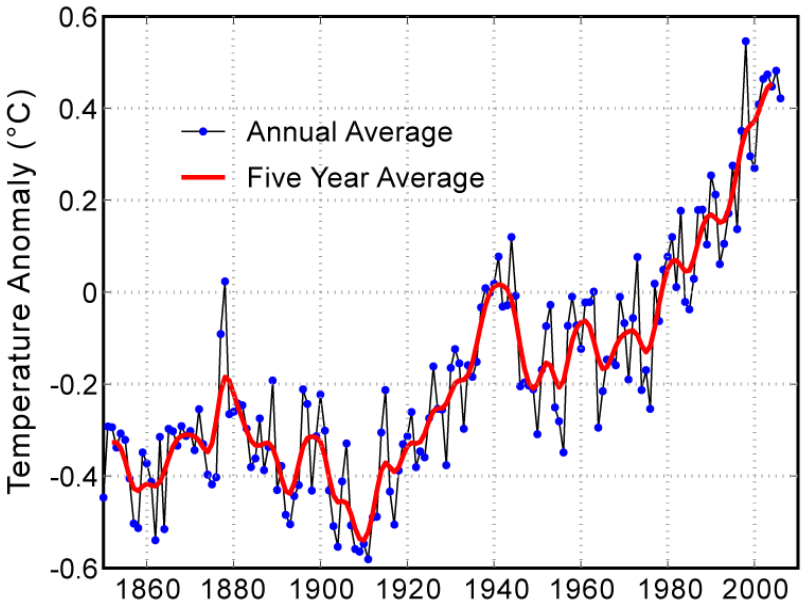


Figure 2.4: The instrumental record of global average temperatures since the last half of 19th century.

The Intergovernmental Panel on Climate Change **IPCC (2007)** concludes, "Most of the observed increase in globally averaged temperatures since the mid-20th century is very likely due to the observed increase in anthropogenic greenhouse gas concentrations" which is pointing to human-induced *greenhouse effect*. The term greenhouse effect introduced by **Joseph Fourier (1824)** is the process which greenhouse gases absorb infrared radiation emitted by the Earth and re-radiate the energy coming from sun as heat back towards to the Earth as a result of warming on the surface. Human induced factors such as fossil fuel combustion, cement manufacture, land use, ozone depletion, animal agriculture and deforestation changes the intensity of greenhouse effect.

2.1. Climate Models and Emissions Scenarios

Scientists have been developing and using global climatic models to provide simulations that can give information about atmospheric circulations at present and future conditions regarding the increase of greenhouse gases in the atmosphere. However, climate models developed so far have been relative simple to quite complex but commonly used models are coupled atmosphere-ocean global climate models (AOGCM) which are the most complex ones and combine both atmosphere global climate models and ocean global climate models. Some recent AOGCMs also include the biosphere, carbon cycle and atmospheric chemistry. A consolidated list of coupled AOGCMs at **IPCC (2001)** includes 31 models. Most commonly used models are HadCM3 (Hadley Centre Coupled Model version 3) **Gordon et al. (2000)**, CGCM2 (The Second Generation Coupled Global Climate Model) **Flato and Boer (2001)**, CSIRO Mk2 (Commonwealth Scientific and Industrial Research Organization Atmospheric Research Mark 2b climate model) **Gordon and O'Farrell (1997)**, DOE PCM (DOE-sponsored parallel climate model) **Washington et al. (2000)**, ECHam4 **Roeckner et al. (1996)**. Estimated temperature change until 2100 of these models can be seen in Figure 2.5.

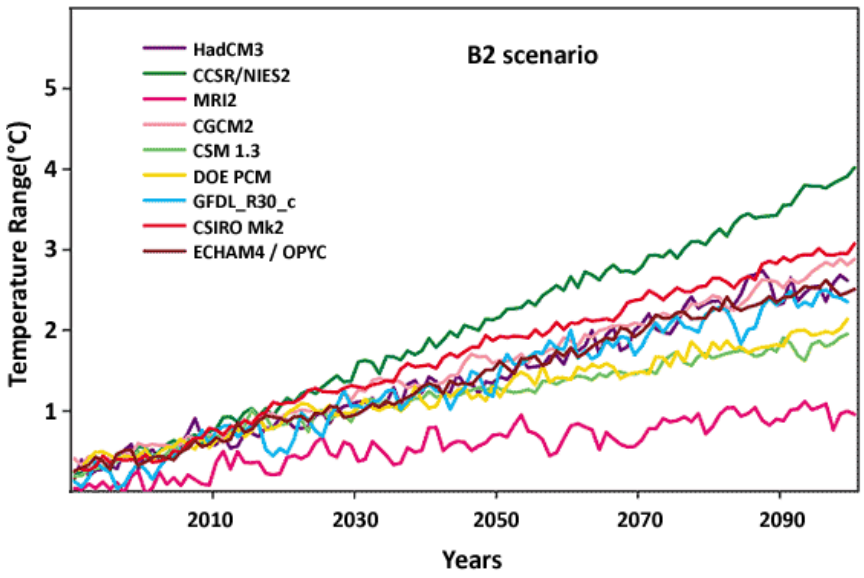


Figure 2.5: The time evolution of the globally averaged temperature change relative to the years (1961 to 1990) of the SRES simulation B2, based on **IPCC (2001)**.

Whilst generating future climate projections, most models use IPCC Special Report on Emissions Scenarios (SRES). **(IPCC, 2000)** There are actually 40 different scenarios with respect to different assumptions for future greenhouse gas pollution, land-use and deforestation. Nevertheless, IPCC's Third Assessment Report and Fourth Assessment Report include six families which are A1FI, A1B, A1T, A2, B1, and B2. (Figure 2.6)

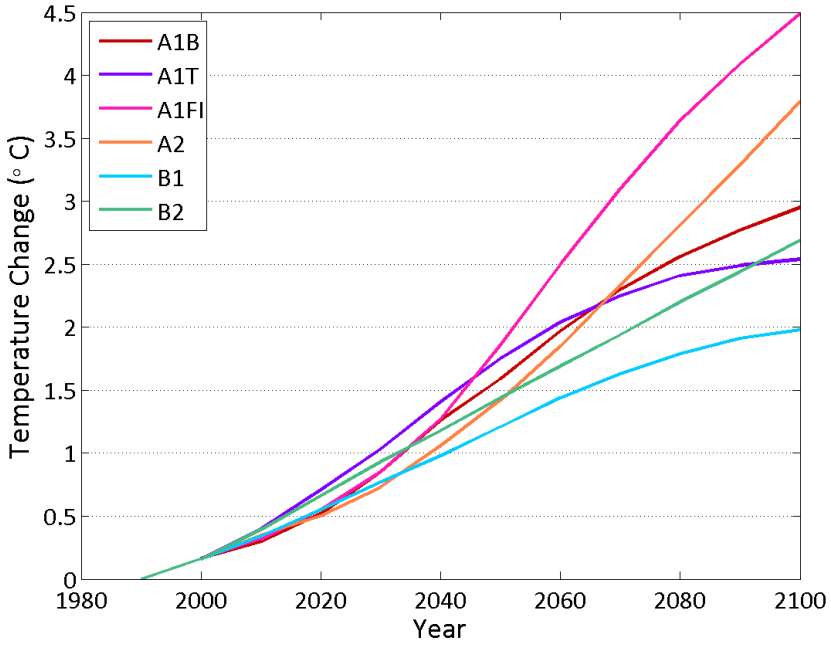


Figure 2.6: The time evolution of the globally averaged temperature change relative to the years (1961 to 1990) of the SRES simulations based on **IPCC (2001)**.

2.2. Effects of Climate Change on Species and Ecosystems

For the wide range of **IPCC (2001)** emissions scenarios, 1.4 to 5.8°C warming at mean surface temperatures, 0.99 to 0.88 m. rise at sea-levels, precipitation increase in high-latitude and equatorial areas, and precipitation decrease in the subtropics with increase in heavy precipitation events are foreseen. Climate Change projections claim that individual organisms, populations, species distributions, and ecosystems will be affected (**IPCC, 2002**). For instance, changes in sea level will directly affect marine and coastal ecosystems (e.g., mangrove/coastal wetlands, sea grasses). El Niño event of the years 1997-1998 affected on several ecosystems, ranging from deserts to tropical rain forests (**Holmgren et al., 2001**). Also, many studies **Huntley(1991), Davis and Shaw (2001), Davis et al. (2001), DiMichele et al.(2001)** are correcting that by means of evidence from fossil records, climate has been key factor on species' distribution, population and extinction. Furthermore, studies on species distribution change over the past 30 years expose that numerous shifts in the distributions and abundances of species have been occurred. For example, **Parmesan et al. (1999)** provided that in a sample of non-migratory European butterflies, 63% of them shifted their range to the north by 35-240 km during this century. Similarly, **Thomas and Lennon (1999**) found that the northern margins of many British bird species have moved further north by an average of 18.9 km over a 20-year period.

3. ECOLOGICAL NICHE MODELING

The possible impacts of climate change on species and ecosystems force scientists to develop predictive models that identify the correlation between species distribution and climate variables. For this reason, there are a number of modelling strategies developed to explore these concerns. These models are mainly relied on species' *ecological niche* concept.

General ecological niche concept can be given as the environmental factors necessary for a species survival. The niche concept was firstly used by **Grinnell (1917)** whose niche definition was focused on influences of physical environment on species. Then, **Elton (1927)** used same concept but included biotical interactions as well as abiotical factors. **Gause (1934)** developed the niche concept and discovered *competitive exclusion principle* which states that two species having identical niches cannot stably coexist. Although recent niche concept reviewed in many studies (e.g. **Malanson et al., 1992; Rutherford et al., 1995; Leibold, 1995**), there are slight differences and they are mainly relied on *Hutchinson's niche concept*. **Hutchinson (1957)** clearly defined the concept as *n-dimensional hypervolume* where *n* is the number of environmental factors vital for species survival. Moreover, he defined this hypervolume as *fundamental niche* of species but he also recognized that interactions with other organisms forced species into occupying a smaller niche then fundamental niche and called it as *realized niche*. In general the distinction between fundamental and realized niche is a result of human influence, geographic barriers and biotic interactions such as competition, predation, disease and parasitism.

Because of complexity and heterogeneity of the nature, designing a model with achieving accurate predictions in every aspect of time and space is a great challenge. **Levins (1966)** discussed a trade-off between generality, reality and precision properties of models. He claims that it is possible to improve just two of these three properties at the same time. This principle implies that models center around three groups: *analytical, mechanistic (physiological)* and *empirical*. Predictive species distribution models are generally classified as empirical models

but can be based on physiologically meaningful parameters (**Guisan and Zimmermann, 2000**). So, the definition indicates that models based on large empirical data sets constitute realized niche of species.

On the other hand, a model can be classified as *static* or *dynamic*. A static model makes time-independent predictions as a function of environmental variables with the assumption of equilibrium, besides a dynamic model provides time-dependent predictions with respect to changing environmental variables. However, most models used for predicting large-scale distribution and abundance of species are static since static models are relatively easy to build, parameterize and test **(Guisan & Zimmermann, 2000)**. Nevertheless, equilibrium assumption is not realistic for static models which have difficulty in most cases but, dynamic models are able to address this issue. For example, forest ecologists have been using dynamic models for years (e.g. **Urban et al., 1991; Gao et al., 1996; He et al., 2002; Gratzer et al., 2004**). Achilles' heel for this approach is, however, requirement of too much information of species which is rarely possible to obtain.

There are many ecological niche modelling techniques in the literature. They are basically relied on statistical or data mining approach. In the following sections details of them are presented.

3.1 Generalized Linear and Additive Models

In statistics, the Generalized Linear Models (GLMs) technique presented by **Nelder and Wedderburn (1972),** is a commonly used tool to model response errors that are not normally distributed or do not have constant variance functions. A GLM is based on three components; a response variable Y, a linear predictor and the link function *g*, which defines functional relationship between the expected value $\mu = E(Y)$ and linear predictors. (Equation 3.1)

$$g(\mu(x)) = \beta_0 + \beta_1 x_1 + \ldots + \beta_m x_m \qquad (3.1)$$

GLMs can provide several link functions and convenient error distributions (see Table 3.1) such as binomial, negative binomial, geometric, exponential, Poisson, gamma, and inverse normal distributions those make GLMs flexible enough for the cases where response is in linear relationship with its predictors.

Table 3.1: Some of link functions used in GLMs

Distribution	Name	Link Function
Normal	Identity	$g(\mu(x)) = \mu(x)$
Exponential	Inverse	$g(\mu(x)) = \mu^{-1}(x)$
Gamma		
Poisson	Log	$g(\mu(x)) = \ln \mu(x)$
Binomial	Logit	$g(\mu(x)) = \ln\left(\frac{\mu(x)}{1-\mu(x)}\right)$
Multinomial		

The assumption that response is in linear relationship with its predictors is not realistic in many cases. **Hastie and Tibshirani (1990)** proposed Generalized Additive Models (GAM) approach which is non-parametric extension to GLMs. Assumption in this approach is that to use a series of non-parametric smoothing functions as components instead of using coefficients. (Equation 3.2)

$$g(\mu(x)) = \beta_0 + f_1(x_1) + \cdots + f_m(x_m) \tag{3.2}$$

Hastie and Tibshirani (1990) discuss several smoothing methods. One such function, for example, is the *cubic splines smoother*, which is a piecewise cubic polynomial with pieces joined at the unique observed value of *x* in a dataset.

In distribution modelling literature, GLM and GAM techniques are broadly applied. For example, **Yee and Mitchell (1991)** employed GAM to model plant distributions. Also, **Bio et al. (1998)** used both GLM and GAM in their study on wetlands and water plants. **Augustin et al. (1996)** applied GLM to red deers' spatial distribution, **Cumming's (2000)** study on bont tick (*Amblyomma hebraeum* Koch) is based on GLM. **Pearce and Ferrier (2000)** applied GLM and GAM to distribution of 24 species from fauna and flora of north-east New Wales.

3.2 Classification and Regression Trees

Classification and Regression Tree (CART) algorithm **(Breiman et al., 1984)** is a widely used statistical procedure for producing classification and regression models with a binary *decision tree* algorithm, in which each *decision node* has exactly two *branches*. A decision tree is a hierarchical model that enables

identification of local regions by means of recursive splits and it is comprised of *decision nodes* and *leaves* (Figure 3.1) where each decision node represents a test of an attribute with a test function $f_m(\boldsymbol{x})$ and a leave represents a class or a class distribution **(Alpaydın, 2004).**

The CART building algorithm is a greedy algorithm in that the best $f_m(\boldsymbol{x})$ which generates the least impure partitions is chosen among all possible test functions. This process is carried out recursively on each sub-partition until sub-partitions are pure enough. To measure impurity **Breiman et al. (1984)** proposes several impurity measures. One of them is *Information Gain (Entropy)*:

$$i(N) = -\sum_{j} P(\omega_j) \log_2 P(\omega_j) \tag{3.3}$$

Another measure is *Gini Index*:

$$i(N) = \sum_{i \neq j} P(\omega_i) P(\omega_j) = \frac{1}{2}\left[1 - \sum_{j} P^2(\omega_j)\right] \tag{3.4}$$

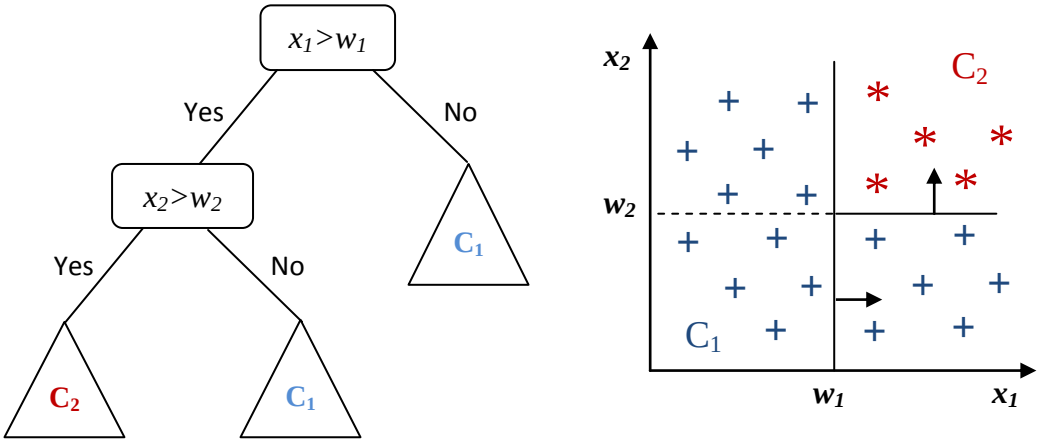


Figure 3.1: A binary decision tree splits a domain into sub-domains by using recursive test until sub-domains are pure enough.

If trees are allowed to grow too much, leaves (final partitions) consist of only few samples which may cause model to over fit. To address this problem *tree pruning* methods are used in which by means of statistical measures the branches having anomalies are eliminated. Two common approaches, *prepruning* and

postpruning, are employed for this purpose (For details see **Han and Kamber, 2001; Alpaydın, 2004**).

In the literature, CART method is used in many studies. For example, **Franklin (1998)** used CART technique with GLM and GAM techniques to predict distribution of shrub species in southern California. **Vayssières et al. (2000)** studied on three major oak species by using GLM and CART. **De'ath and Fabricius (2000)** employed CART to analyze survey data of soft coral taxa from the Australian central Great Barrier Reef. **Thuiller et al (2003)** investigated predictive ability of CART vs. GLM and GAM on tree species.

3.3. Artificial Neural Networks

3.3.1 Brief History

In the 1940s, first computers are built which made possible processing large amount of data. At the same time, it was found that processing architecture of computers were mimic to physiology of human brain. While processing element is called as *bit* in computers, in the brain processing element is the *neuron*. The similarity between computers and human brain motivated to develop brain-like systems that can learn on its own. **McCulloch and Pitts (1943)** showed networks of neurons can compute and learn any function. **Hebb (1949)** stated that neural pathways are strengthened each time they are used (*Hebbian learning*) which contributed to the development of artificial neural networks. **Rosenblatt (1958)** introduced *perceptron*, a hardware neural net for character recognition. **Widrow and Hoff (1960)** developed ADALINE (*adaptive linear combiner*) for adaptive control of noise on telephone lines. In the 1960s, many implementations of neural computers were built. However, hardware capabilities were limited and mathematical theorems were too weak to cope with complex problems. **Minsky and Papert (1969)** criticized limitations of perceptron model. Thus, neural networks studies entered a recession phase.

In 1980s a handful of researchers' studies led to renaissance of neural networks. A number of studies of **Grossberg and Carpenter (1983, 1987, 1988, 1990)** established a new principle of self-organization called *Adaptive Resonance Theory* (ART). **Hopfield (1982)** developed a class of recurrent networks as an

associative memory using statistical mechanics; now known as *Hopfield networks*. **Kohonen (1982)** introduced *Self-Organized Maps* using one and two dimensional lattice structures. **Kirkpatrick et al. (1983)** introduced *Simulated Annealing* for solving combinatorial optimization problems. This concept was later used by **Ackley et al. (1985)** to develop the *Boltzmann machine* which is the first successful realization of multi-layered neural network. **Rumelhart et al. (1986b)** announced the discovery of a method that allowed a network to learn to discriminate between not linearly separable classes. They called *backward propagation of errors* which is a generalization of the *Least Mean Squares* (LMS), a mathematical optimization technique. After those developments, interest on artificial neural networks boosted and thousands of research papers have been published so far.

3.3.2 Basic Concepts

Kohonen (1988) defined artificial neural networks as "massively parallel interconnected networks of simple (usually adaptive) elements and their hierarchical organizations which are intended to interact with the objects of the real world in the same way as biological nervous system does." A neural network is basically comprised of *nodes* and *weights* where they are equivalent to *neurons* and *synapses* in central nervous system respectively. In neuron model, weighted inputs of a neuron are accumulated and then emitted through an activation function as response of neuron. Figure 3.2 illustrates neuron model **(Haykin, 1999)**.

In the neural model there is also a single *bias unit*, denoted by b_k, that is connected to each unit other than the inputs.

$$u_k = \sum_{i=1}^{m} x_i w_{ji} \tag{3.1}$$

$$y_k = \varphi(u_k + b_k) \tag{3.2}$$

where $x_i \in R$, $i=1,...m$ are input signals, $w_{ki} \in R$, $i=1,...m$ are *connection weight* or *synaptic weight* and y_k is ouput signal of neuron *k.*

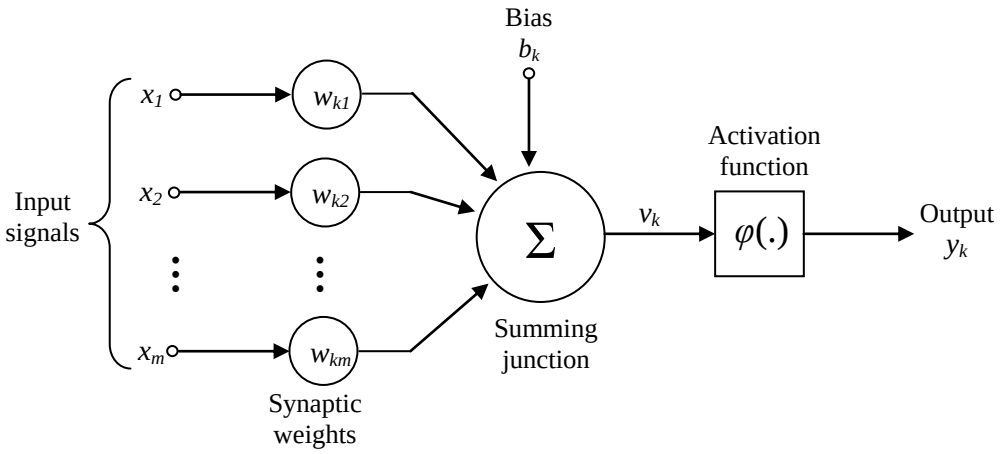


Figure 3.2: Nonlinear model of a neuron

On the other hand, $\varphi(.)$ is referred to an activation function also called transfer function which determines relationship between inputs and output of a node and a network. Although selection of activation function can be problem dependent, an activation function should provide some properties such as nonlinearity, continuity, smoothness, monotonicity and *saturation*, which means having some maximum and minimum output value (**Duda et al., 2000**). One of commonly used activation function is *logistic function* (Equation 3.3) which varies smoothly from 0 at $-\infty$ to 1 at ∞ (Figure 3.3). Also, another *sigmodial* activation function is tanh(v) function which varies between -1 at $-\infty$ to 1 at ∞.

$$\varphi(v) = \frac{1}{1 + \exp(-\lambda v)} \tag{3.3}$$

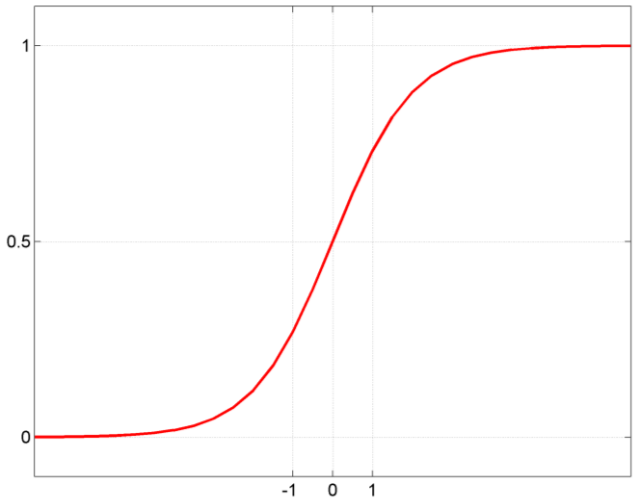


Figure 3.3: Sigmoid (logistic) function with λ=1

Neural networks can be classified in terms of directions of signal flow: *feed-forward networks* and *recurrent networks*. Signals in feed-forward networks propagate from input units to output units and signal flow is just one direction. Nevertheless, in recurrent networks connections between the units shape a directed cycle which means that signals emitted from any neuron may propagate to input neuron. Multilayer perceptron network is an example of feed-forward networks while Hopfield network is a recurrent network (Figure 3.4).

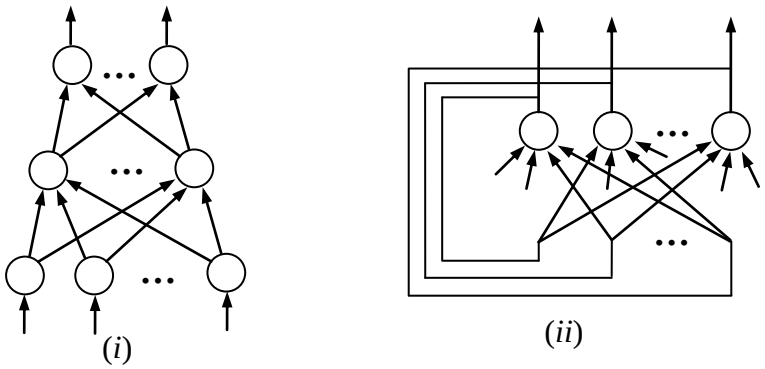


Figure 3.4: Neural network with different flow types. (*i*) a multilayer perceptron network, (*ii*) a Hopfield network

Perceptron concept of **Rosenblatt (1958)** which is comprised of a single neuron is the simplest kind of feed-forward neural network to achieve classification of a group patterns into subgroups having similar feature values (*linear classifier*). However, a perceptron cannot solve classification problems having nonlinear discriminant or cannot be used for nonlinear regression. These limitations can be overcome by using one or more *hidden layers* between input and output layers. Such feed-forward networks are referred to as m*ultilayer perceptrons* (MLPs) and able to solve some difficult problems by training them with *error back-propagation algorithm* **(Haykin, 1999)**. On the other hand, according to continuity theorem of **Kolmogorov (1957)**, any continuous $g(\mathbf{x})$ function on the unit hypercube can be represented as Equation 3.4 where Ξ_j and ψ_{ij} properly chosen functions.

$$g(x) = \sum_{j=1}^{2n+1} \Xi_j \left(\sum_{i=1}^{d} \psi_{ij}(x_i) \right) \tag{3.4}$$

Thus, application of this theorem to neural networks implies that any arbitrary function with continuous inputs and outputs can be approximated with a multilayer

perceptron (*universal approximation*). Furthermore, **Hornik et al. (1989)** proved that one hidden layer MLP can learn any nonlinear function.

3.3.3 Back-propagation Algorithm

"The back-propagation algorithm is one of the simplest and most general methods for supervised training of multilayer neural networks" (**Duda et al., 2000**). It achieves its generality because of the *gradient-descent* technique used to train the network. Gradient descent is an optimization algorithm which tries to reach a local minimum by iteratively moving as much as negative gradient at each step point.

$$x_{k+1} = x_k - \eta_k \nabla J(x_n) \tag{3.5}$$

Similar to gradient descent technique, the Back-propagation algorithm learns to generate a mapping from input pattern space to the out pattern space by minimizing the error between the actual output produced by the network and the desired output across a set of pattern vector pairs or exemplars.

As shown in Figure 3.4, the MLP is a layered, feed-forward network, comprised of one input layer, one or more hidden layers, one output layer and also a bias unit on each layer. The process of the network proceeds according to following algorithm: A vector pair from training set pairs (**x**, **t**) is selected where **x** is input vector and **t** is output (target) vector. Then, net activation is computed using Equation 3.6 and for each hidden unit, activation function emits an output as Equation 3.7.

$$net_j = \sum_{i=1}^{nI} x_i w_{ji} + b_j \tag{3.6}$$

$$y_j = \varphi(net_j) \tag{3.7}$$

On output layer net activations and output values are computed by using the Equations 3.8 and 3.9.

$$net_k = \sum_{i=1}^{nH} y_j w_{kj} + b_k \tag{3.8}$$

$$z_k = \varphi(net_k) \tag{3.9}$$

Then, errors of output and hidden layers evaluated using Equation 3.10 and 3.11.

$$\delta_k = (t_k - z_k)\varphi'(net_k) \tag{3.10}$$

$$\delta_j = \varphi'(net_k)\sum_{k=1}^{nH} w_{kj}\delta_k \tag{3.11}$$

Finally weights are updated using errors at hidden and output layers *(Delta rule*) where η is learning rate. (Equation 3.12, 3.13 and 3.14)

$$\Delta w_{kj} = \eta\delta_k y_j \tag{3.12}$$

$$\Delta w_{ji} = \eta\delta_j x_i \tag{3.13}$$

$$w^{t+1} = w^t + \Delta w^t \tag{3.14}$$

Sometimes $\Delta \boldsymbol{w}^t$ update may cause large oscillations and this leads to slow convergence. To address this issue **Rumelhart et al. (1986a)** modified delta rule as adding *momentum* term. (Equation 3.15-3.16)

$$\Delta w_{kj}{}^t = \alpha\Delta w_{kj}{}^{t-1} + \eta\delta_k{}^t y_j{}^t \tag{3.15}$$

$$\Delta w_{ji}{}^t = \alpha\Delta w_{ji}{}^{t-1} + \eta\delta_j{}^t x_i{}^t \tag{3.16}$$

Haykin (1999) stated that although convergence of back-propagation algorithm cannot be proved, some practical stopping criteria might be derived. For example, **Haykin (1999)** proposed to monitor absolute rate of change in the average *squared error* (Equation 3.17) per epoch. Squared error is a good measure for assessing error since it has helpful properties such as smoothness and differentiability (**Michie et al., 1994**).

$$J(w) \equiv \frac{1}{2}\sum_{k=1}^{nO}(t_k - z_k)^2 \tag{3.17}$$

3.3.4 Artificial Neural Networks in Ecological Modelling Literature

In the literature, artificial neural networks (ANNs) have been used in several studies regarding ecological modelling. **Mastrorillo et al. (1997)** employed ANNs to model spatial presence/absence of three small-bodied fish. **Manel et al. (1999)** applied this technique to prediction of presence/absence of Himalayan river bird. They also compared ANNs with multiple discrimant analysis and logistic regression

techniques and their results shows that logistic regression always outperformed ANNs. Also, **Özesmi and Özesmi (1999)** developed spatial model for habitat selection of marsh-breeding bird species using both ANNs. **Olden and Jackson (2001)** studied on fish habitat model in lakes from south-central Ontoria, Canada with the aid of ANNs technique. They show that ANNs provides greater predictive power than regression techniques do. **Pearson et al. (2002)** employed ANNs to determine impacts of climate change on tree species in Great Britain. **Thuiller (2003)** proposed a computation framework which includes ANNs, GLM, GAM and CART techniques. In that study 61 tree species across Europe were considered for modelling and performance of four techniques was presented. Results show that despite of slight differences, ANNs technique has the best performance.

3.4. Other Techniques

Genetic Algorithm for Rule-set Prediction (**Stockwell and Peters, 1999)** is a machine learning algorithm developed to predict habitat modelling. Algorithm develops a set of rules by using several predictive techniques such as atomic, logistic regression and range rules. Then, best rules are selected by using a genetic algorithm. GARP technique is employed in a number of studies in the literature (e.g. **Peterson and Cohoon, 1999; Peterson, 2001; Lim et al. 2002; Anderson, 2003; Anderson et al., 2003**).

The most recent technique is *Maximum Entropy* (MaxEnt), a statistical mechanics based approach, introduced by **Phillips et al. (2004)**. This technique is based on finding a distribution function of occurrence samples which has maximum entropy, i.e., closest to uniform. Then, it turns into optimization problem:

$$\max_{p \in \Delta} H(p) \text{ subject to } p[f] = \tilde{\pi}[f] \tag{3.18}$$

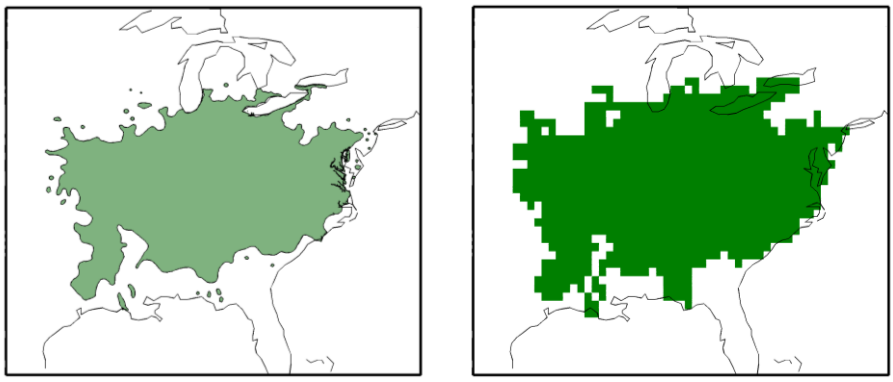
where $H(p) = -\sum_{x \in X} p(x) \ln p(x)$ is entropy of a distribution p; $\boldsymbol{f}$ the vectors of all n features; $\tilde{\pi}[f]$ is empirical average of f. This optimization problem is solved by a sequential update algorithm **(Dudik et al., 2004)** to find distribution p. **Phillips et al. (2006)**, employed Maxent technique to predict distribution of two mammal species from South America.

4. METHODOLOGY

This study focuses on making accurate prediction for tree species in North America in response to climate change expected until the end of 21st century. To be able to make a prediction for a given species with its given spatial distribution, feed-forward artificial neural networks (ANNs) are employed with back-propagation learning algorithm.

To solve this problem ANNs is chosen as its robustness on noisy data and ability to determine complex nonlinear relationships which, in this problem, corresponds to correlation between environmental factors and species' distribution. Thus, ANNs use environmental features to classify presence or absence information of a selected species at a unit spatial region. Also, ANNs technique outperforms the other techniques as mentioned in the previous section.

Tolerance to noisy data property of ANNs is important in this case because surveys are not always perfectly carried out and species distribution data might have observation errors. In addition, a specific issue for this study is resolution error. Environmental data set (see Section 4.1) obtained has 0.5˚ resolution but species' spatial distribution data has higher resolution and needs scaling which induces noise in data. Figure 4.1 illustrates the difference between original data and 0.5˚ data.



Original distribution **0.5˚ resolution of distribution**

Figure 4.1: Resolution error of *Asimina triloba*

On the other hand, Multilayer Perceptrons are universal approximators that can learn any continuous function (See details in Section 3.3.2). Besides, back-propagation learning algorithm is chosen since it is one of the simplest and most general methods for supervised training of multilayer neural networks (**Duda et al., 2000**).

Comparative literature on habitat modelling includes many studies exposing that ANNs achieve better results than traditional linear models (**Manel et al., 1999; Segurado and Araújo, 2004; Thuiller, 2003**). However, those studies did not deeply investigate optimal values of the algorithm parameters (number of hidden units, learning rate and momentum). For example, **Pearson et al. (2002)** used 11 hidden node and learning parameter as 0.2 without a momentum term for 32 plant species. It is very likely that ANNs with those parameters generate low performance for some species. Many other studies (**Manel et al., 1999; Segurado and Araújo, 2004; Thuiller, 2003)** suffer from the same issue. In this study this problem is fixed by performing a discrete scan in parameter space for each study species.

On the other hand, to construct a niche based model, both spatial distribution and environmental data must to be obtained. In reality, data provided in museums, atlases or public databases do not have detailed information about species. Especially, data of time series of a species' distribution rarely exists and this fact prevents development of dynamic models. In this study, distribution data is obtained from "Atlas of United States Trees" by **U.S. Geological Survey (1999) (**see Section 4.1.2**)** which derived from three other atlases published in 1970s. Thus, available data forces us to use a static model based on time independent training data on assumption that data represents an equilibrium state.

Choosing environmental variables is another issue. **Pearson and Dawson (2003)** summarized a hierarchical structure for ecological systems, which previously discussed by **Wu and Loucks (1995)**, **Collingham et al. (2000), Whittaker et al. (2001)** and **Willis and Whittaker(2002)**, and **Pearson and Dawson (2003)** suggested to use climatic factors on global or continental scales (Table 4.1). For this reason, climate variables are chosen as predictor variables in this study.

Table 4.1: Relation between environmental variables and scale domains

	Global >10000km	Continental 2000-10000km	Regional 200-2000km	Landscape 10-200km	Local 1-10km
Climate	✓	✓	✓		
Topography			✓	✓	✓
Land-use				✓	✓
Soil-type					✓
Biotic interactions					✓

While modelling plant species distributions, **Pearson et al. (2002)** employed a climate-hydrological model which consists of mean temperature of coldest one in any year, maximum temperature of warmest month, and growing degree days (average of the daily maximum and minimum temperatures compared to a base temperature). However, instead of monthly averages, using yearly averages as **Pearson et al. (2002)** did may miss some relationships between spatial distribution and climate. For example, in blooming season a plant may be very sensitive to temperature or precipitation values of that period and yearly maximum or minimum values can not represent this relation.

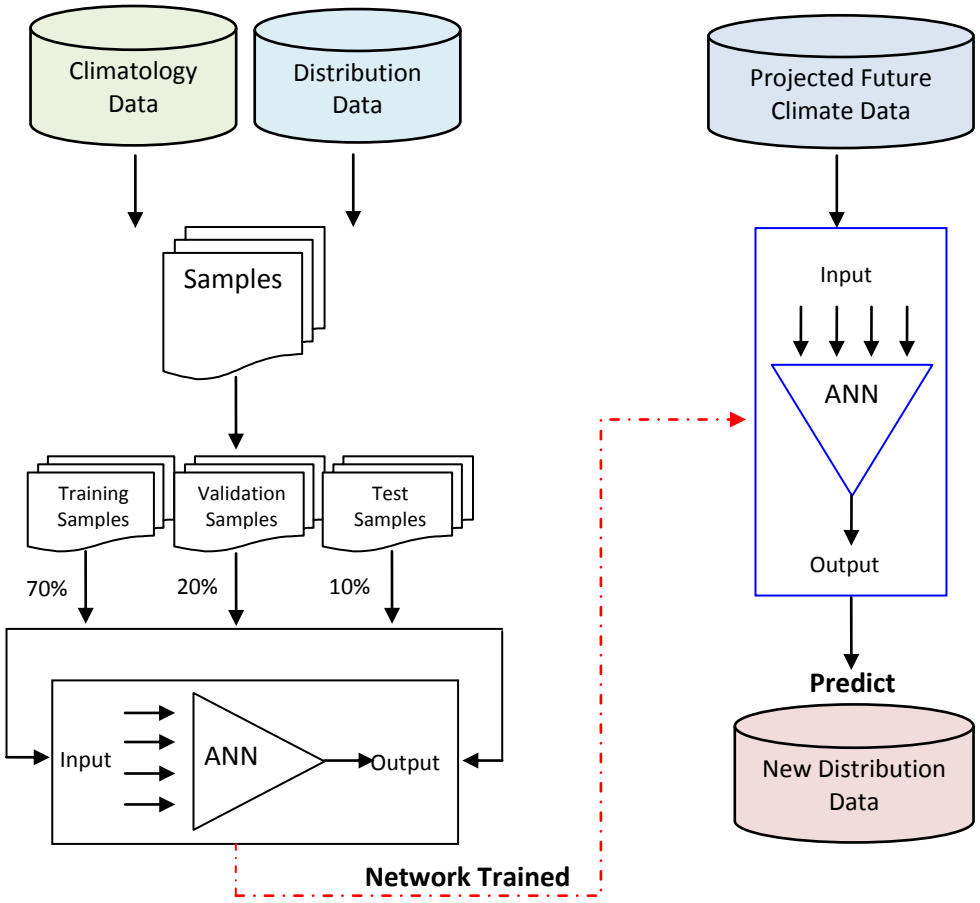


Figure 4.2: Model outline

To summarize, the model developed in this work is basically runs as follows: It receives climatology and distribution data files and produces a sample set. Then, sample set is randomly split into three parts for network training, validation and test with the ratios of 70%, 20%, and 10% respectively. Network trained with the first part and calibrated with the second part during the training process to identify when network begins to over-train. Then, the performance of the produced network is tested with the third part of samples. Finally, future climate data set is given to the final network as input and afterwards distribution of a selected species is evaluated. Figure 4.2 outlines the model details. Data and procedure details are presented in the following sections.

The model procedures in this study are realized by Java programming language-based software which is specifically developed for this study. Java programming language is preferred because of its many advantages as stated in Java language white paper by **Sun Microsystems (1995)**: "Java is a simple, object-oriented, distributed, interpreted, robust, secure, architecture neutral, portable, multithreaded, and dynamic." Java program handles with data pre-processing, training and test procedures. For further information, see Java codes and documentation in attached CD.

4.1. Data

4.1.1 Climatology Data

As input parameters of the ANNs, Climate Research Unit (CRU) 1.0 Dataset **(Hulme et al. 1995)** at 0.5˚ latitude/longitude resolution containing 1961-1990 climatology of three basic climate variables are used: *mean temperature*, *diurnal temperature range* (variation in temperature that occurs from the highs of the day to the cool of nights) and *precipitation*.

Data is provided as separate ASCII files for each climate parameter. The first and second lines of the files contain information of the grid size and the other part of file includes 12 (for each month) global grid matrices, in size of 360×720 rows and columns (Totally, 12×360=4320 lines + 2 header lines exist in the files), in which each elements in grid matrices represent monthly average value of global grid regions. Also, grid matrices just include land values which is called *land/sea mask*.

For example, precipitation data file starts as follows (after header lines just 10 columns are placed here).

grd_sz	xmin	Ymin	xmax	ymax	n_cols	n_rows	n_months	missing	
0.5	0.25	-89.75	359.75	89.75	720	360	12	-9999	
-9999	-9999	-9999	-9999	-9999	-9999	-9999	44	47	43 ...
-9999	-9999	46	46	47	50	51	49	49	48 ...
-9999	-9999	43	-9999	49	51	51	50	48	48 ...

"-9999" values represent grid regions in sea and the other integer values represent (mm/day) ×10 (e.g. "46" means 4.6 mm/day). Similarly mean temperature and diurnal temperature range data files have similar format. They consist of 12 grid matrices (360×720) having °C×10 integer values.

4.1.2 Species' Distribution Data

To complete the training, test and validation stages, binary output variables for observed species' distribution which is obtained from digital representation of "Atlas of United States Trees" (**U.S. Geological Survey**, **1999**) are used. The North American aspect is chosen because this region is diverse in terms of its geological, ecological and climatic properties. Temperature extremes in the region span the range of -40 to +40°C. The Great Plains (including Canadian prairies) and south eastern United States experience more severe weather conditions (e.g., thunder-storms, tornadoes, and hail) than any other region in the world. Virtually all sectors within North America are vulnerable to climate change to some degree in some sub-regions. Six North American plants from different geographic regions in the digital atlas are selected as pilot species:

Amelanchier alnifolia (Western serviceberry), is native to Western North America in a diverse range of habitats extending from near sea level to sub-alpine. It is a deciduous shrub that grows in dense, vegetatively propagated clumps. The low, many-stemmed shrub ranges in height from 1-4 m. Leaves are round in shape, have jagged or toothed margins, are about 1-5cm long, and have stipules.

Asimina triloba (pawpaw), is a large shrub native to eastern North America. It grows to a height of 11 meters (rarely to 14 m) with a trunk diameter of 20-30 cm. The leaves are deciduous, broad lanceolate, 15-30 cm long. The flowers produced in

early spring are dark red, 2-5 cm diameter, with a fetid smell. The fruit is a large yellow-green berry 5-15 cm long, containing several brown seeds.

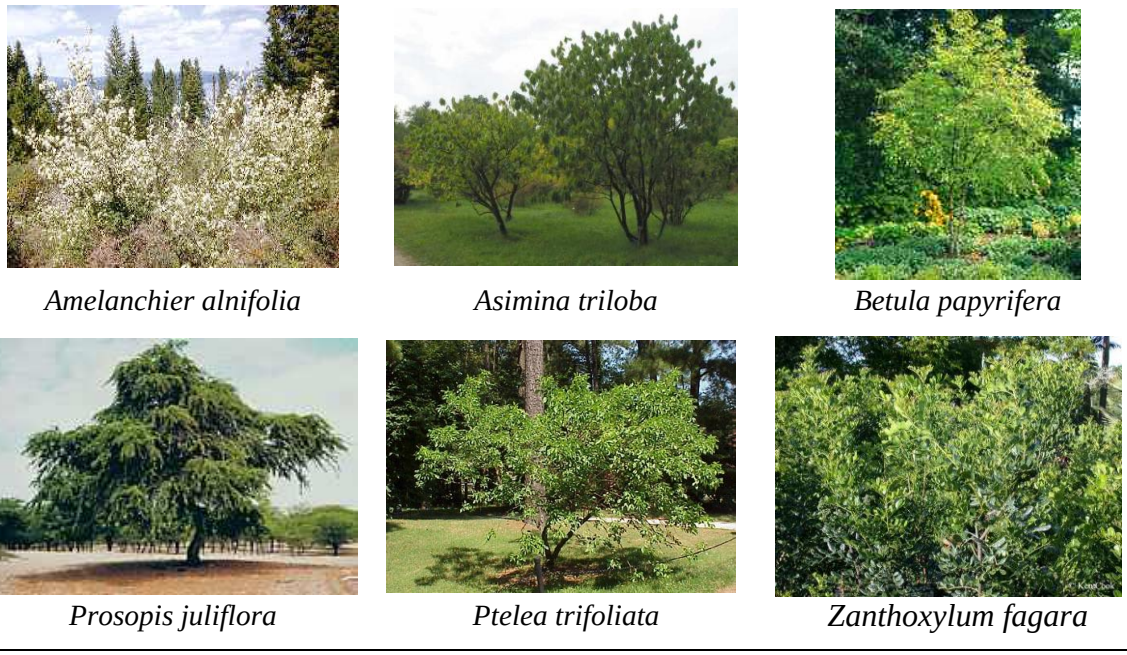


Figure 4.3: Six pilot species used in model

Betula papyrifera (paper birch), is a medium-sized (16 m) deciduous tree, broadleaved hardwood with a small, open crown of spreading and ascending branches. It is native to northern North America, from coast to coast, from Pacific Alaska across all of Canada to Atlantic Newfoundland, as far north as trees grow. It has 5-10cm long, 4-5cm wide, ovate, long-pointed, coarsely and doubly saw-toothed, usually with five to nine veins on each side leaves. The paper birch has both male and female flowers called catkins. These turn into little winged nutkins, which ripen in early August to mid September.

Prosopis juliflora (Mesquite) is a perennial deciduous thorny shrub up to 9m indigenous to arid Central America from Mexico south to Columbia. The leaves, 7.6 cm to 20.3 cm long, are first divided into 1 to 2 pairs of primary divisions. Each of these is again divided into about 10 to 28 pairs of finely hairy or hairless secondary leaflets.

Ptelea trifoliata (Hoptree), is a small, deciduous tree or large shrub (5-8m), native to eastern coast of United States. It has pinnate, trifoliate, deciduous leaves with ovate to elliptical leaflet shape and shiny, dark green color also has small,

greenish-white flowers blooms in early June. Seeds are green, changing to brownish in late summer and persist through most of the winter.

Zanthoxylum fagara (lime prickly ash) is a perennial medium tree (3-6m) with a cylindrical crown from many small, irregularly-shaped branches originating in the hammocks of the American Tropics. Leaves compound, aromatic, about 7-10 cm long. It has greenish-yellow, inconspicuous, fragrant, dioecious flower characteristics also produces orange-brown glandular punctate follicles opening to reveal small, shiny, black seeds.

Data files of six species are provided as vector data in ArcView® *shapefiles* which contains *GT-polygon composed of chains* (see **SDTS (1998)** terminology). Distribution maps of species from shape files are illustrated in Figure 4.4.

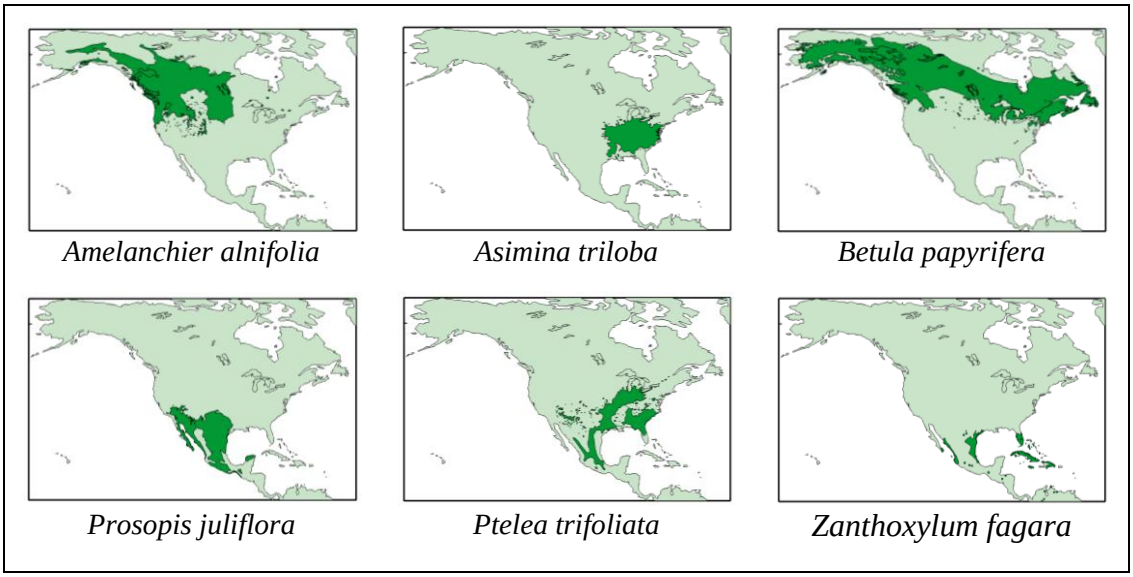


Figure 4.4: Distribution Maps of each species.

4.1.3 Climate Scenarios Data

To evaluate future distribution of species, future values of climate variables must be given to the ANNs after training process is achieved. For this purpose, TYN SC 2.0 data-set produced by **Mitchell and Jones (2005)** from Tyndall Centre for Climate Change Research is used. The TYN SC 2.0 data-set comprises monthly grids of modelled climate, for the period 2001-2100, and covering the global land surface at 0.5° resolution. There are five climatic variables available: cloud cover, diurnal temperature range, precipitation, temperature, vapour pressure. There is one control scenario and 20 climate change scenarios which are made up of all permutations of

five global climate models: CGCM2, CSIRO mk 2, DOE PCM, HadCM3 and ECHam4 with four SRES emissions scenarios: A1FI, A2, B2 and B1 (see Section 2.1)

The dataset, however, is supplied in the form of a set of raw ASCII files which must be unpacked into the scenarios by means of a supplied fortran90 code. When the code is compiled and run, it receives four type of arguments: time period (in the rage of 2001-2100), climate variable such as precipitation, SRES emissions scenario and climate model. Then, with respect to given arguments fortran program produces files which have the format having 5 header lines and global grid-boxes data. The header lines for a global 0.5° grid file look likes as follows:

```
Tyndall Centre grim file created on 11.05.2007 at 18:02 by Dr. Tim Mitchell
.dtr = diurnal temperature range (degrees Celsius)
SRES=A1FI GCM=CGCM2 Period=2050-2050 Variable=.dtr
[Long=-180.00, 180.00] [Lati= -90.00,  90.00] [Grid X,Y= 720, 360]
[Boxes=   67420] [Years=2050-2050] [Multi=    0.1000] [Missing=-999]
```

Each grid box data is given as a line with a grid-reference for the box which holds the position on the x, y axes whose bottom-left corner has the reference: 1, 1 in the grid and also grid box has a second line including 12 integers of 5 digit which represents values of each months. In addition, integer values must be multiplied by the *Multi* number in the headers to obtain correct values. Followings are a couple of lines of from the file.

```
Grid-ref=   1, 148
   79   94   77   82   78   73   77   74   80   84   84   96
Grid-ref=   1, 311
   94   88   91   92   65   78   87   78   79   58   64   88
Grid-ref=   1, 312
   95   89   91   93   65   77   87   77   79   57   65   88
```

4.2 Data Preprocessing

Both 1961-1990 climatology and 2001-2100 scenario data-set files have excess information that is not used in the model. Thus, the global 720×360 grid climate data in each file is clipped to region from longitude 180˚W to 50˚E and from latitude 7˚ to 83˚N (230×76 grid), which are the boundaries of North America.

On the other hand, since the distribution data was vectoral as ArcView® shapefiles, it must be converted to raster data. For this purpose, *vec2mtx* function in

Matlab® Mapping Toolbox is used and shapefiles are converted to 0.5˚ resolution clipped (76×230) grid data files where each grid box contains 0 or 1 values, representing occurance of a species at each grid box.

Another issue to be considered on data is difference between magnitudes of input values coming from different data sets. Input data-set used for training (climatology data) consists of monthly values of mean temperature, diurnal temperature range and precipitation in which minimum / maximum values are −438/341, 36/235, 0/257 respectively. There is a slight difference between values so, to prevent classifier from preferring one of the features over the others, *min-max normalization* carried on the input data set, and data range scaled to [0,1] range by using Equation 4.1.

$$v' = \frac{v - min}{max - min} \tag{4.1}$$

On the other hand, species' distribution data has a characteristic that number of presence occurrences is always much smaller than number of absence occurrences. To address this issue, presence data is randomly reproduced until presences form nearly half of the data. Thus, it is ensured that the network is not trained from too many absences.

All data pre-processing procedures are achieved by Java program. Program processes input and output data files and produces a sample set with 14,951 samples. Each sample has a feature vector with 36 elements representing monthly climate values (12 months×3 climate variables) and a binary target value and representing species occurrence. All samples saved in memory as *HashMap* data structure of Java in which indexes of HashMaps associated with reference values of grid boxes.

4.3 Training Artificial Neural Network

4.3.1 The Network Architecture

An artificial neural network is typically composed of layers of nodes. In the back-propagation network all input nodes are in one input layer and all output nodes are in one output layer while hidden nodes are in one or more hidden layers. Firstly, selection of number of hidden layers arises as first issue. Some theoretical works

such as **Cybenko (1989)** and **Hornik et al. (1989)** show that one hidden layer is sufficient for ANNs to approximate any continuous function (see Section 3.3.2). However, several researchers claim that two hidden layers may be more advantageous for some type of problems (**Barron, 1994**; **Srinivasan et al. 1994**; **Zhang, 1994**). Especially when one hidden layer needs too many hidden nodes, it is better to use two hidden layers with smaller number of hidden nodes which described by **Alpaydın (2000)** as **"**preferring long and narrow networks to short and fat networks". In the following section, investigations on optimal number of hidden nodes expose that the network does not need too many hidden nodes on one hidden layer. Thus, one hidden layer is chosen for the sake of simplicity.

The number of input nodes correlates with the number of features in the input vector of samples. Since samples have 36 features as previously described, the number of input nodes is selected as 36 nodes. Samples have also a binary occurrence value (target value) which corresponds to output layer of the network. **Alpaydın (2000)** states "When there are two classes, one output unit suffices" and we have two classes which are presences and absences so; network is defined with one output node. Consequently, basic architecture of the three-layer network is comprised of 36 units on input layer, one unit on output layer and a hidden layer with parametric number of units (Figure 4.5).

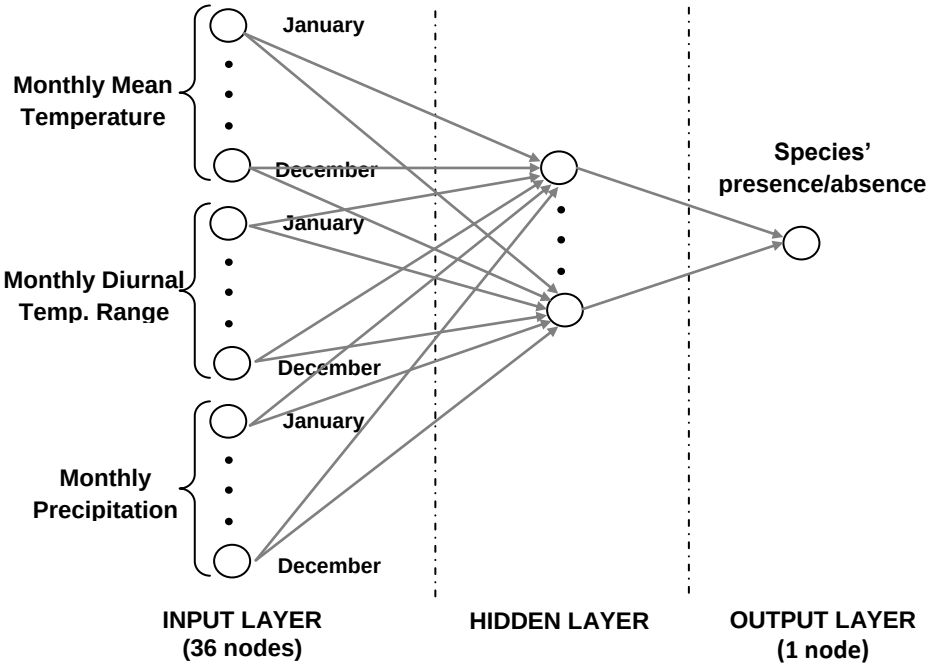


Figure 4.5: Network architecture with 36 input nodes and one output node.

4.3.2 Back-propagation Parameters

Network performance also relies not only on network architecture but also on parameters of learning algorithm. As discussed in section 3.3.2 an activation function should provide some properties and Logistic function, a favourite activation function, is well suited for output nodes of classification problems having binary target values **(Zhang et al., 1998)**. For this reason, Logistic function with $\lambda=1$ (see Equation 3.3) is chosen as activation function in both hidden and output layers.

In addition, weights between nodes should be initialized before training but choosing initial values is important in order to have fast and *uniform learning*, that is, all weights converges to their final values at same speed. **Duda et al. (2000)**, discuss how to choose initial weights: If initial weights are chosen too small, net activation of hidden units become small and hidden layers lose their importance and network turns into linear model. On the other hand, if weights are chosen too large then, hidden units saturate. **Duda et al. (2000)** suggest that initials of input-to-hidden should be in range of $-1/\sqrt{n_I}<w_j<1/\sqrt{n_I}$ and initials of hidden-to-output should be in range of $-1/\sqrt{n_H}<w_k<1/\sqrt{n_H}$ where n_I number of input units and n_H number of hidden units. Thus, weights are randomly initialized within those ranges before training starts.

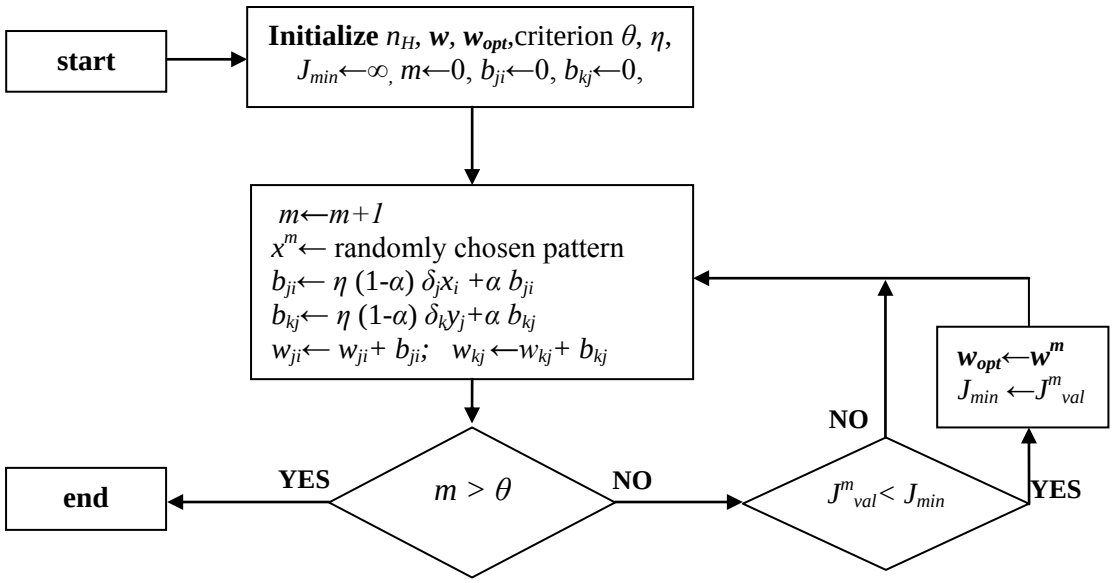


Figure 4.6: Stochastic Back-propagation Algorithm

In training process, selection of samples at each epoch is performed with respect to *stochastic training* protocol (**Duda et al., 2000)**, in which samples are

chosen randomly from training set and weight update carried out after each sample is used. Figure 4.6 illustrates steps taken in the algorithm. The best network state is defined as minimum average error of all samples in the validation set which is measured once at each 100 epochs. Instead of using a criterion associated with error of individual patterns, the algorithm runs for certain number of epochs and the weight set satisfying minimum validation error becomes optimal weights. Thus, the system is prevented from over-training and also algorithm does not stop at a local minimum of error surface.

In order to employ the algorithm illustrated in Figure 4.5, the best values of three critical parameters, which are number of hidden units, learning rate and momentum, must be determined. However, no rule exists to find optimal values of these parameters. For this reason, the only way is to perform an exhaustive training for each combination of discrete values of parameter space. **Sharda and Patil (1992)**, for example, tried nine combinations of three learning rates (0.1, 0.5, 0.9) and three momentum values (0.1, 0.5, 0.9). Similarly, in this study 105 combinations of five learning rates (0.01, 0.05, 0.1, 0.5, 0.9), three momentum values (0.1, 0.5, 0.9) and seven values for number of hidden units (5, 25, 50, 100, 175, 250, 500) are experimented one by one. The network described above is trained three times for each six species with 10,000 epochs for each combination. Then, for each species parameter combination (average value of three experiments of each combination is taken) producing minimum validation error is chosen as the best. Error term used here is quadratic error which is formulated in Equation 3.17. In the algorithm, square of difference between target value and evaluated value gives error of a pattern.

$$J(w) = (t - z)^2 \tag{4.2}$$

When calculating training error or validation error, sum of errors of all individual patterns are divided by the number of patterns in the pattern-set as formulated in Equation 4.3. It is clear that training and validation sets are different pattern sets.

$$J = \frac{1}{n}\sum_{p=1}^{n} J_p \tag{4.3}$$

Experiments described above produce following optimal values presented in the Table 4.2. Moreover, Appendix A (Figure A.1-A.18) contains error plots of each combination.

On the other hand, trying these exhaustive works on a desktop computer is not feasible. For this reason, all computer intensive works in the study are run on HP DL360 G5 Cluster having 1032 Xeon 2.33 Ghz. processors with Infiniband network in National Computing and High Performance Center, Istanbul Technical University.

Table 4.2: Optimal n_H, η and α values obtained from the experimental search.

	n_H	η	α	**Validation Error**	**Training Error**
Amelanchier alnifolia	25	0.9	0.5	0.0854	0.0860
Asimina triloba	50	0.9	0.5	0.0215	0.0207
Betula papyrifera	50	0.9	0	0.0532	0.0542
Prosopis juliflora	25	0.5	0.9	0.0265	0.0285
Ptelea trifoliata	5	0.9	0.9	0.0691	0.0720
Zanthoxylum fagara	25	0.9	0	0.0405	0.0430

After optimal n_H, η and α values are evaluated, full training task is ready to be performed for each species. In the experiments, training is conducted for 10,000 epochs. If training is let continue, validation and training errors go on decreasing and finally reach a horizontal asymptote. Thus, network run is performed with optimal parameters for 100,000 epochs as repeating training three times for each species. Then, optimal weights for each species were selected amongst the three runs that achieve the lowest minimum validation set error. In Table 4.3, minimum training and validation errors reached for each species are given. Also, Figure 4.7illustrates training and validation errors of *Prosopis juliflora* measured at each 100 epochs. The other species' error plots are presented in Appendix B (Figure B.1-B.6).

Table 4.3: Training and validation errors during 100,000 epochs

	Validation Error	**Training Error**
Amelanchier alnifolia	0.0540	0.0504
Asimina triloba	0.0114	0.0140
Betula papyrifera	0.0434	0.0475
Prosopis juliflora	0.0166	0.0161
Ptelea trifoliata	0.0533	0.0554
Zanthoxylum fagara	0.0328	0.0363

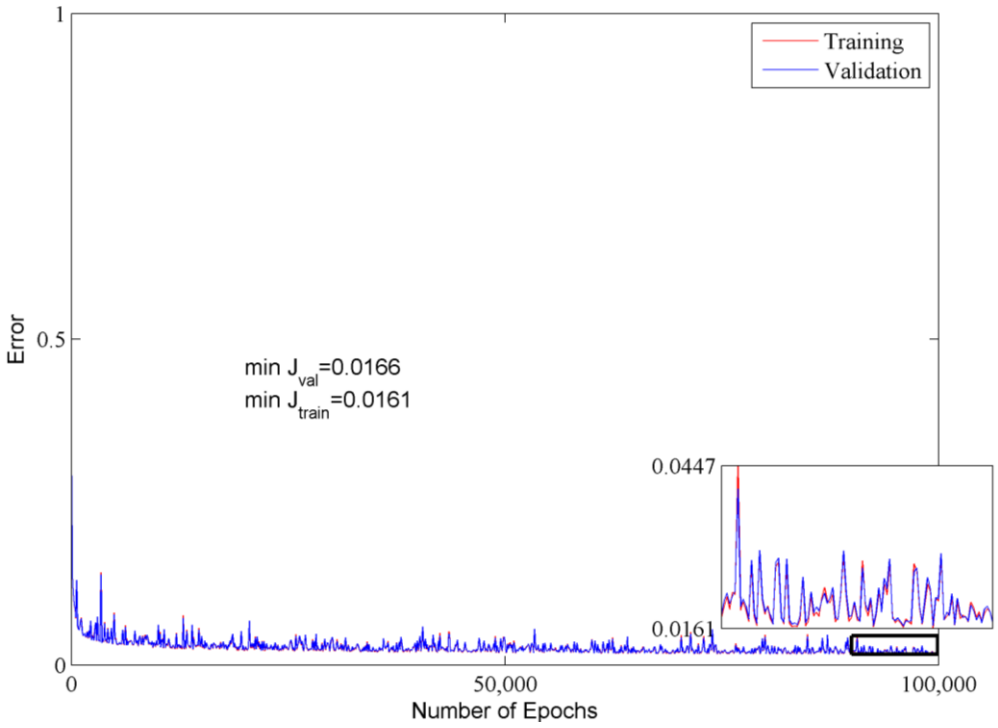


Figure 4.7: Training and validation errors during 100,000 epochs for *Prosopis juliflora*

At last, training step is completed by achieving optimal weights. In following section, optimal network is used for measuring performance with the aid of test set, unseen 10% of data, and finally used for prediction.

5. RESULTS

5.1 Performance Measures

5.1.1 ROC Analysis

The ROC (Receiver Operating Characteristic) technique developed in signal processing that used during World War II for the purpose of radar detection. A radar system is a "receiver" that identifies aircrafts in its range whether they are enemy or not. Its "operating characteristics" is identification of an enemy aircraft while it is indeed an enemy, referred as *true-positive rate* or s*ensitivity*, or identification of non-enemy plane while it is indeed a non-enemy, referred *as true-negative rate* or *specificity* **(Dayhoff and DeLeo, 2001)**. (See Table 5.1 and Equation 5.1-5.2)

Table 5.1: Confusion Matrix

	Predicted Positive (Presence)	**Predicted Negative (Absence)**
Positives	TP	FN
Negatives	FP	TN

$$\text{Sensitivity} = TP/(TP+FN) \tag{5.1}$$

$$\text{Specificity} = FN/(TP+FN) \tag{5.2}$$

A ROC plot is obtained by representing sensitivity values on the *y* axes and (1-specifity) on the *x* axis. This plot can characterize performance of a model by means of distance to upper left corner, coordinate (0,1), where curve of a perfect model reaches. On the other hand, line from bottom left corner to upper right corner is a random classifier that gets half the positives and half the negatives correct. So, curve of a model placing up to the line can be classified as a good model while curve under the line can be classified as poor model. Furthermore, a ROC curve provides a

valuable property: AUC (Area Under the ROC Curve) which assesses the ranking in terms of separation of the classes. A model having AUC value very next to 1.0 provides a good separation between classes.

5.1.2 Cohen's Kappa Statistics

Cohen's Kappa statistics (Equation 5.3) is a commonly used method to assess the level of agreement between observed and predicted data **(Cohen, 1960)**.

$$\kappa = [\ (a+d) - ((a+c)(a+b)+(b+d)(c+d))/N\ ] \ /\ [\ N - ((a+c)(a+b)+(b+d)(c+d))/N\ ] \quad (5.3)$$

where a =TP, b = FP, c = FN, d = TN and N = TP+FP+FN+TN;

Monserud and Leemans (1992) suggest a model assessment presented in Table 5.2.

Table 5.2: Model assessment by using Kappa measure

$\kappa < 0.05$	no agreement
$0.05 < \kappa < 0.20$	very poor
$0.20 < \kappa < 0.40$	Poor
$0.40 < \kappa < 0.55$	Fair
$0.55 < \kappa < 0.70$	Good
$0.70 < \kappa < 0.85$	very good
$0.85 < \kappa < 0.99$	Excellent

5.2 Model Performance and Threshold Selection

Final network configuration with optimal weights is tested with 10% of the data (test data). Output values produced by network is classified with respect to a cut-point threshold value and results including TP, TN, FP, FN numbers are combined in Kappa statistics in order to measure best threshold value and overall model performance. In Table 5.3 Kappa statistics of six species with thresholds in range [0, 1] presented. Range is [0, 1] since network produces output values in that range (see Section 4). Maximum values are marked in the table and they are chosen as cut-point threshold value for prediction. Also, confusion matrices, sensitivity and specificity values for best threshold values are presented in Table 5.4.

Table 5.3: Kappa statistics of each species

Threshold	*Amelanchier alnifolia*	*Asimina triloba*	*Betula papyrifera*	*Prosopis juliflora*	*Ptelea trifoliata*	*Zanthoxylum fagara*
0.05	0.677734	0.951659	0.781720	0.900916	0.796520	0.863653
0.10	0.737926	0.953645	0.810038	0.928254	0.835226	0.878444
0.15	0.763575	0.957623	0.825427	0.937140	0.841531	0.885866
0.20	0.783647	0.961611	0.838624	0.937947	0.850050	0.886664
0.25	0.794367	0.963609	0.844000	0.946860	0.849961	0.890434
0.30	0.814626	0.963609	0.854036	0.950888	0.856622	0.898003
0.35	0.820847	0.964587	0.864168	0.953915	0.860357	0.897810
0.40	0.829494	0.961516	0.867496	0.956922	0.863946	0.900663
0.45	0.836114	0.964521	0.864765	0.955774	0.869717	0.906387
0.50	0.837474	0.965524*	0.870268*	0.958771*	0.872458*	0.910157
0.55	0.846230	0.965502	0.862749	0.952469	0.872230	0.913036*
0.60	0.852063*	0.963450	0.859016	0.950280	0.870569	0.912981
0.65	0.849161	0.962376	0.849209	0.943926	0.866204	0.910719
0.70	0.832718	0.959287	0.830995	0.942764	0.860676	0.903126
0.75	0.811030	0.949954	0.837645	0.936282	0.862860	0.900937
0.80	0.795772	0.940482	0.825199	0.928815	0.704667	0.884181
0.85	0.752727	0.934075	0.789311	0.914860	0.674522	0.859238
0.90	0.642726	0.918026	0.735424	0.898360	0.613968	0.830906
0.95	0.361105	0.833392	0.603065	0.843964	0.295921	0.621644

Table 5.4: Species' prediction success on test data with respect to selected cut-point threshold. *Amelanchier alnifolia (i), Asimina triloba (ii), Betula papyrifera (iii), Prosopis juliflora (iv), Ptelea trifoliate (v), Zanthoxylum fagara (vi).*

Species	Threshold	TP	FP	TN	FN	Sensitivity	Specificity	Kappa
i	0.60	702	93	1130	46	0.938503	0.923957	0.852063
ii	0.50	744	30	1383	4	0.994652	0.978769	0.965524
iii	0.50	693	56	943	55	0.926471	0.943944	0.870268
iv	0.50	726	18	1367	22	0.970588	0.987004	0.958771
v	0.50	729	108	1269	19	0.974599	0.921569	0.872458
vi	0.55	740	80	1378	8	0.989305	0.945130	0.913036

In addition to numeric representation of model performance in Table 5.3, Figure 5.1 illustrates TP, FP, TN and FN values in the map of study region. Orange and yellow points are false predictions. In terms of Kappa criterion, model achieves excellent performance for all species since their kappa values are greater than 0.85. However, Figure 5.1 identifies that model works fine for *Prosopis juliflora*, *Betula papyrifera*, *Amelanchier alnifolia* and *Asimina triloba* species while for *Ptelea*

trifoliate and *Zanthoxylum fagara* species model cannot identify positive samples well.

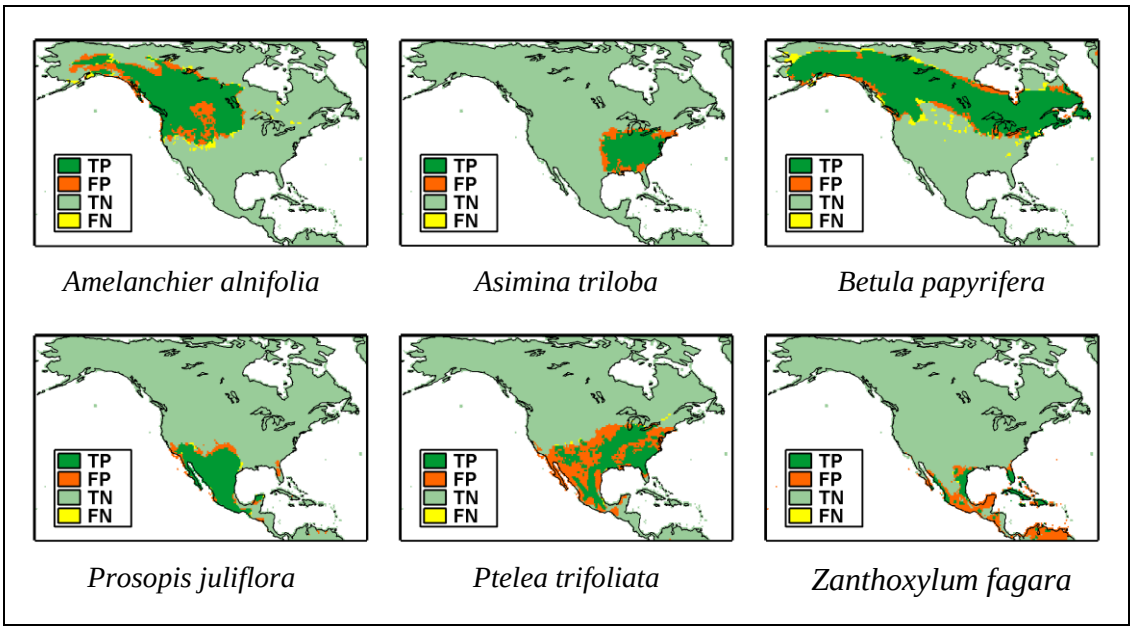


Figure 5.1: Maps of current range of each species predicted by the model.

5.3 Model Comparison

Generally, statistical measurements give an idea about model performance. However, comparison with at least one common technique in the literature provides further assessment about model. For this reason, two common techniques, Generalized Linear Modelling (GLM) and Classification and Regression Trees (CART) are employed to model same problem (See Section 3.1 and 3.2 for details about techniques).

GLM and CART techniques are applied to the data-set with similar data configuration. Complete data set is randomly partitioned into two parts with 70% for training and 30% for testing. Two techniques are applied by means of Matlab® Statistical Toolbox. GLM training in Matlab® is achieved by *glmfit* function with arguments: *binomial* distribution which is ideal for binary classification and *probit* link function. Then, *glmval* function produced outputs for test data. Similarly, CART technique applied by processing training data with *treefit* function of Matlab® Statistical Toolbox and then *treeprune* function is used for tree-pruning and finally, *treeval* function is used to produce output values of testing data.

Six species' output data produced by three models (GLM, CART and ANNs) is used to generate sensitivity and specificity values at threshold values in range [0, 1]. (same as Table 5.2) Then, for each species, sensitivity vs. (1-specificity) pairs of three models generate ROC curves presented in Figure 5.2. AUC and Kappa values of three models for six species are given in Table 5.5 where maximum values for each model per species are marked. Kappa measure expose that ANNs is the best model while AUC measure shows a bit different results. Figure 5.2 implies that ANNs technique is slightly better than GLM while CART technique is showing the lowest performance.

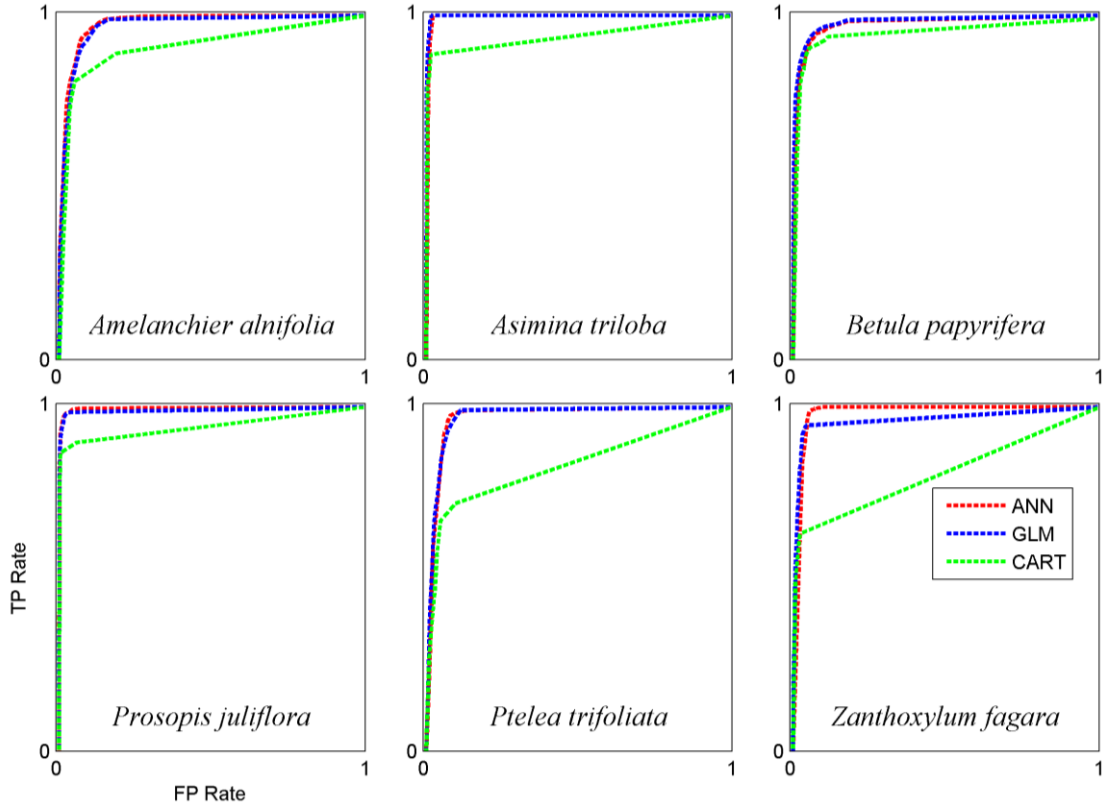


Figure 5.2: ROC curves of each 6 species with three modelling techniques.

Table 5.5: AUC and κ measures of GLM, CART and ANNs models on pilot species

	GLM		CART		ANNs	
	AUC	**κ**	**AUC**	**κ**	**AUC**	**κ**
Amelanchier alnifolia	0.971225	0.772270	0.929885	0.731002	0.975271*	0.852063*
Asimina triloba	0.997920*	0.902110	0.922051	0.827438	0.994727	0.965524*
Betula papyrifera	0.978491*	0.860645	0.955922	0.843425	0.975480	0.870268*
Prosopis juliflora	0.996506*	0.916358	0.952507	0.873997	0.995735	0.958771*
Ptelea trifoliata	0.966915	0.685789	0.914219	0.626359	0.969226*	0.872458*
Zanthoxylum fagara	0.978061	0.666678	0.817363	0.456974	0.980665*	0.913036*

5.4 Future Predictions

After presentation of model performance which shows satisfactory results, prediction for future conditions, in other words classification of species availability (presence vs. absence) in future climatic conditions, is produced from the network. In Figure 4.2 this prediction procedure is illustrated. Future climate data used here contains the best and worst SRES scenarios (A1FI and B1) with CGCM2, DOE PCM and HadCM3 global climate models. These six climate model-scenario combinations are given to the network which was previously trained and tested with the past data. Then, network produces output values which are classified with respect to best cut-point threshold value measured in Section 5.2. This prediction process is repeated for each species for 2050 and 2100 years. Distribution predictions of *Betula papyrifera* for 2100 year are presented in Figure 5.3. Other projections can be found in Appendix C (Figure C.1 - C.6).

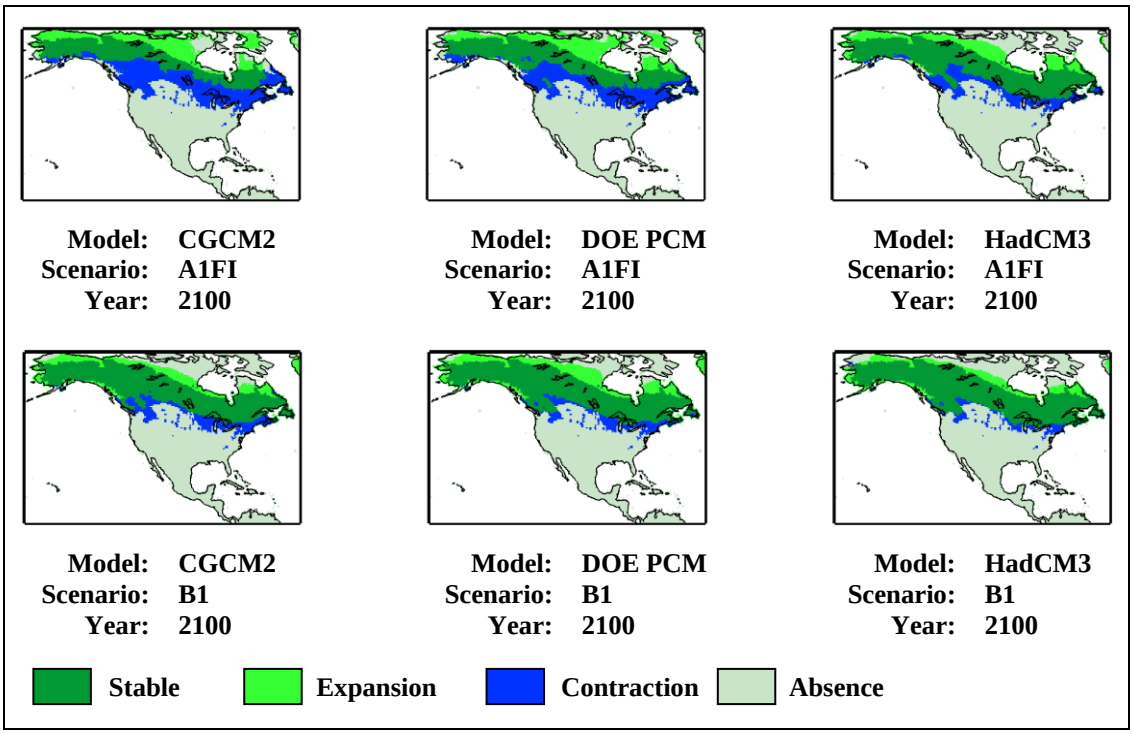


Figure 5.3: 2100 Projections of *Betula papyrifera* with different SRES scenarios and models

6. DISCUSSION AND CONCLUSION

In this study, ANNs technique is employed to evaluate future distributions of six plant species from North America. For this purpose, network having 36 input nodes and one output node is optimized to find the best values of network parameters including number of hidden nodes, learning rate and momentum, by conducting exhaustive training runs. ANNs applications in niche modelling literature, network parameters are not well-investigated and all available researches (e.g. **Manel et al., 1999; Pearson et al., 2002; Thuiller, 2003; Segurado and Araújo, 2004**) apply network parameters, that is claimed to be optimized, to all species in their studies. This study, however, shows importance of parameter selection for each different species by measuring validation errors of dozens combination of the parameters and choosing the combination that achieves minimum validation error per species. Besides, evaluated results compared with GLM and CART techniques and comparisons expose that ANNs technique has better performance as previous studies indicate (**Manel et al., 1999; Thuiller, 2003; Segurado and Araújo, 2004**).

On the other hand, although **Pearson and Dawson (2003)** suggest using climatic parameters on continental scale, effects of the other parameters such as topography, soil types and biotic interactions should be examined. Model performance on *Zanthoxylum fagara* and *Ptelea trifoliata* species is not very high and that motivates a further investigation on environmental parameters. For instance, these two species are native to tropic regions where biotic interactions like competition for resources (water, light, etc.) are extremely intensive. Furthermore, while modelling future distributions of plants, physiological effects of increasing CO_2 concentration in the atmosphere, *fertilisation effect*, is not taken into account **(Catovsky and Bazzaz, 1999)**.

Another difficulty of niche modelling is data availability. A species' high resolution distribution data as time series or abundance quantity is very rarely found and that prevent researches from developing dynamic models. In addition, surveys

representing habitat of animal species just provide presence data and this obstruct to develop classification based models from that type of data.

Future work on this study would be to implement this model on species from Turkey because there are many endangered endemic species. Furthermore, Turkey will be one of the most effected countries by climate change. For this reason, to be able to identify effects of climate change on endangered species will be very important to take all necessary precautions.

REFERENCES

Ackley, D.H., Hinton, G.E. and Sejnowski, T.J., 1985. A Learning Algorithm for Boltzmann Machines, *Cognitive Science*, **9**, 147-169.

Alpaydın, E., 2004. Introduction to Machine Learning, MIT Press, Cambridge, MA.

Anderson, R.P., 2003**.** Real vs. artefactual absences in species distributions: tests for Oryzomys albigularis (Rodentia: Muridae) in Venezuela, *Journal of Biogeography*,**30**, 591-605.

Andres, R.J., Fielding, D.J., Marland, G., Boden, T.A. and Kumar, N., 1999. Carbon dioxide emissions from fossil-fuel use, 1751-1950, *Tellus* 51B:759-765.

Augustin, N.H., Mugglestone, M.A. and Buckland, S.T., 1996. An autologistic model for spatial distribution of wildlife, *Journal of Applied Ecology*, **33**, 339-347.

Alley, R.B., Gow, A.J., Johnsen, S.J., Kipfstuhl, J., Meese, D.A. and Thorsteinsson, T.H., 1995, Comparison of deep ice cores, *Nature* **373,** 393-394.

Alley, R.B., 2000. Ice-core evidence of abrupt climate changes, *Proceedings of the National Academy of Sciences*, **97** (4), 1331-1334.

Anderson, R.P., 2003. Real vs. artefactual absences in species distributions: tests for *Oryzomys albigularis* (Rodentia: Muridae) in Venezuela, J*ournal of Biogeography*, **30**, 591-605.

Anderson, R.P., Lew, D. and Peterson, A.T., 2003. Evaluating predictive models of species' distributions: criteria for selecting optimal models, *Ecological Modelling*, **162**, 211-232.

Araújo, M.B. and Williams, P.H., 2000. Selecting areas for species persistence using occurrence data, *Biological Conservation*, **96**, 331-345.

Barron, A.R., 1994. A comment on "Neural networks: A review from a statistical perspective", *Statistical Science*, **9**, 33–35.

Bio, A.M.F., Alkemande, R. and Barendregt, A., 1998. Determining alternative models for vegetation response analysis - a non-parametric approach, *Journal of Vegetation Science*, **9**, 5-16.

Bradley, R.S. and Jones, P.D., 1995. Recent developments in studies of climate since A.S. 1500, in *Climate Since A.D. 1500 2nd Edition*, pp 666-679 Routledge, London.

Breiman, L., Friedman, J., Olshen R.A. and Stone, C.J., 1984. Classification and regression trees, Wadsworth.

Brohan, P., Kennedy, J.J., Harris, I., Tett, S.F.B. and Jones, P.D., Uncertainty estimates in regional and global observed temperature changes: a new dataset from 1850, *Journal of Geophysical Research*, **111**, D12106.

Carpenter, G.A. and Grossberg, S., 1983. A massively parallel architecture for a self-organizing neural pattem recognition machine, *Computer Vision, Graphics, and Image Processing*, **37**, 54-115.

Carpenter, G.A. and Grossberg, S., 1987. ART 2: Self-organization of stable category recognition codes for Lalog output pattems, *Applied Optics*, **26**, 4919-4930.

Carpenter, G.A. and Grossberg, S., 1988. The ART of adaptive pattem recognition by a self-organizing neural network, *Computer*, 77-88, Mar. 1988.

Carpenter, G.A. and Grossberg, S., 1990. ART 3 hierarchical search Chemical transmitters in self-organizing pattem recognition architectures, in Proc. *Int. Joint Con5 on Neural Networks*, **2**, 30-33.

Catovsky, S. and Bazzaz, F.A., 1999. Elevated CO_2 influences the response of two birch species to soil moisture: implications for forest community structure, *Global Change Biology*, **5**, 507-518.

Cohen J., 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, **20**, 37-46.

Collingham, Y.C., Wadsworth, R.A., Huntley, B. and Hulme, P.E., 2000. Predicting the spatial distribution of non-indigenous riparian weeds: issues of spatial scale and extent, *Journal of Applied Ecology*, **37**, 13-27.

Crowley, T.J. and Lowery, T.S., 2000. How warm was the Medieval warm period?, *Ambio*, **29**, 51-54.

Cumming, G.S., 2000. Using between-model comparisons to fine-tune linear models of species ranges, *Journal of Biogeography*, **27**, 441-455.

Cybenko, G., 1989. Approximation by superpositions of a sigmoidal function, *Mathematical Control Signals Systems*, **2**, 303-314.

Davis, A.J., Jenkinson, L.S., Lawton, J.L., Sharrocks, B. and Wood, B., 1998. Making mistakes when predicting shifts in species range in response to global warming, *Nature*, **391**, 783-786.

Davis B.D. and Shaw R.G., 2001. Range shifts and adaptive responses to Quaternary climate change, *Science*, **292**, 673-679.

Dayhoff, J.E. and DeLeo, J.M., 2001. Artificial neural networks: opening the black box, *Cancer*, **91**, 1615–1635.

Dawson, T.P., Curran, P.J. and Plummer, S.E., 1998. The biochemical decomposition of slash pine needles from reflectance spectra using neural networks, *International Journal of Remote Sensing*, **19** (7), 1433-1438.

De'ath, G. and Fabricius, E., 2000. Classification and regression trees: A powerful yet simple technique for ecological data analysis, *Ecology*, **81**, 3178-3192.

DiMichele ,W.A., Pfefferkorn, W.H. and Gastaldo, A.R., 2001. Response of late Carboniferous and early Permian plant communities to climate change, *Annual Review of Earth and Planetary Sciences*, **29**, 461-487.

Duda, R.O., Hart, P.E. and Stork, D.G., 2000**.** Pattern Classification, A Wiley-Interscience Publication.

Dudik, M., Phillips, S.J., and Schapire, R.E., 2004. Performance guarantees for regularized maximum entropy density estimation, *Proceedings of the 17th Annual Conference on Computational Learning Theory.*

Elton, C., 1927. Animal Ecology, London: Sidgewick & Jackson.

EPICA community members, 2004. Eight glacial cycles from an Antarctic ice core, *Nature* 429:6992, 623-628.

Flato, G., Boer, G.J., Lee, W.G., McFarlane, N.A., Ramsden, D., Reader M.C. and Weaver, A.J., 2000**.** The Canadian Centre for Climate Modelling and Analysis Global Coupled Model and its Climate, *Climate Dynamics*, **16**, 451-468**.**

Franklin, J., 1998. Predicting the distribution of shrub species in southern California from climate and terrain-derived variables, *Journal of Vegetation Science*, 9, 733-748.

Gao, Q., Li, J. and Zheng, H., 1996. A dynamic landscape simulation model for the alkaline grasslands on songnen plain in northeast China, *Landscape Ecology*, **11**, 339-349.

Gause, G. F., 1934. The Struggle for Existence, Baltimore: Williams & Wilkins. Reprinted by Hafner Publishing Company, New York, 1969.

Gordon, H. B. and O'Farrell, S. P., 1997. Transient climate change in the CSIRO coupled model with dynamic sea ice, Monthly Weather Review, **125**, 875-907.

Gordon, C., Cooper, C., Senior, C.A., Banks, H.T., Gregory, J.M., Johns, T.C., Mitchell, J.F.B., and Wood, R.A., 2000. The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. *Climate Dynamics*, **16**, 147-168.

Gratzer, G., Canham, C., Dieckmann, U., Fischer, A., Iwasa, Y., Law, R., Lexer, M.J., Sandmann, H., Spies, T.A., Splechtna, B.E., Szwagrzyk, J., 2004. Spatio-temporal development of forests – current trends in field methods and models, *Oikos*, **107,** 3-15.

Grinnell, J., 1917. The niche relationships of the California Thrasher, *Auk*, **34**, 427-433.

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

Han, J., and Kamber, M., 2001. Classification and prediction, in *Data Mining: Concepts and Techniques*, pp. 279-328, Morgon Kaufmann Publishers.

Hastie, T., Tibshirani, R., 1990. Generalised Additive Models. Chapman and Hall, London.

Haykin, S., 1999. Neural Network - A Comprehensive Foundation, Second Edition, Prentice Hall.

He, H.S., Larsen, D.R. and Mladenoff, D.J., 2002. Exploring component-based approaches in forest landscape modeling, *Environmental Modelling and Software*, **17**, 519-529.

Hebb, D.O., 1949. The Organization of Behavior: A Neuropsychological Theory, New York, John Wiley.

Holmgren, M., Scheffer M., Ezcurra E., Gutierrez J. R. and Mohren G. M. J., 2001. El Nino Effects on the Dynamics of Terrestrial Ecosystems, *Trends in Ecology & Evolution*, **16**, 89-94.

Hopfield, J.J., 1982. Neural networks and physical systems with emergent collective computational abilities, *Proceedings of the National Academy of Sciences*, **79**, 2554-2558.

Hornik, K., Stinchcombe, M. and White, H., 1989. Multilayer feedforward networks are universal approximators, *Neural Networks*, **2**, 359-366.

Huettmann, F. and Diamond, A.W., 2001. Seabird colony locations and environmental determination of seabird distribution: a spatially explicit breeding seabird model for the Northwest Atlantic, *Ecological Modelling*, **141**, 261–298.

Hughes, M.K. and Diaz, H.F., 1994. Was there a medieval warm period, and if so, where and when?, *Climate Change*, **26**, 109−142.

Hulme, M. and Jenkins, G.J., 1998. Climate Change Scenarios for the United Kingdom: Scientific Report, *UK Climate Impacts Programme Technical Report No. 1*, Climatic Research Unit, Norwich, 80 pp.

Hulme, M., Conway, D., Jones, P.D., Jiang, T., Barrow, E.M. and Turney, C., 1995. Construction of a 1961-1990 European climatology for climate change modeling and impact applications, *International Journal of Climatology*, **15**, 1333-1363.

Huntley, B., 1991. How plants respond to climate change: migration rates, individualism and the consequences for plant communities, *Annals of Botany*, **67**, 15-22.

Hutchinson, G. E., 1957. Concluding Remarks, *Cold Spring Symposia on Quantitative Biology*, **22**, 415-427.

IPCC, 2000. Special Report on Emissions Scenarios (SRES): a Special Report of Working Group III of the Intergovernmental Panel on Climate Change. Cambridge University Press. Cambridge.

IPCC, 2001. Climate Change 2001: the Scientific Basis. Cambridge University Press. Cambridge.

IPCC, 2002. Climate Change and Biodiversity—Technical Paper V. Cambridge University Press. Cambridge.

IPCC, 2007. Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, *Intergovernmental Panel on Climate Change*, 5th February, 1-18.

Iverson, L.R., Prasad, A. and Schwartz, M.W., 1999. Modeling potential future individual tree-species distributions in the eastern United States under a climate change scenario: a case study with *Pinus virginiana*, *Ecological Modelling*, **115**, 77-93.

Keeling, C.D., Whorf, T.P., Wahlen, M. and Plicht, J.V., 1995. Interannual extremes in the rate of rise of atmospheric carbon dioxide since 1980, *Nature*, **375**, 666-670.

Kennett, J. P. and Stott, L. D., 1991. Abrupt deep-sea warming, palae-oceanographic changes and benthic extinctions at the end of the Palaeocene. *Nature*, **353**, 225-229.

Kirkpatrick, S., Gelatt Jr., C.D. and Vecchi, M.P., 1983. Optimization by Simulated Annealing, *Science*, **220**, 671-680.

Kohonen, T., 1982. Self-organized formation of topologically correct feature maps, *Biological Cybernetics*, **43**, 59-69.

Kohonen, T., 1988. An introduction to neural computing, *Neural Networks*, **1**, 4.

Kolmogorov, A.N., 1957. On the representation of continuous functions of several variables by superposition of continuous functions of one variable and addition, *Doklady Akademiia Nauk SSSR*, **114**, 953-956.

Leathwick, J.R. and Whitehead, D., 2001. Soil and atmospheric water deficits and the distributions of New Zealand's indigenous tree species, *Functional Ecology*, **15**, 233-242.

Leibold, M.A., 1995. The niche concept revisited: mechanistic models and community context, *Ecology*, **76**, 1371-1382**.**

Lek, S., Delacoste, M., Baran, P., Dimopoulos, I., Lauga, J. and Aulagnier, S., 1996. Application of neural networks to modelling nonlinear relationships in ecology, *Ecological Modelling*, **90**, 39-52.

Levins, R., 1966. The strategy of model building in population biology, *American Scientist*, **54**, 421-431.

Lim, B.K., Peterson, A.T. and Engstrom, M.D., 2002. Robustness of ecological niche modeling algorithms for mammals in Guyana, *Biodiversity and Conservation*, **11**, 1237-1246.

Loiselle, B. A., Howell, C. A., Graham, C. H., Goerck, J. M., Brooks, T., Smith, K. G. and Williams, P. H., 2003.Avoiding pitfalls of using species distribution models in conservation planning, *Conservation Biology*, **17**, 1591-1600.

Malanson, G.P., Westman, W.E. and Yan, Y.-L., 1992. Realized versus fundamental niche functions in a model of chaparral response to climatic change, *Ecological Modelling*, **64**, 261-277.

Manel, D., Dias, J.M., Buckton, S.T. and Ormerod, S.J., 1999. Alternative methods for predicting species distribution: an illustration with Himalayan river birds, *Journal of Applied Ecology*, **36**, 734-747.

Mastrorillo, S., Lek, S., Dauba, F. and Belaud, A., 1997. The use of artificial neural networks to predict the presence of small-bodied fish in a river, *Freshwater Biology*, **38**, 237-246.

McCulloch, W.S. and Pitts,W.A., 1943. A logical calculus of the ideas imminent in nervous activity, *Bulletin of Mathematics and Biophysics*, **5**, 115-133.

Michie, D., Spiegelhalter, D.J. and Taylor, C.C., 1994. Machine Learning, Neural and Statistical Classification. Ellis Horwood Limited.

Minsky, M. and Papert, S., 1969. Perceptrons, Cambridge, MA: MIT Press.

Mitchell, J. F. B. and Johns, **T. C.,** 1997. On modification of global warming by sulphate aerosols, *Journal of Climate* **10**, 245-267.

Mitchell, J. F. B., Johns, T. C., Gregory, J. M. and Tett, S., 1995. Climate response to increasing levels of greenhouse gases and sulphate aerosols, *Nature*, **376**, 501-504.

Mitchell, aT.D. and Jones, P.D., 2005. An improved method of constructing a database of monthly climate observations and associated high-resolution grids, *Int. J. Climatol*., **25,** 693-712.

Monserud, R. A. and Leemans, R., 1992. Comparing global vegetation maps with Kappa statistic, *Ecological Modelling*, **62**, 275-293.

Nelder J. A. and Wedderburn R. W. M., 1972. *J. Roy. Statist. Soc. Ser. A.*, **135**, 370-384

North Greenland Ice Core Project Members, 2004. High-resolution record of Northern Hemisphere climate extending into the last interglacial period, *Nature*, **431**, 147-151.

Olden, J.D., Jackson, D.A., 2001. Fish-habitat relationships in lakes: gaining predictive and explanatory insight by using artificial neural networks, *Trans. Am. Fish. Soc.*, **130**, 878–897.

Özesmi, S.L. and Özesmi, U, 1999. An artificial neural network approach to spatial habitat modelling with interspecific interaction, *Ecological Modelling*, **116**, 15-31.

Parmesan, C. and 12 others, 1999. Poleward shifts in geographical ranges of butterfly species associated with regional warming, *Nature*, **399**, 579-583.

Pearce, J. and Ferrier, S., 2000. An evaluation of alternative algorithms for fitting species distribution models using logistic regression, *Ecological Modelling*, **128**, 127-147.

Pearson, R.G., Dawson, T.P., Berry, P.M. and Harrison, P.A., 2002. SPECIES: a spatial evaluation of climate impact on the envelope of species, *Ecological Modelling*, **154**, 289-300.

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

Peterson, A.T., 2001. Predicting species' geographic distributions based on ecological niche modeling, *The Condor*, 103, 599-605.

Peterson, A.T. and Cohoon, K.P., 1999. Sensitivity of distributional prediction algorithms to geographic data completeness, *Ecological Modelling*, **117**, 159-164.

Peterson, A.T., Stockwell, D.R.B. and Kluza, D.A., 2002. Distributional prediction based on ecological niche modeling of primary occurrence data. In: Predicting Species Occurrences: Issues of Accuracy and Scale. Editors:

Petit J.R. and 18 others, 1999. Climate and Atmospheric History of the Past 420,000 years from the Vostok Ice Core, Antarctica, *Nature*, **399**, 429-436.

Phillips, S.J., Dudik, M. and Schapire, R.E., 2004. A maximum entropy approach to species distribution modeling, in: *Proceedings of the 21st International Conference on Machine Learning*, ACM Press, New York, pp. 655-662.

Phillips, S.J.,, **Anderson, R.P. and Schapire, R.E.,** 2006. Maximum entropy modeling of species geographic distributions, *Ecological Modelling*, **190**, 231-259

Quinlan, J. R., 1986. Induction of decision trees, *Machine Learning*, **1**, 81-86.

Roeckner, E., J.M. Oberhuber, A. Bacher, M. Christoph and I. Kirchner, 1996**.** ENSO variability and atmospheric response in a global coupled atmosphere-ocean GCM, *Clim. Dyn*., **12**, 737-754.

Rohde, A.R., 2005. Global Warming Art, retrieved July 17, 2007 from http://www.globalwarmingart.com

Rosenblatt, F., 1958. The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain, Cornell Aeronautical Laboratory, *Psychological Review*, **6**, 386-408.

Rumelhart, D.E., Hinton, G.E. and Williams, R.J., 1986a. Learning representations by backpropagating errors, *Nature*, 323, 533-536.

Rumelhart, D.E., Hinton, G.E. and Williams, R.J., 1986b. Learning internal representations by error propagation, in *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Volume I: Foundations*, Cambridge, MA: MIT Press, pp. 318-362.

Rutherford, M.C., Callaghan, M.O., Hurford, J.L., Powrie, L.W. and Schulze, R.E., 1995. Realized niche spaces and functional types: a framework for prediction of compositional change, *J. Biogeogr*, **22**, 523-531.

Scott, J.M., Heglund, P.J., Samson, F., Haufler, J., Morrison, M., Raphael, M. and Wall, B., 2002. Predicting Species Occurrences: Issues of Accuracy and Scale, Island Press, Covelo, California.

SDTS, 1998. Spatial Data Transfer Standard by U.S. Geological Survey, NCITS 320-1998, *American National Standards Institute (ANSI)*, Reston, VA.

Segurado, P., Araújo M.B., 2004. Evaluation of methods for modelling species probabilities of occurrence, *Journal of Biogeography*, **31**, 1555– 1568.

Sharda, R. and Patil, R.B., 1992. Connectionist approach to time series prediction: An empirical test, *Journal of Intelligent Manufacturing*, **3**, 317–323.

Srinivasan, D., Liew, A.C. and Chang, C.S., 1994. A neural network short-term load forecaster. *Electric Power Systems Research*, **28**, 227–234.

Skapura, D.M., 1996. Building Neural Networks, ACM Press, New York.

Stewart, R.H., 2005. Importance of the Deep Circulation, Chapter 13.1, Retrieved July 17, 2007, from http://oceanworld.tamu.edu/resources/ocng_textbook

Stockwell, D. and Peters, D., 1999. The GARP modelling system: problems and solutions to automated spatial prediction, *International Journal of Geographical Information Science*, **13**, 143–158.

Stockwell, D.R.B. and Peterson A.T., 2002. Effects of sample size on accuracy of species distribution models, *Ecological Modelling*, **148**, 1-13.

Stuiver, M. and Grootes, P. M., 2000. GISP2 oxygen isotope ratios, *Quat. Res.* **53**, 277-284.

Sun Microsystems, 1995. The Java Language Environment: A White Paper, available from "http://javasoft.com/whitePaper/java-whitepaper-1.html.

Thomas, C. D. and Lennon, J. J., 1999. Birds extend their ranges northwards, *Nature* **399**, 213.

Thuiller, W., 2003. BIOMOD - Optimising predictions of species distributions and projecting potential future shifts under global change, *Global Change Biology*, **9**, 1353-1362.

Thuiller, W., Araújo, M.B., and Lavorel, S., 2003. Generalized models vs. classification tree analysis: Predicting spatial distributions of plant species at different scales, *Journal of Vegetation Science*, **14**, 669-680.

Urban D.L., Bonan G.B., Smith T.M. and Shugart H.H., 1991. Spatial applications of gap models, *Forest Ecology and Management*, **42**, 95-110.

U.S. Geological Survey, 1999. *Digital representation of Atlas of United States Trees* by Elbert L. Little, Jr. available at http://esp.cr.usgs.gov/data/atlas/little/

Vayssières, M.P., Plant, R.E. and Allen-Diaz, B.H., 2000. Classification trees: an alternative non-parametric approach for predicting species distributions, *Journal of Vegetation Science*, **11**, 679-694.

Walker, P.O., 1990. Modelling wildlife distribution using a geographic information system: kangaroos in relation to climate, *Journal of Biogeography*, **17**, 279-289.

Washington, W.M. and 10 others, 2000. Parallel Climate Model (PCM): Control and Transient simulations, *Climate Dynamics*, **16**, 755-774.

Widrow, B. and Hoff, M.E., 1960. Adaptive switching circuits, WESCON *Convention*, Record Part IV, 96-104, 1969.

Willis, K.J. and Whittaker, R.J., 2002. Species diversity -scale matters, *Science*, **295**, 1245-1248.

Whittaker, R.J., Willis, K.J. and Field, R., 2001. Scale and species richness: towards a general, hierarchical theory of species diversity, *Journal of Biogeography*, **28**, 453-470.

Woodward, F.I. and Rochefort, L., 1991. Sensitivity analysis of vegetation diversity to environmental change, *Global Ecology and Biogeography Letters* **1**, 7-23.

Wu, J. and Loucks, O.L., 1995. From balance of nature to hierarchical patch dynamics: a paradigm shift in ecology, *The Quarterly Review Biology*, **70**, 439-466.

Yee, T.W., Mitchell, N.D., 1991. Generalized additive models in plant ecology, *Journal of Vegetation Science*, **2**, 587-602.

Zhang, J., Walter, G.G., Miao, Y. and Wayne, W.N., 1995. Wavelet neural networks for function learning, *IEEE Transactions on Signal Processing*, **43**, 1485-1497.

Zhang, G., Patuwo, B.E. and Hu, M.Y., 1998. Forecasting with artificial neural networks: The state of the art, *International Journal of Forecasting*, **14**, 35–62.

APPENDIX A

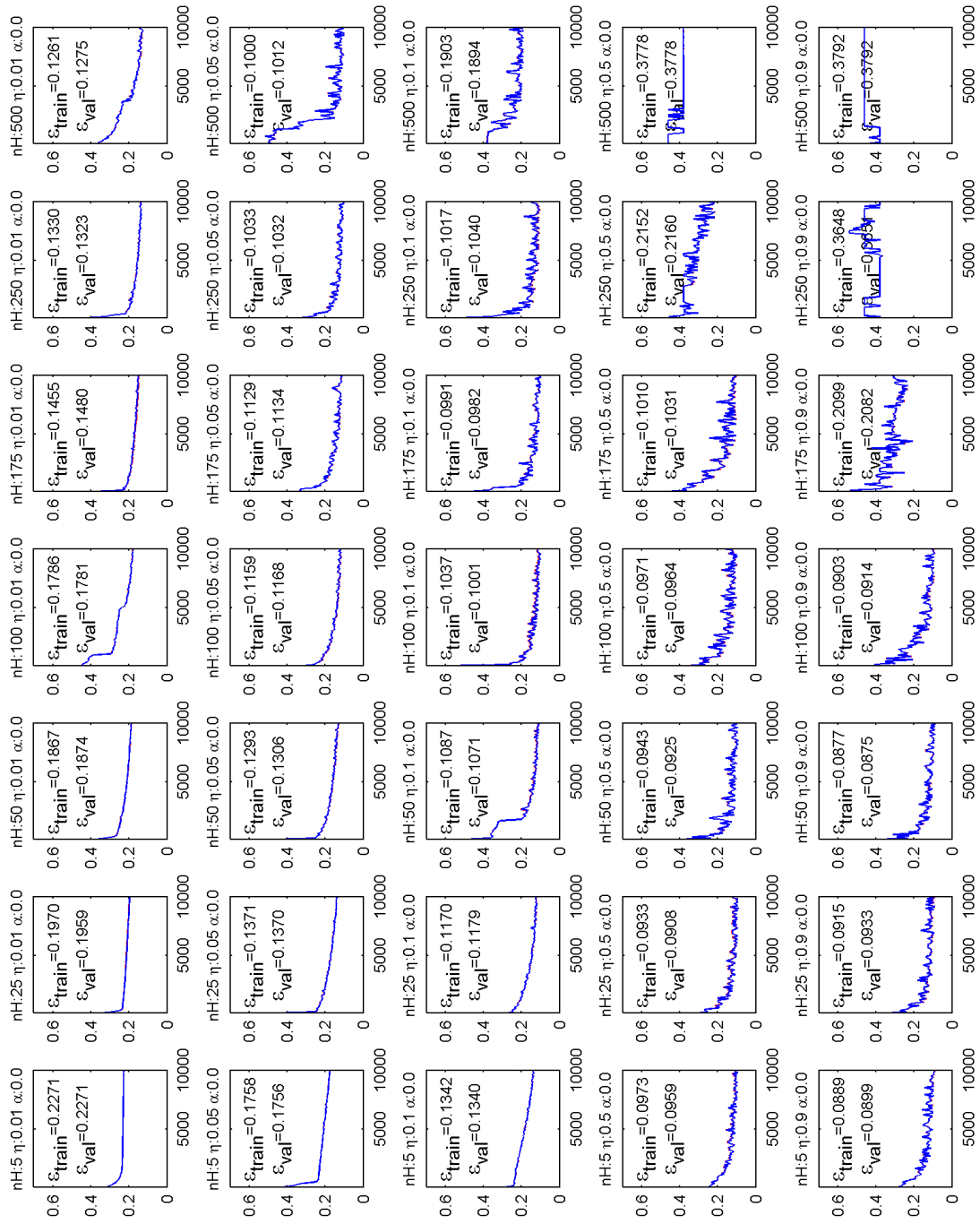


Figure A.1: Training and validation errors during 10000 epochs of *Amelanchier alnifolia* for number of hidden units: [5, 25, 50, 100, 175, 250, 500], learning rate:[0.01, 0.05, 0.1, 0.5, 0.9] and momentum: 0 .

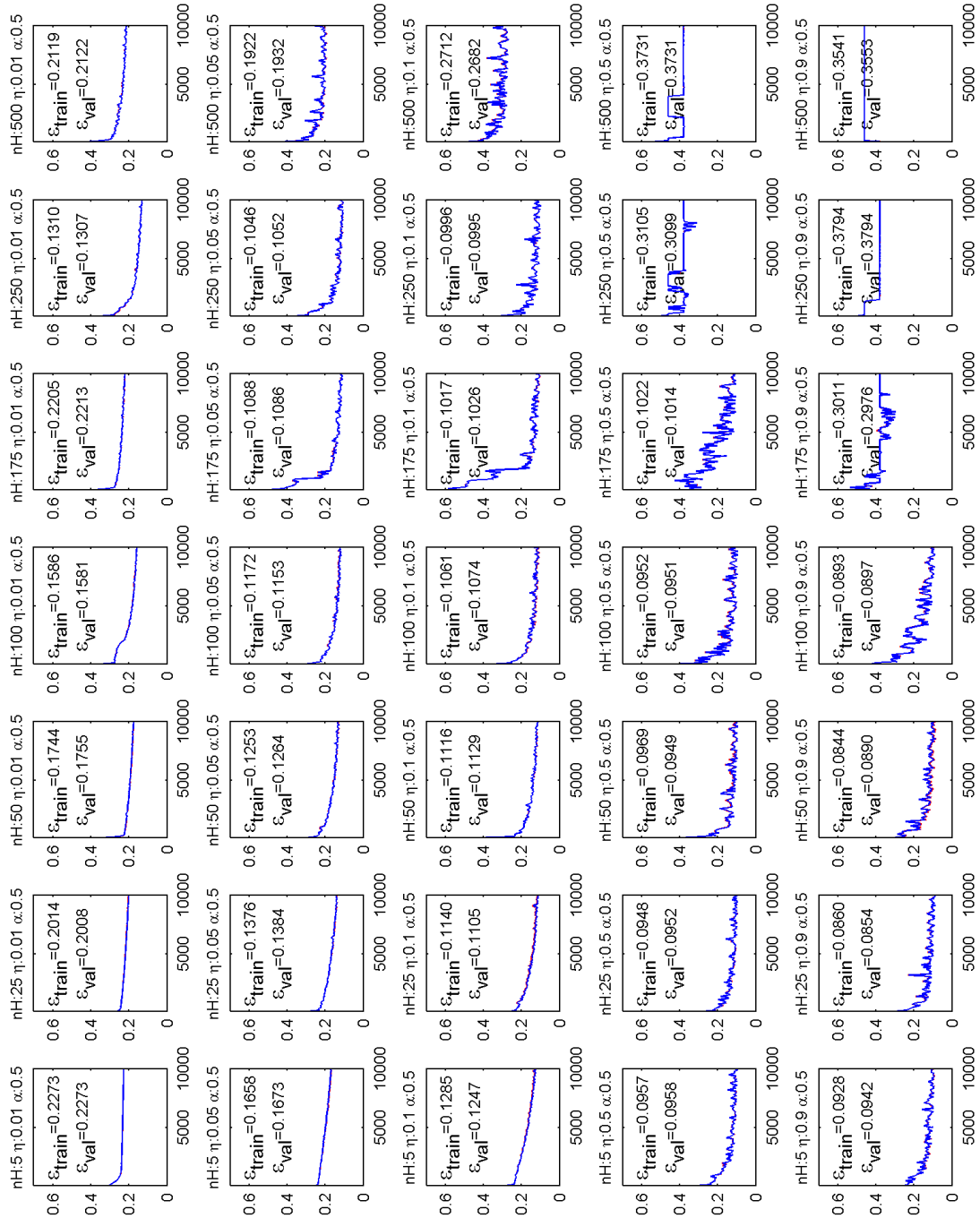


Figure A.2: Training and validation errors during 10000 epochs of *Amelanchier alnifolia* for number of hidden units: [5, 25, 50, 100, 175, 250, 500], learning rate:[0.01, 0.05, 0.1, 0.5, 0.9] and momentum: 0.5.

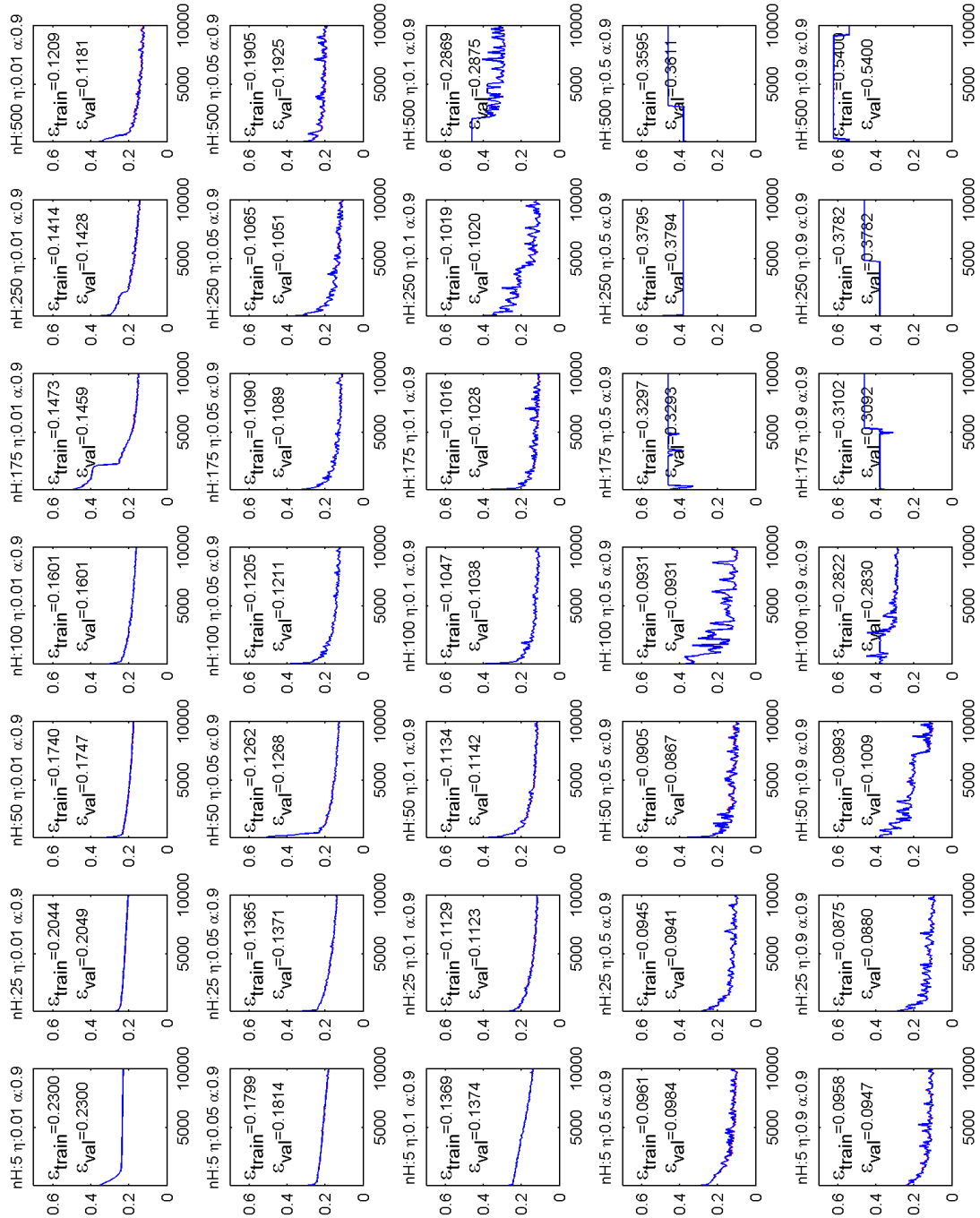


Figure A.3: Training and validation errors during 10000 epochs of *Amelanchier alnifolia* for number of hidden units: [5, 25, 50, 100, 175, 250, 500], learning rate:[0.01, 0.05, 0.1, 0.5, 0.9] and momentum: 0.9.

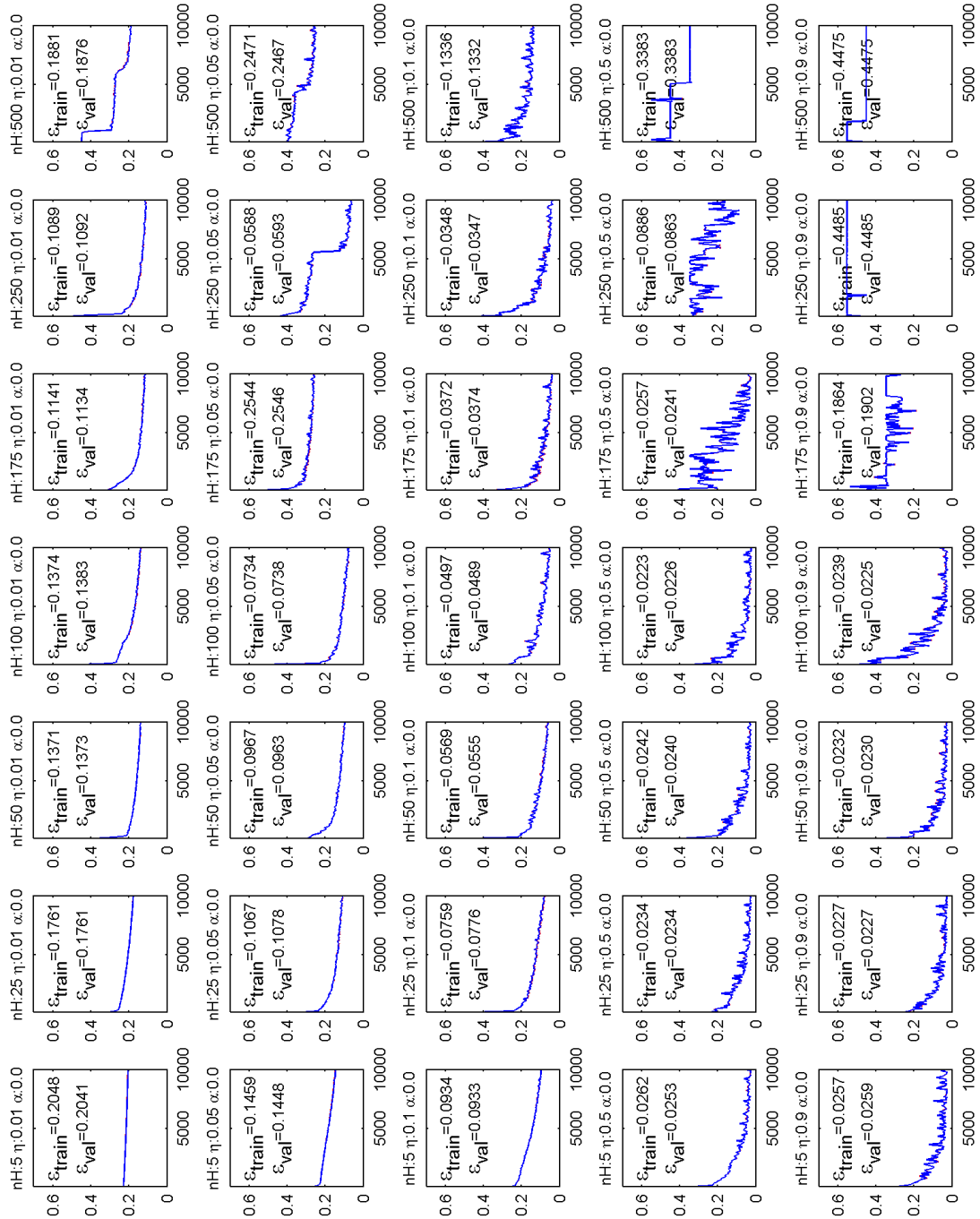


Figure A.4: Training and validation errors during 10000 epochs of *Asimina triloba* for number of hidden units: [5, 25, 50, 100, 175, 250, 500], learning rate:[0.01, 0.05, 0.1, 0.5, 0.9] and momentum: 0.

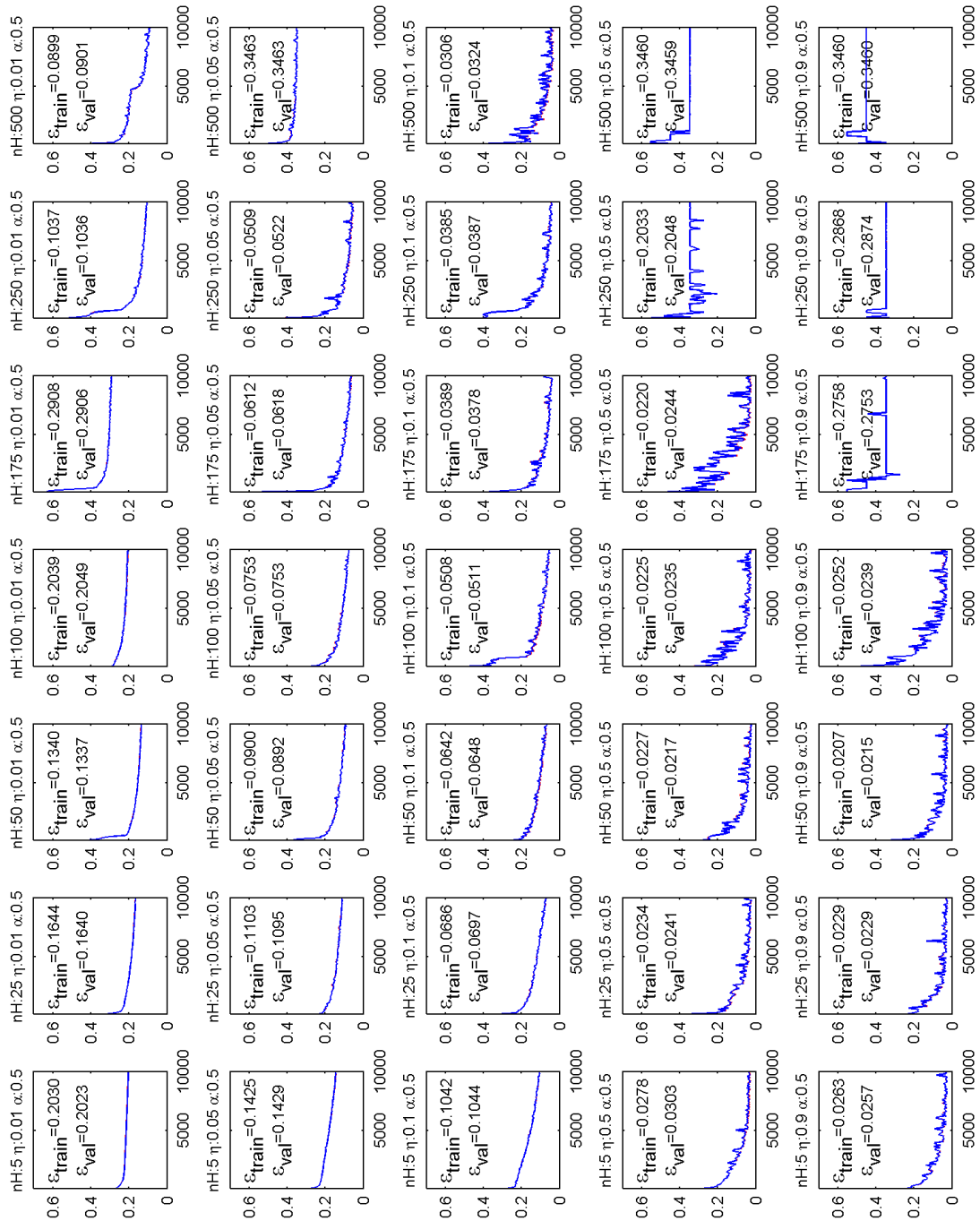


Figure A.5: Training and validation errors during 10000 epochs of *Asimina triloba* for number of hidden units: [5, 25, 50, 100, 175, 250, 500], learning rate:[0.01, 0.05, 0.1, 0.5, 0.9] and momentum: 0.5.

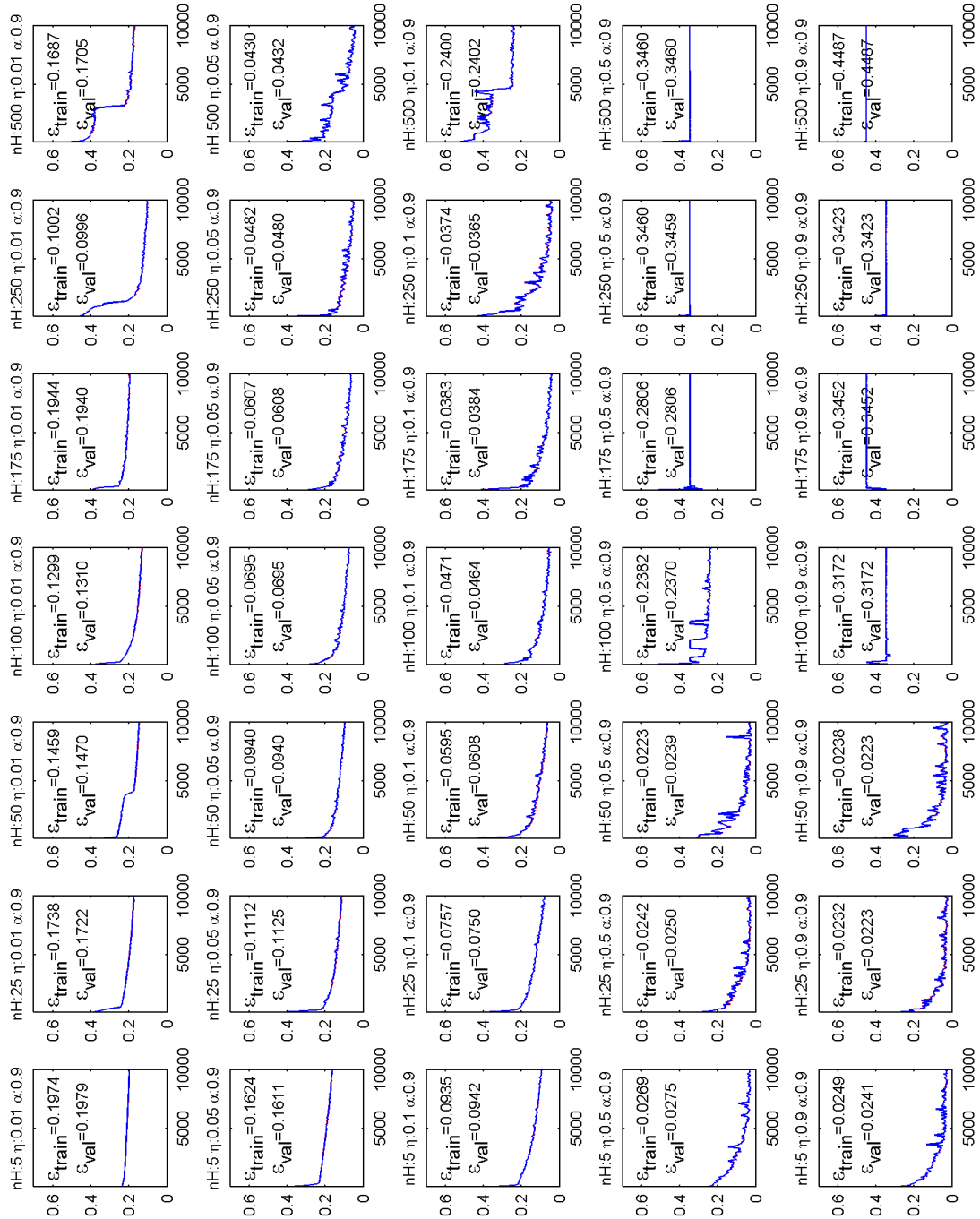


Figure A.6: Training and validation errors during 10000 epochs of *Asimina triloba* for number of hidden units: [5, 25, 50, 100, 175, 250, 500], learning rate:[0.01, 0.05, 0.1, 0.5, 0.9] and momentum: 0.9.

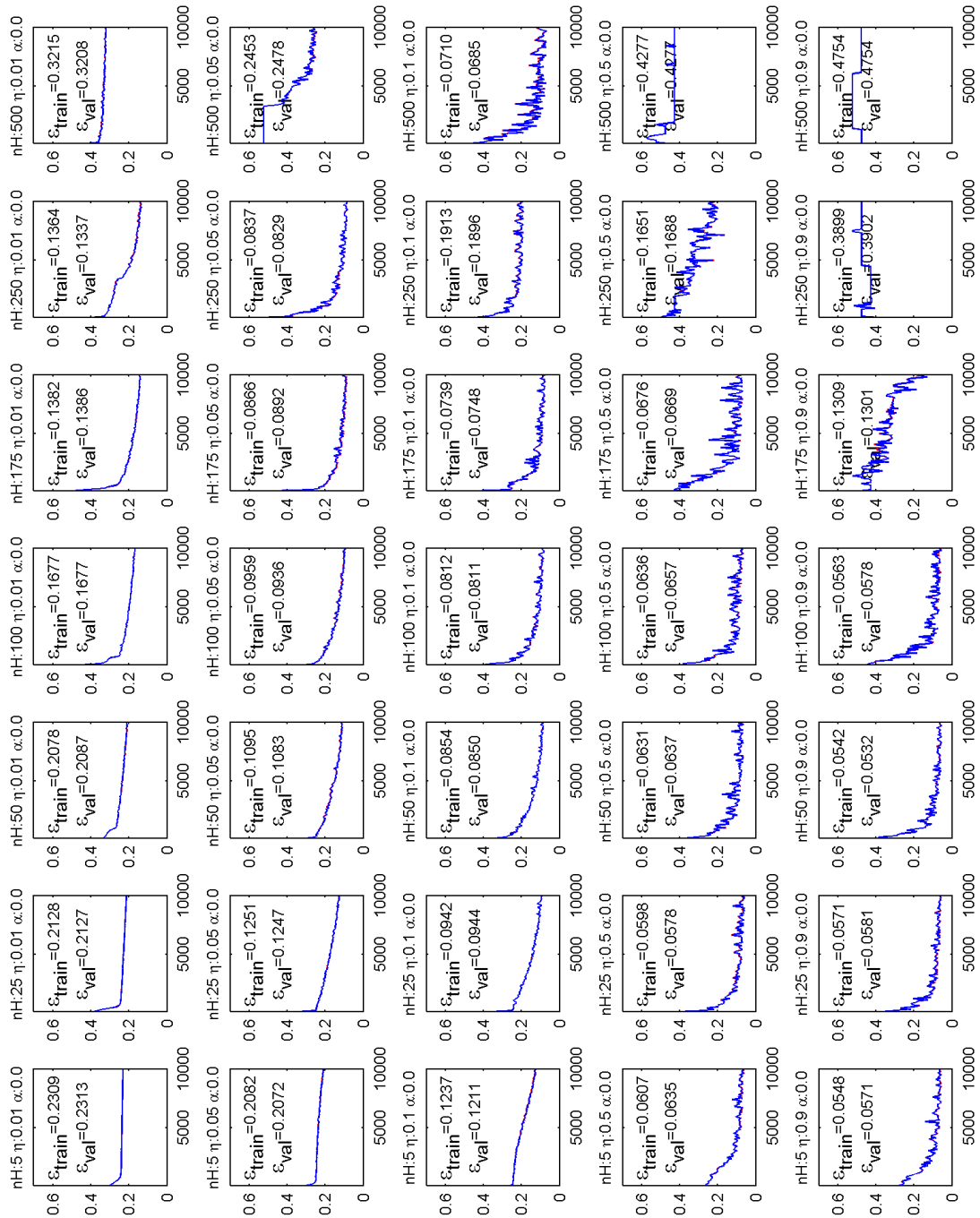


Figure A.7: Training and validation errors during 10000 epochs of *Betula papyrifera* for number of hidden units: [5, 25, 50, 100, 175, 250, 500], learning rate:[0.01, 0.05, 0.1, 0.5, 0.9] and momentum: 0.

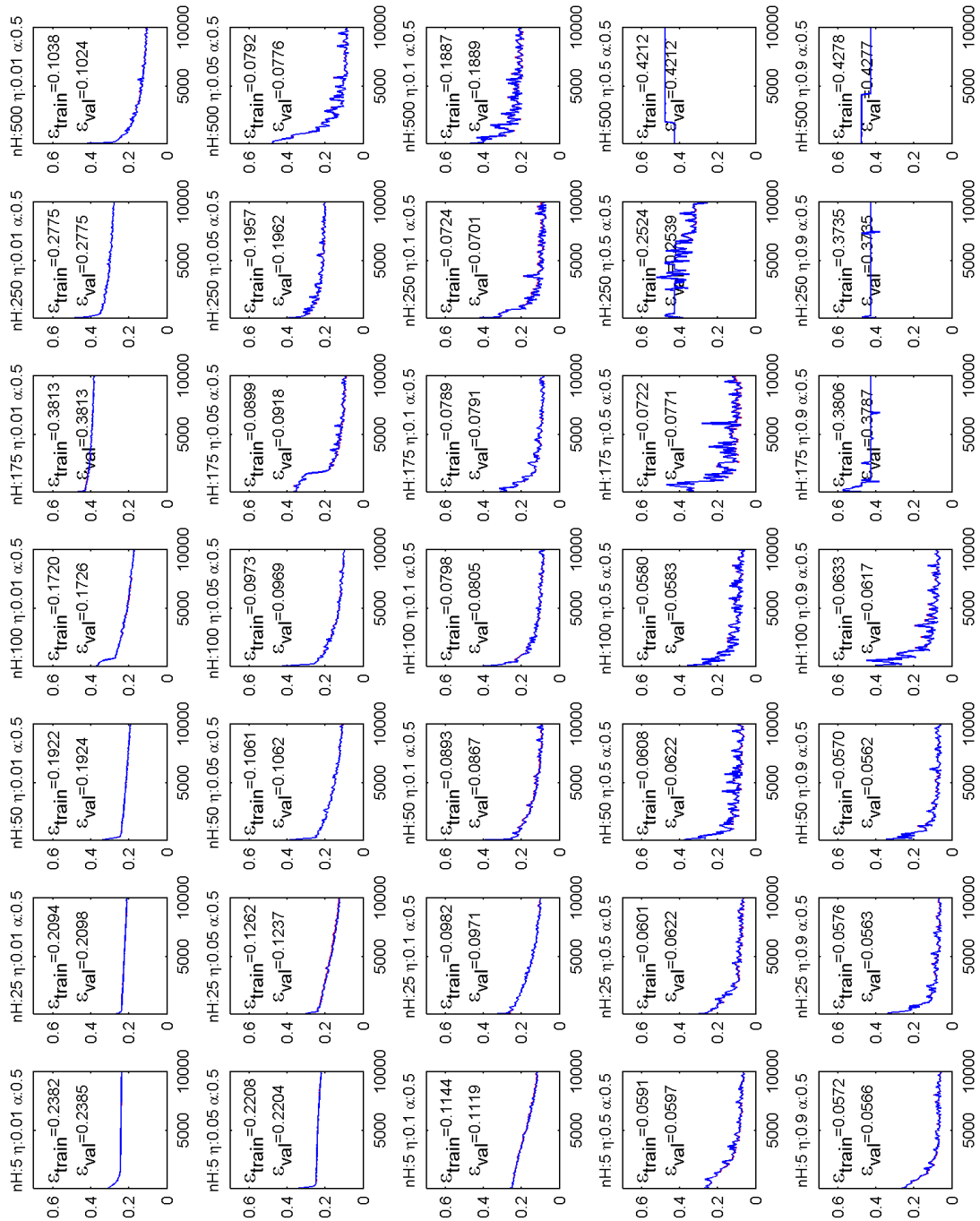


Figure A.8: Training and validation errors during 10000 epochs of *Betula papyrifera* for number of hidden units: [5, 25, 50, 100, 175, 250, 500], learning rate:[0.01, 0.05, 0.1, 0.5, 0.9] and momentum: 0.5.

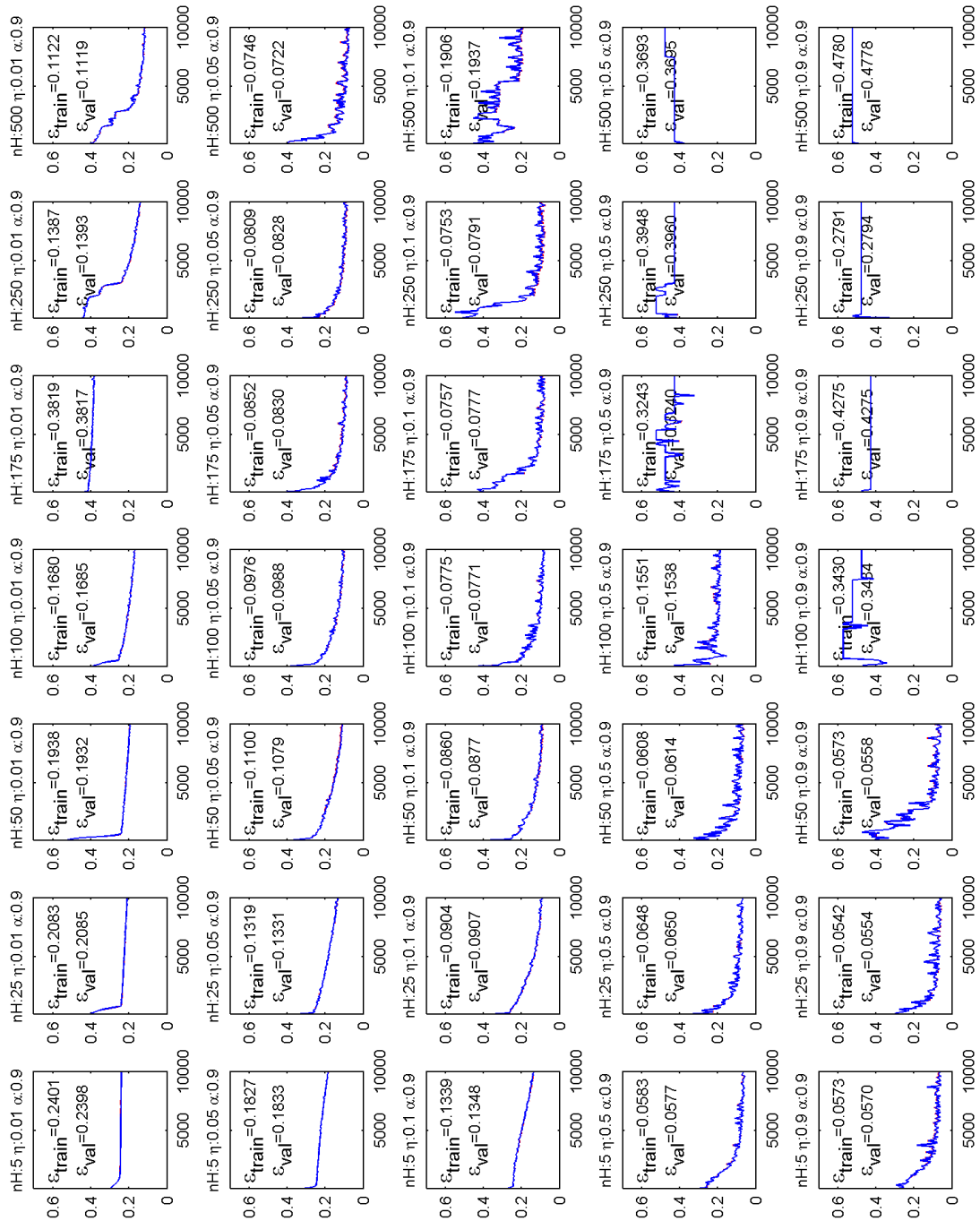


Figure A.9: Training and validation errors during 10000 epochs of *Betula papyrifera* for number of hidden units: [5, 25, 50, 100, 175, 250, 500], learning rate:[0.01, 0.05, 0.1, 0.5, 0.9] and momentum: 0.9.

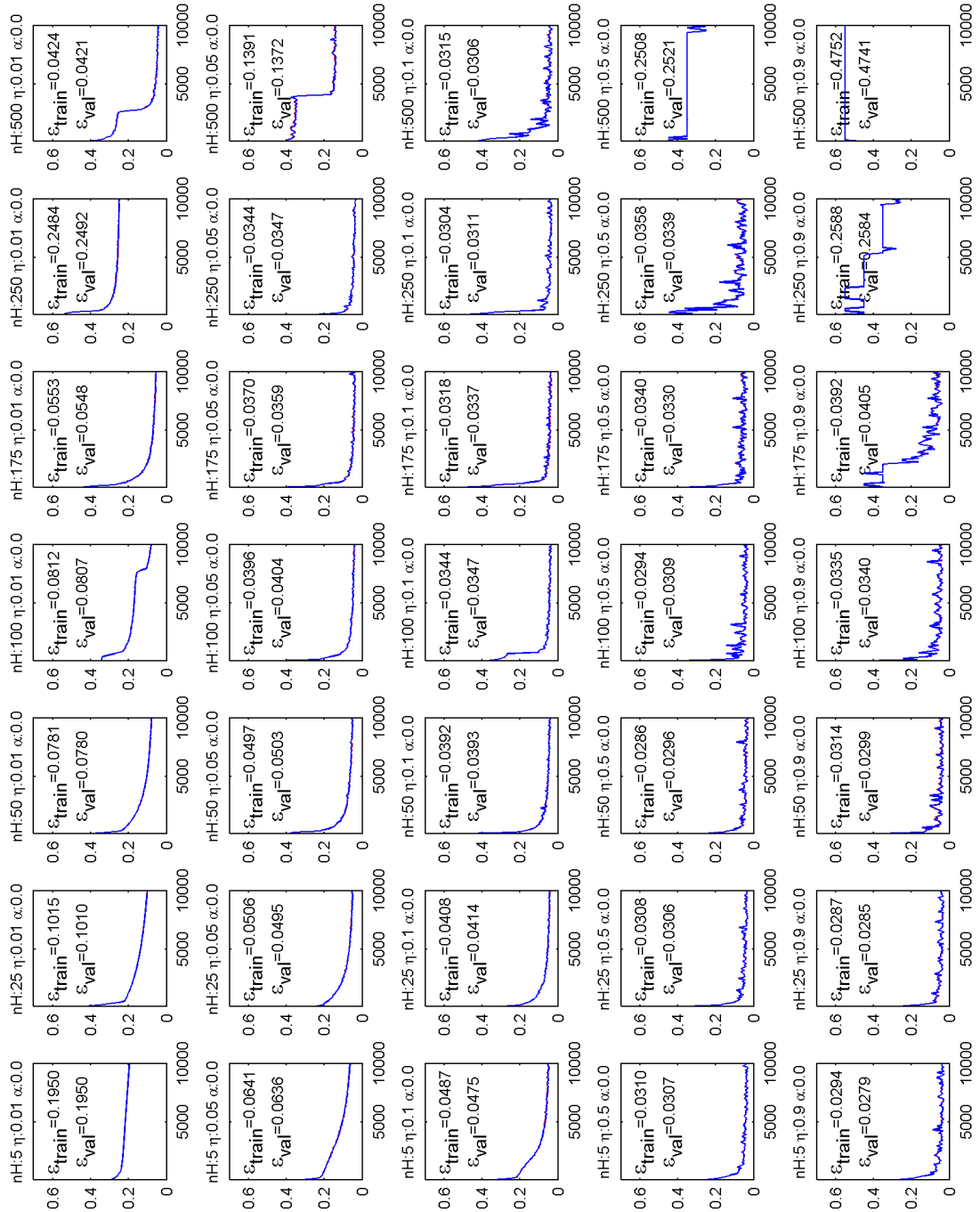


Figure A.10: Training and validation errors during 10000 epochs of *Prosopis juliflora* for number of hidden units: [5, 25, 50, 100, 175, 250, 500], learning rate:[0.01, 0.05, 0.1, 0.5, 0.9] and momentum: 0.

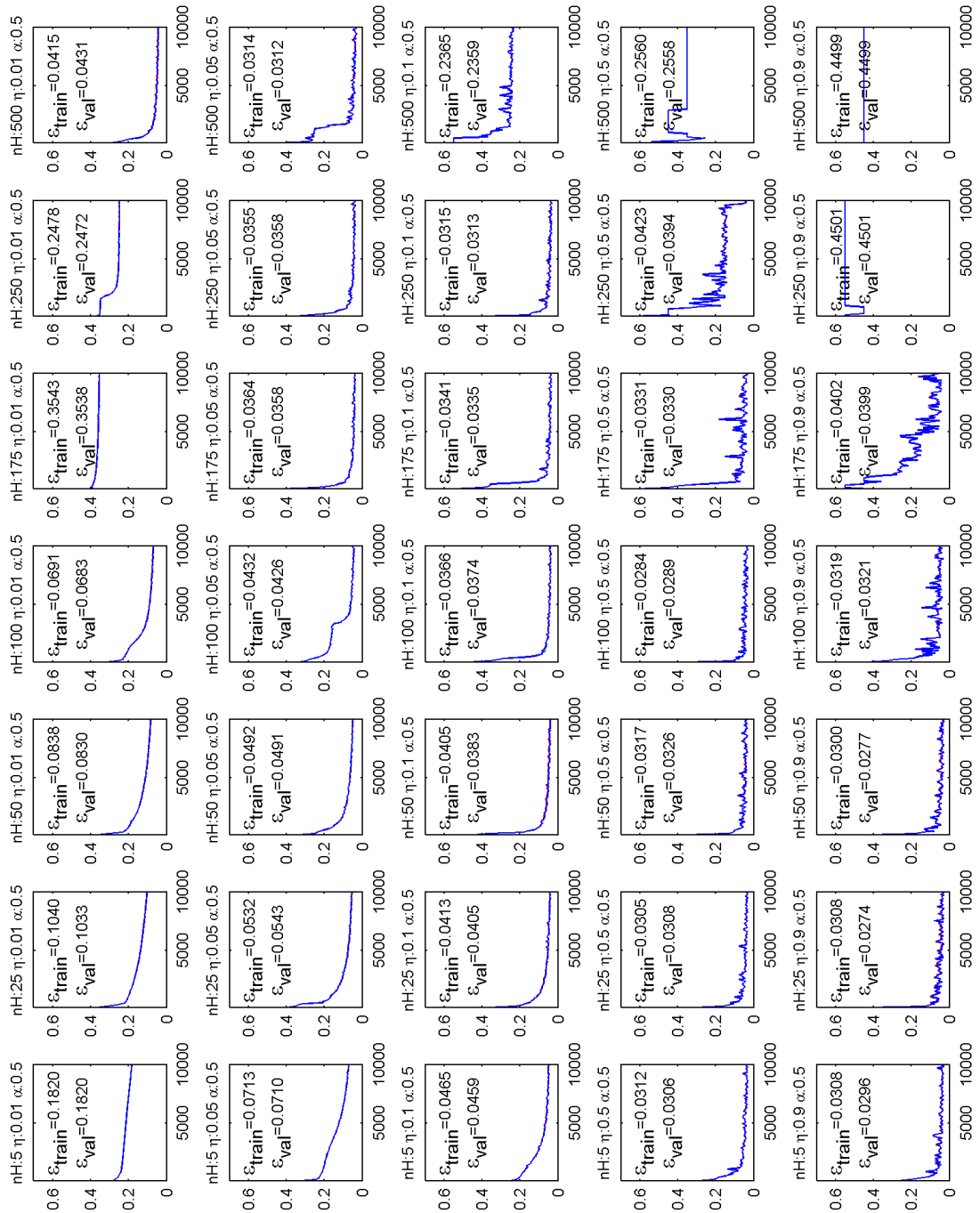


Figure A.11: Training and validation errors during 10000 epochs of *Prosopis juliflora* for number of hidden units: [5, 25, 50, 100, 175, 250, 500], learning rate:[0.01, 0.05, 0.1, 0.5, 0.9] and momentum: 0.5.

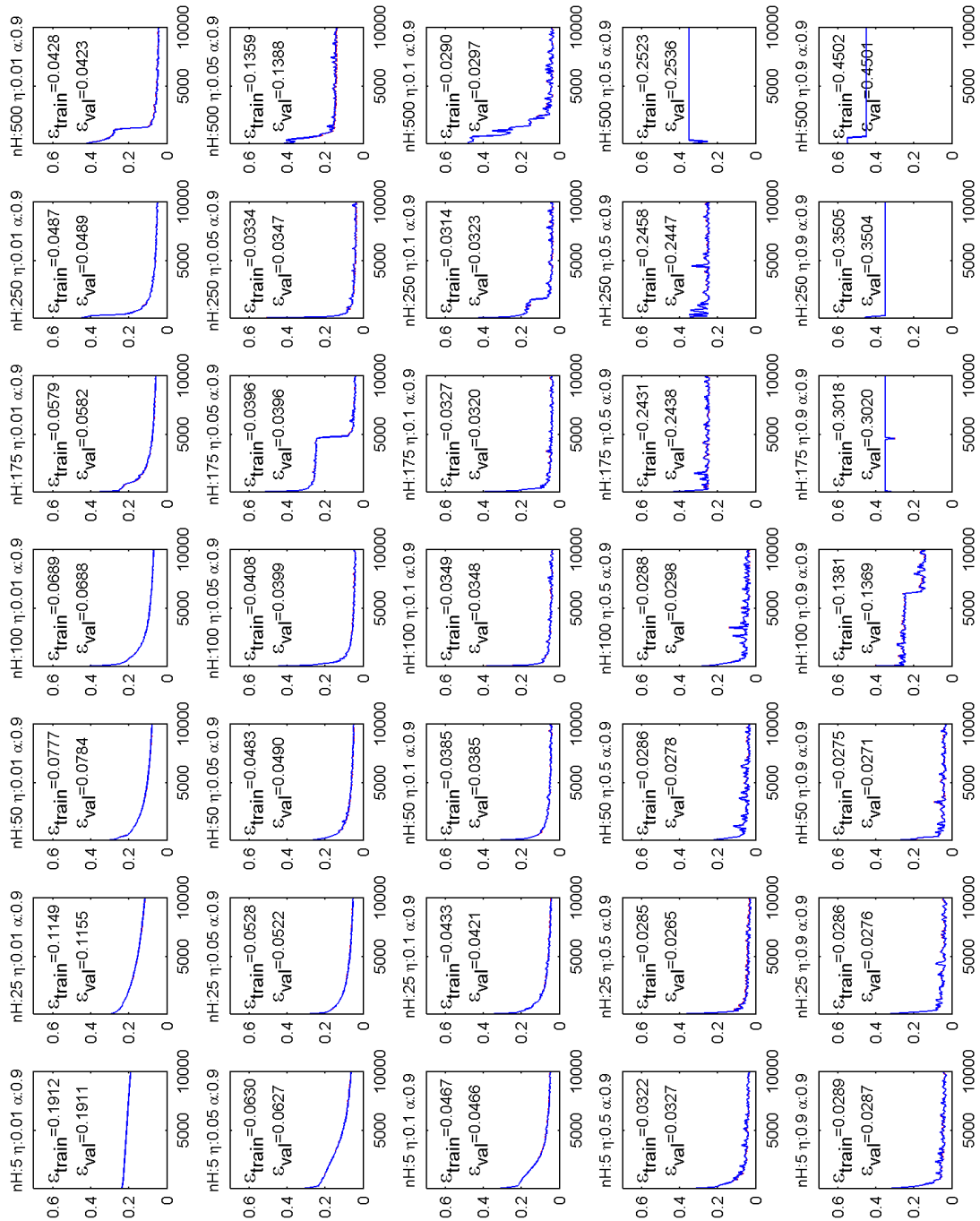


Figure A.12: Training and validation errors during 10000 epochs of *Prosopis juliflora* for number of hidden units: [5, 25, 50, 100, 175, 250, 500], learning rate:[0.01, 0.05, 0.1, 0.5, 0.9] and momentum: 0.9.

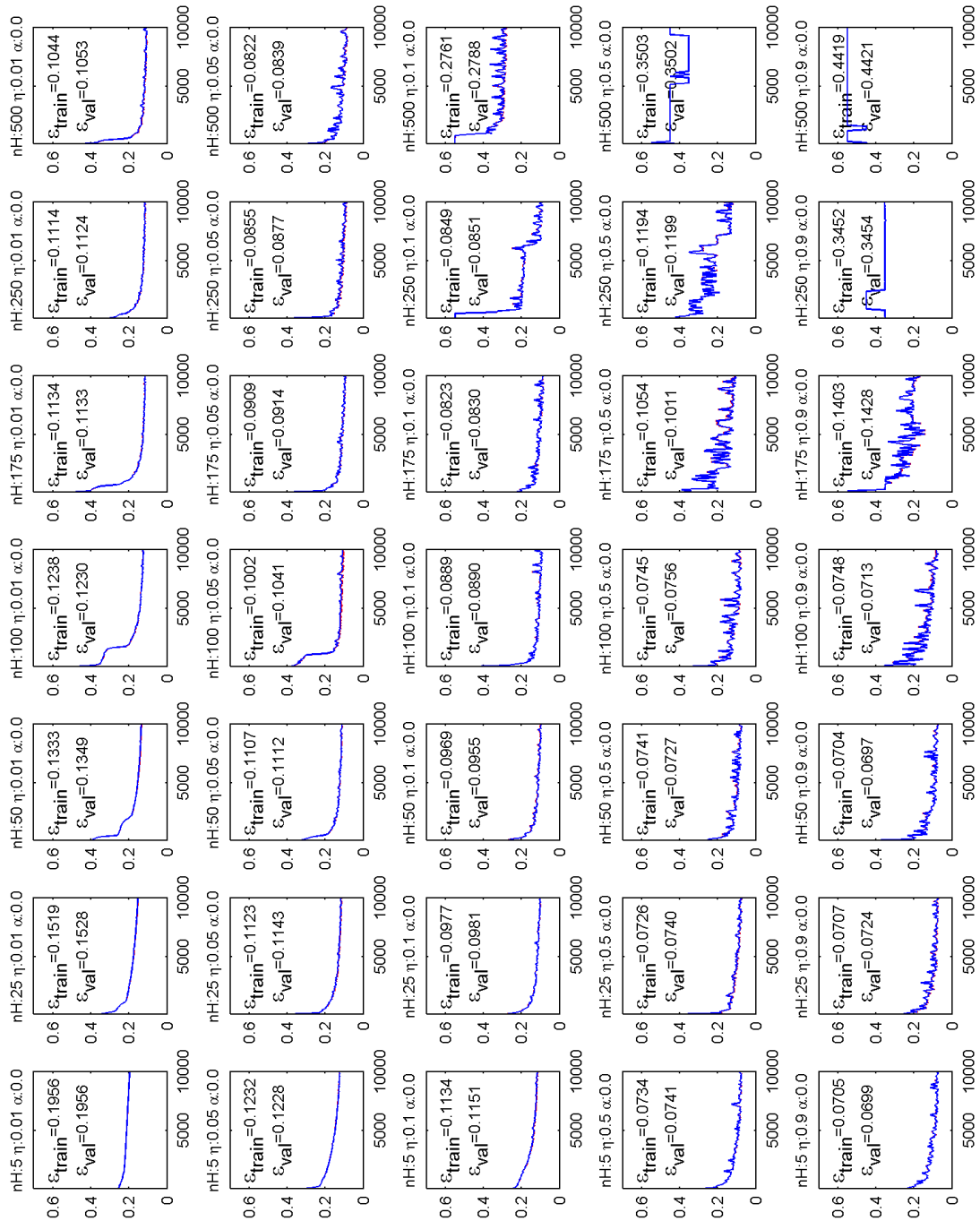


Figure A.13: Training and validation errors during 10000 epochs of *Ptelea trifoliate* for number of hidden units: [5, 25, 50, 100, 175, 250, 500], learning rate:[0.01, 0.05, 0.1, 0.5, 0.9] and momentum: 0.

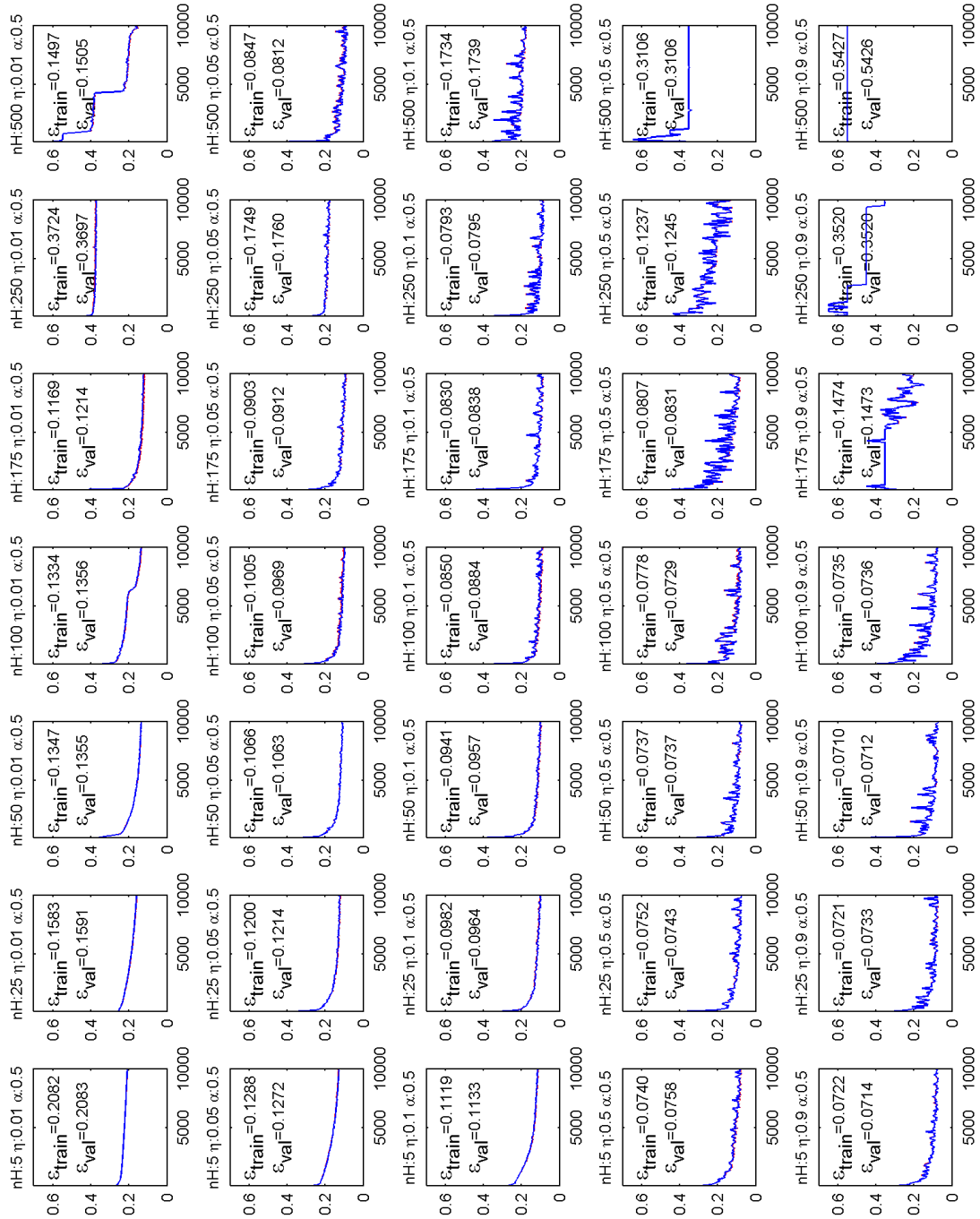


Figure A.14: Training and validation errors during 10000 epochs of *Ptelea trifoliate* for number of hidden units: [5, 25, 50, 100, 175, 250, 500], learning rate:[0.01, 0.05, 0.1, 0.5, 0.9] and momentum: 0.5.

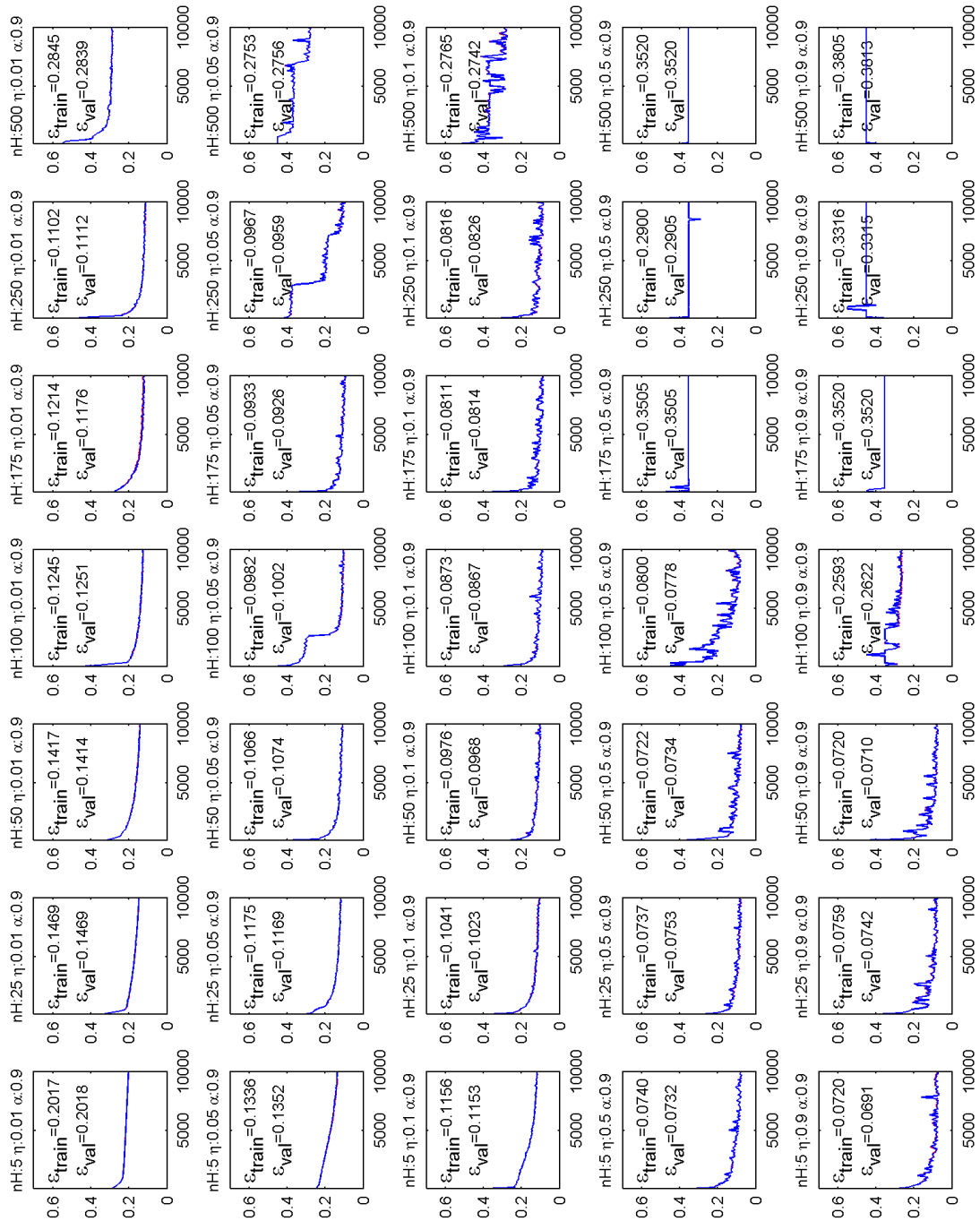


Figure A.15: Training and validation errors during 10000 epochs of *Ptelea trifoliate* for number of hidden units: [5, 25, 50, 100, 175, 250, 500], learning rate:[0.01, 0.05, 0.1, 0.5, 0.9] and momentum: 0.9.

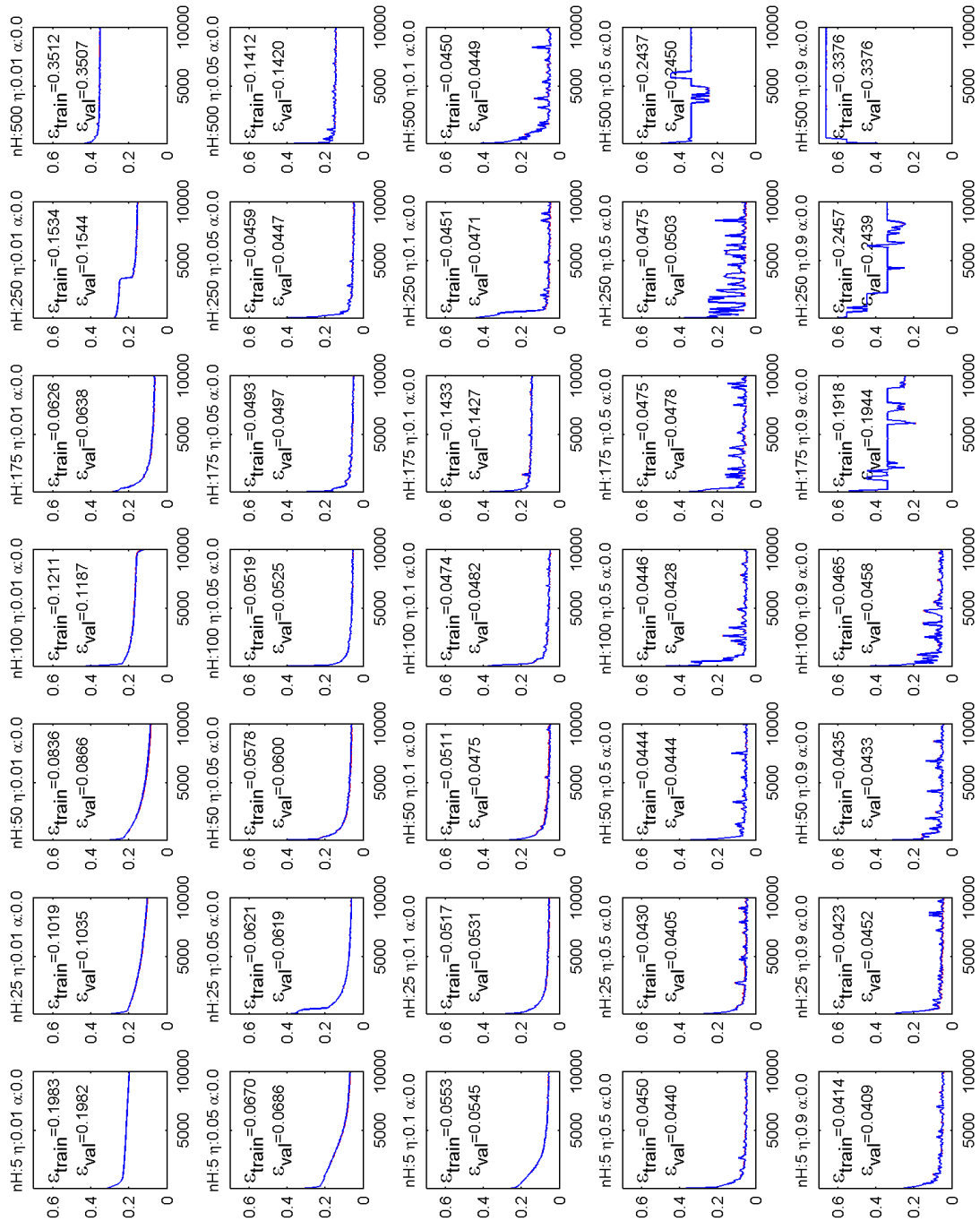


Figure A.16: Training and validation errors during 10000 epochs of *Zanthoxylum fagara* for number of hidden units: [5, 25, 50, 100, 175, 250, 500], learning rate:[0.01, 0.05, 0.1, 0.5, 0.9] and momentum: 0.

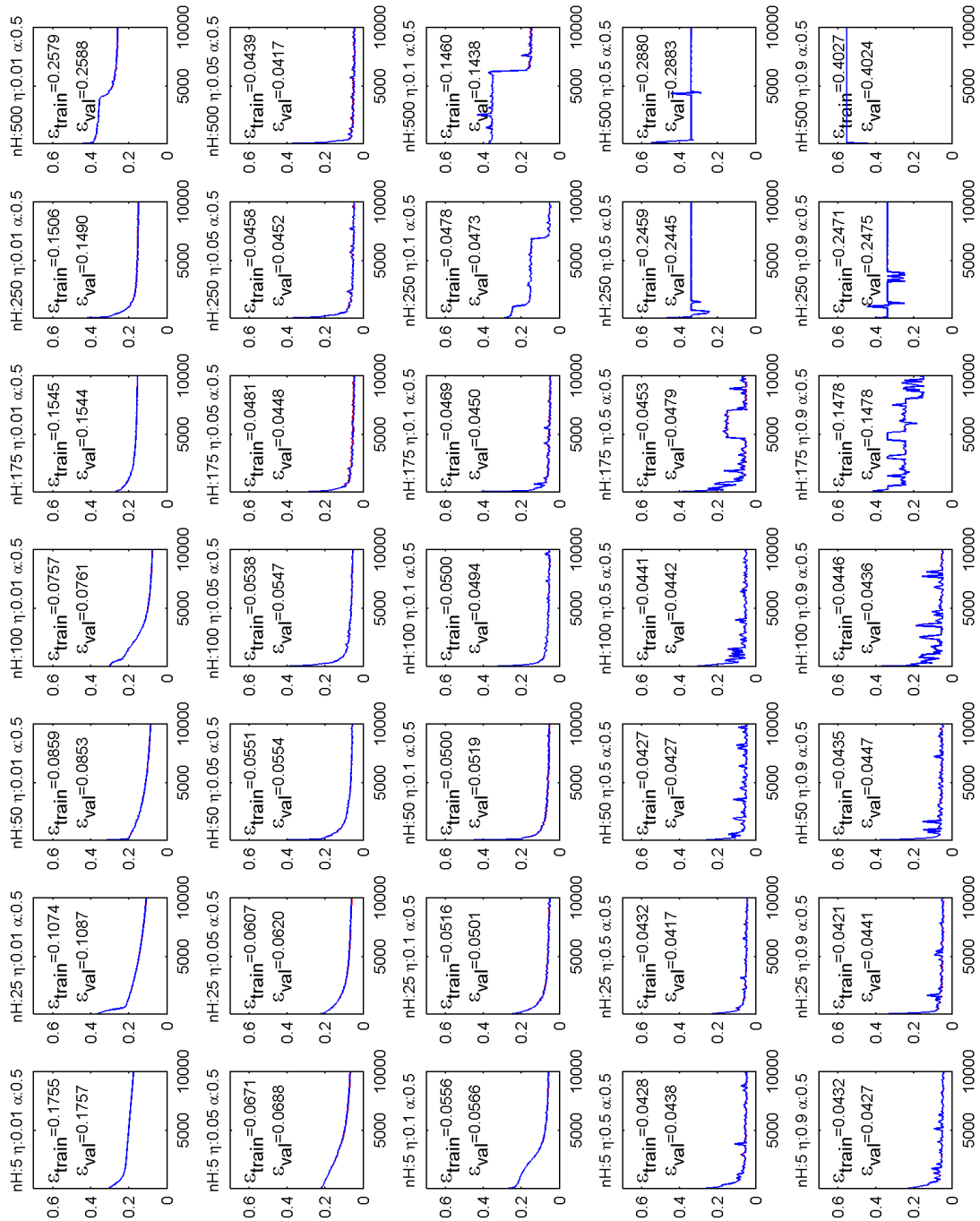


Figure A.17: Training and validation errors during 10000 epochs of *Zanthoxylum fagara* for number of hidden units: [5, 25, 50, 100, 175, 250, 500], learning rate:[0.01, 0.05, 0.1, 0.5, 0.9] and momentum: 0.5.

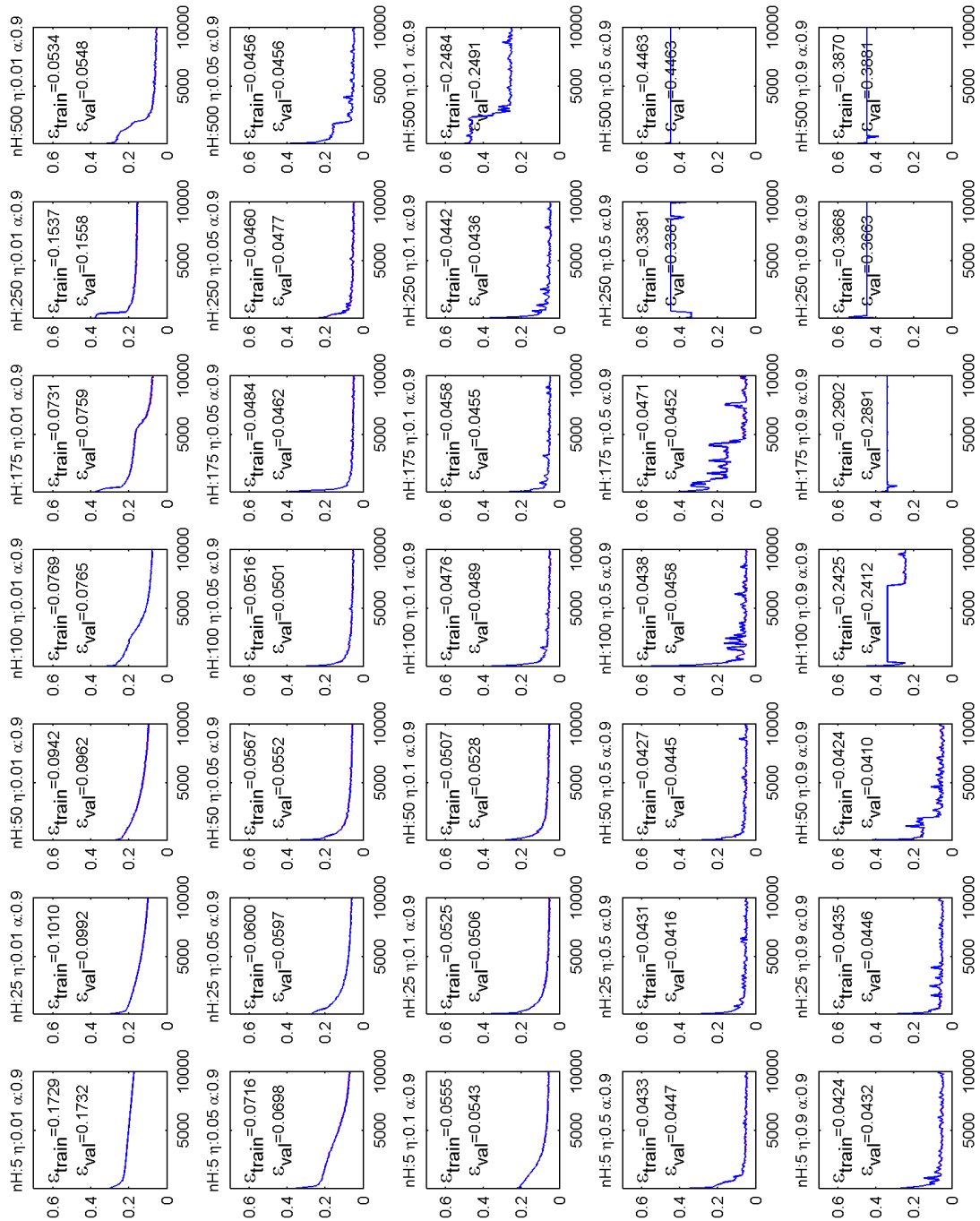


Figure A.18: Training and validation errors during 10000 epochs of *Zanthoxylum fagara* for number of hidden units: [5, 25, 50, 100, 175, 250, 500], learning rate:[0.01, 0.05, 0.1, 0.5, 0.9] and momentum: 0.9.

APPENDIX B

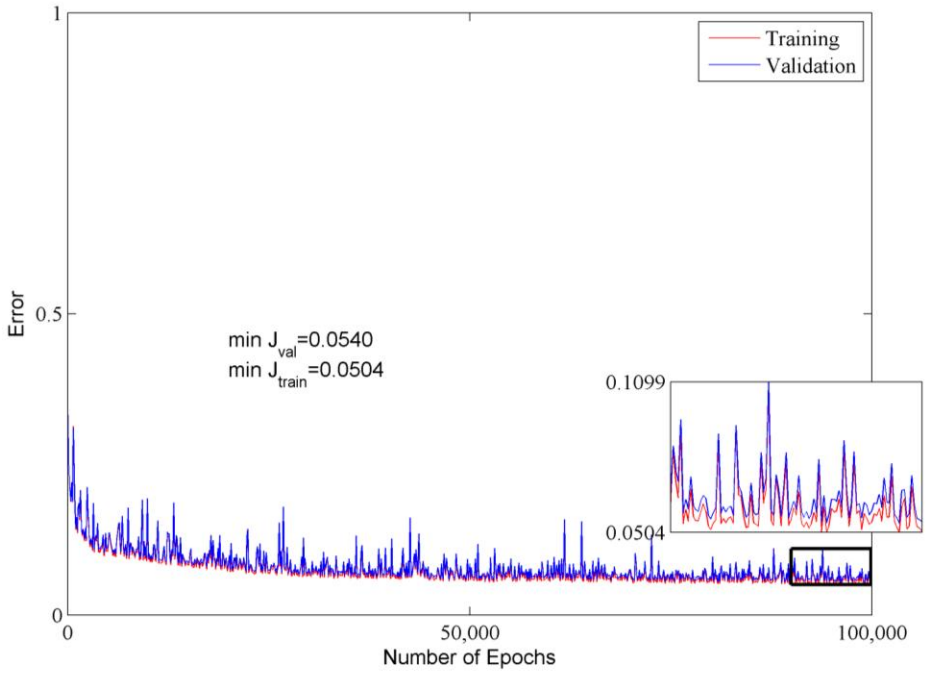


Figure B.1: Training and validation errors during 100,000 epochs for *Amelanchier alnifolia*

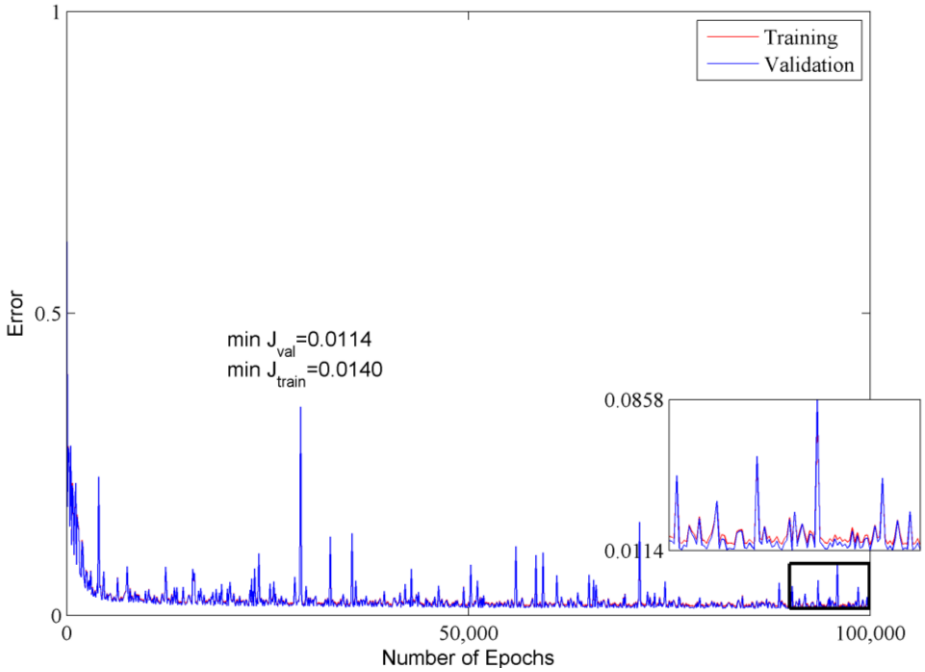


Figure B.2: Training and validation errors during 100,000 epochs for *Asimina triloba*

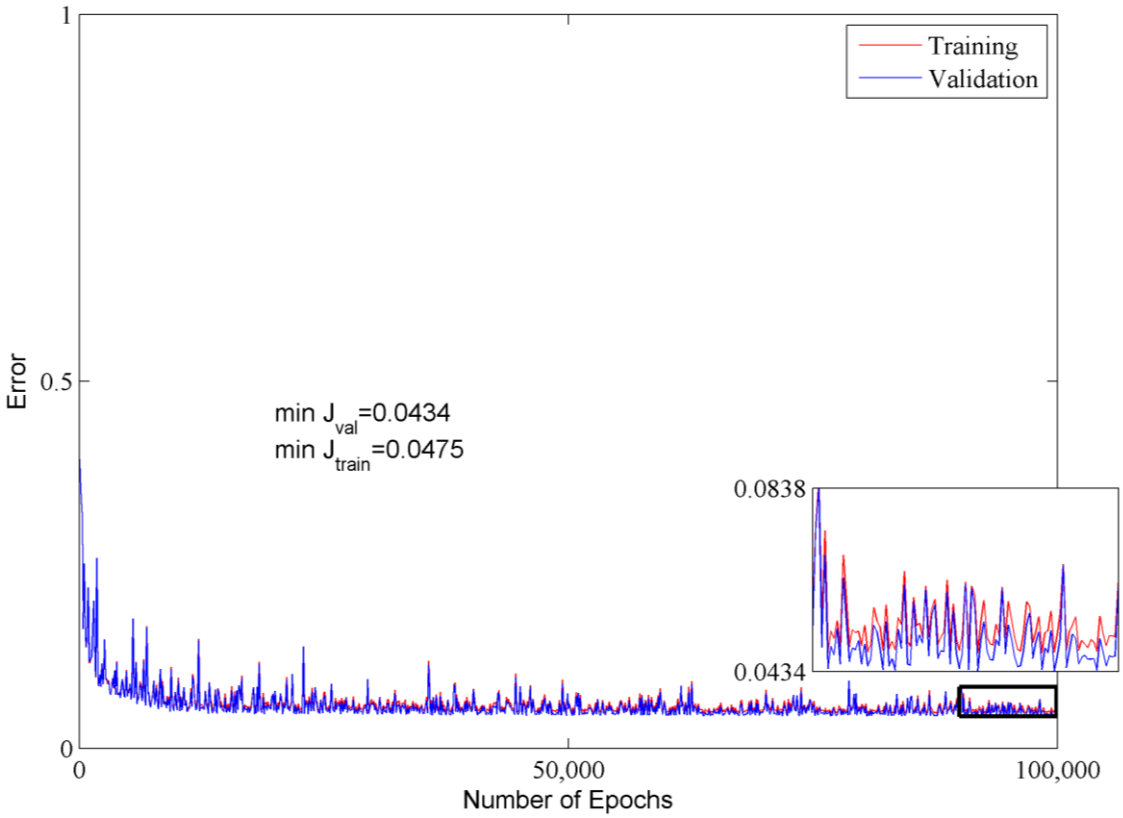


Figure B.3: Training and validation errors during 100,000 epochs for *Betula papyrifera*.

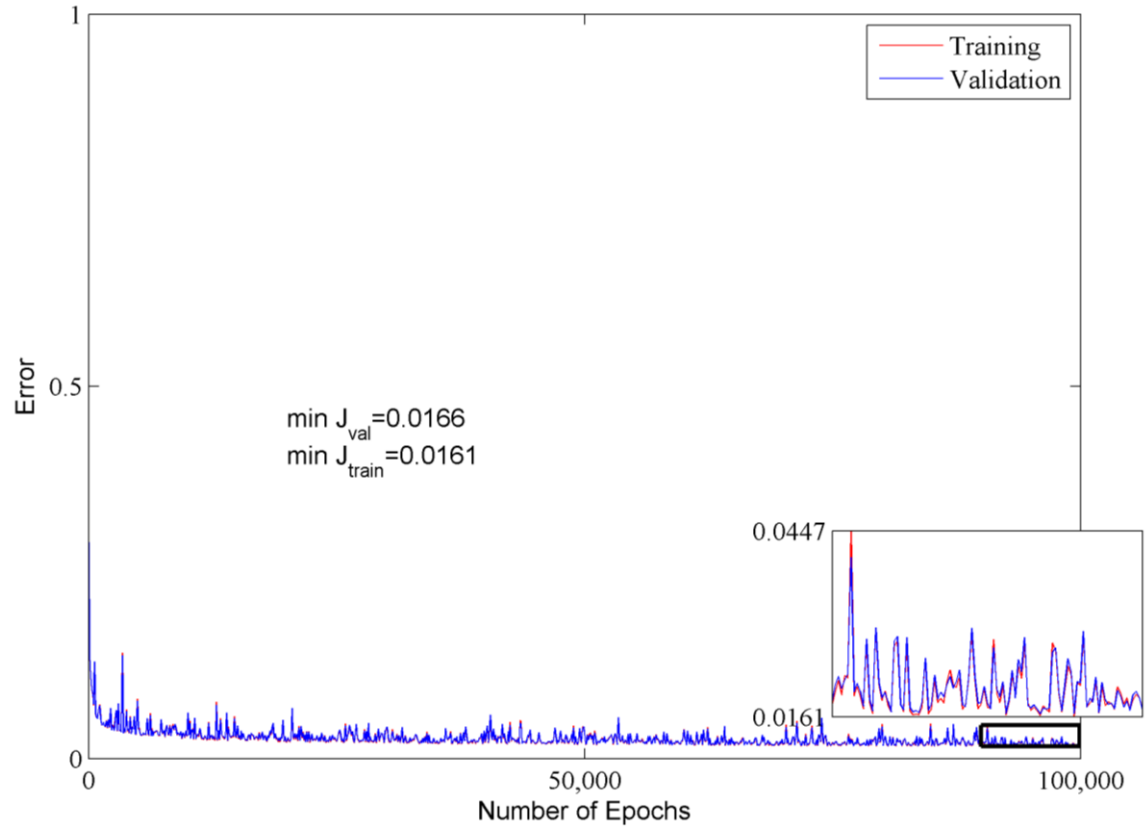


Figure B.4: Training and validation errors during 100,000 epochs for *Prosopis juliflora*.

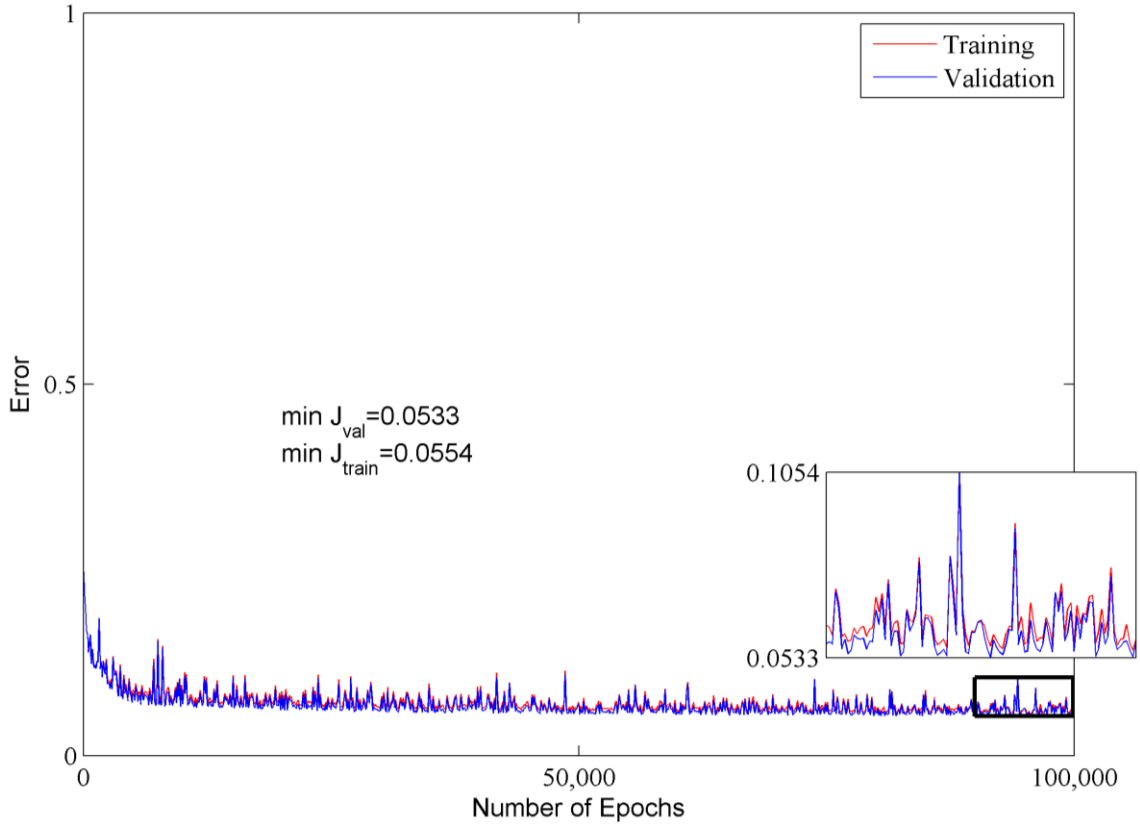


Figure B.5: Training and validation errors during 100,000 epochs for *Ptelea trifoliate*.

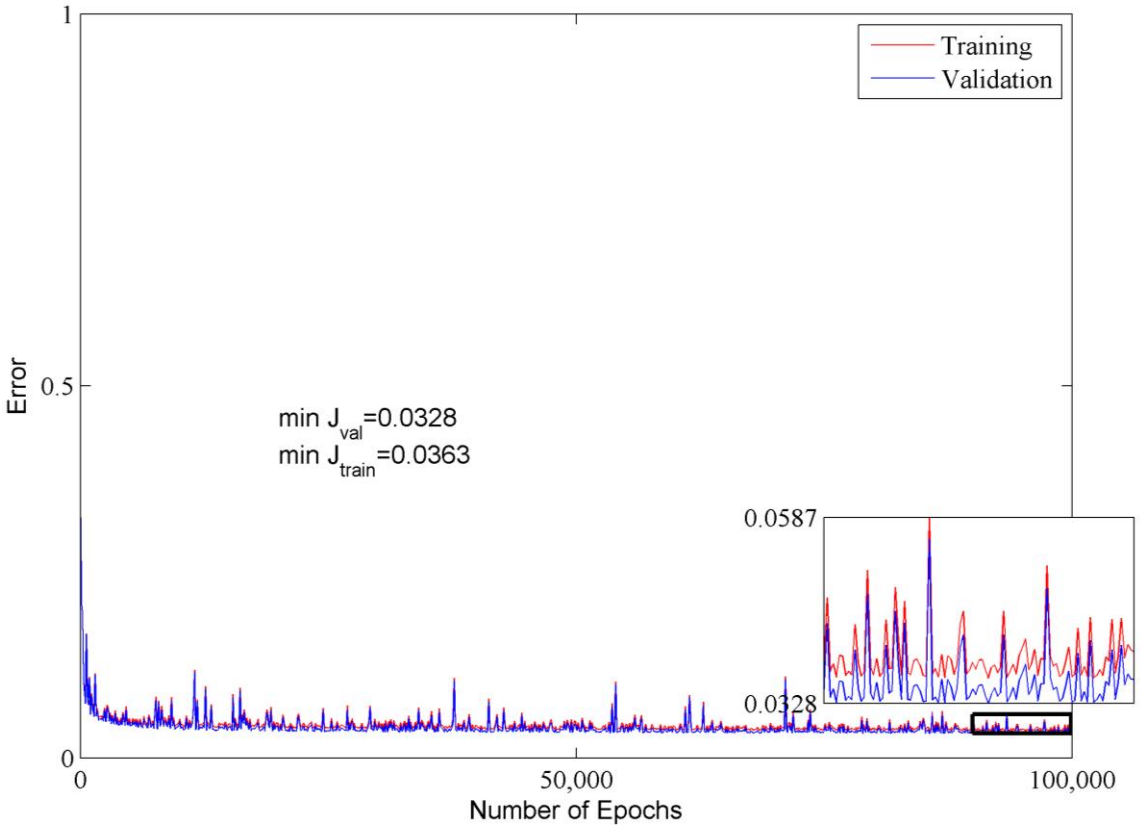


Figure B.6: Training and validation errors during 100,000 epochs for *Zanthoxylum fagara*.

APPENDIX C

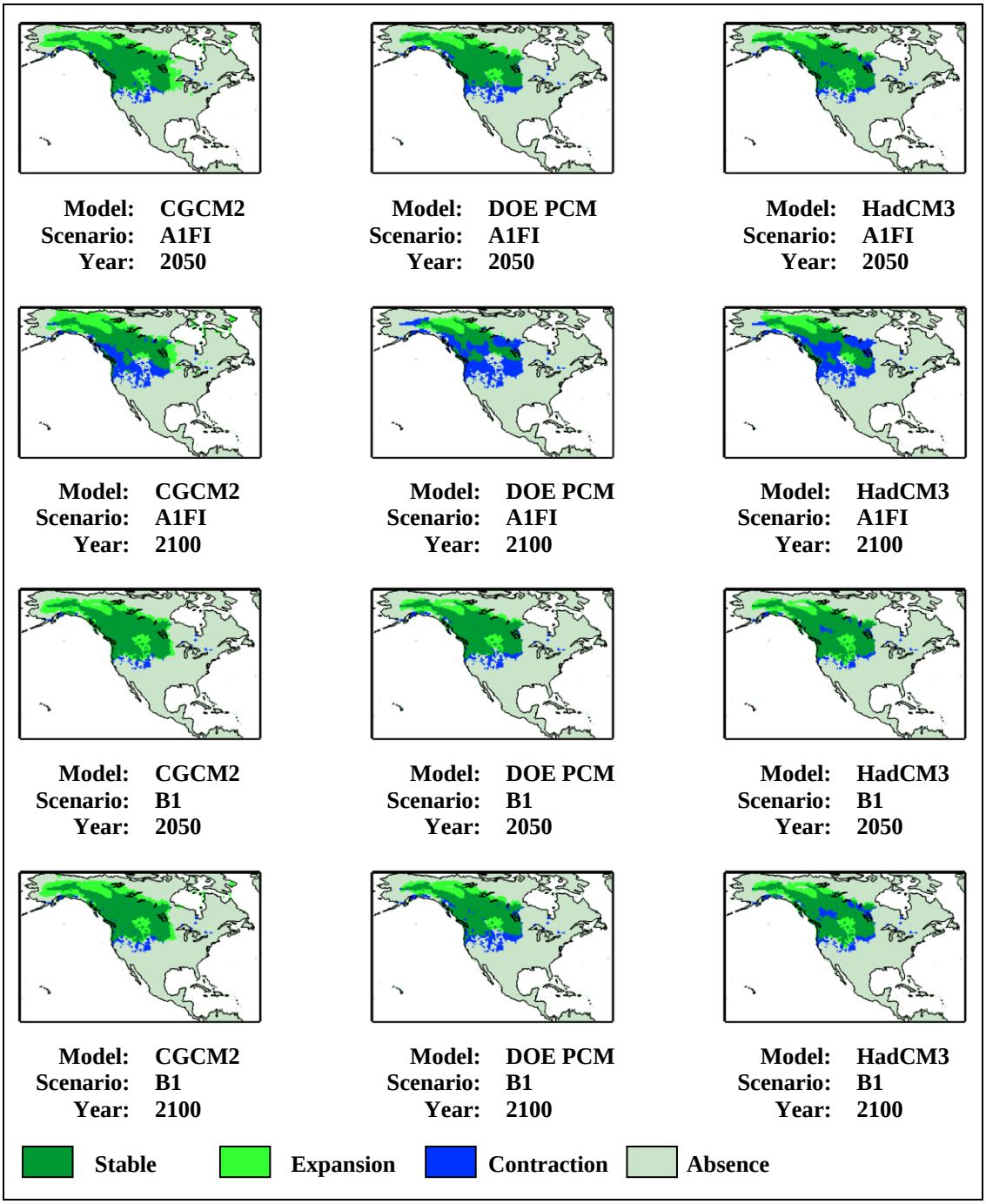


Figure C.1: Future distribution projections of *Amelanchier alnifolia* with different climate models and emission scenarios.

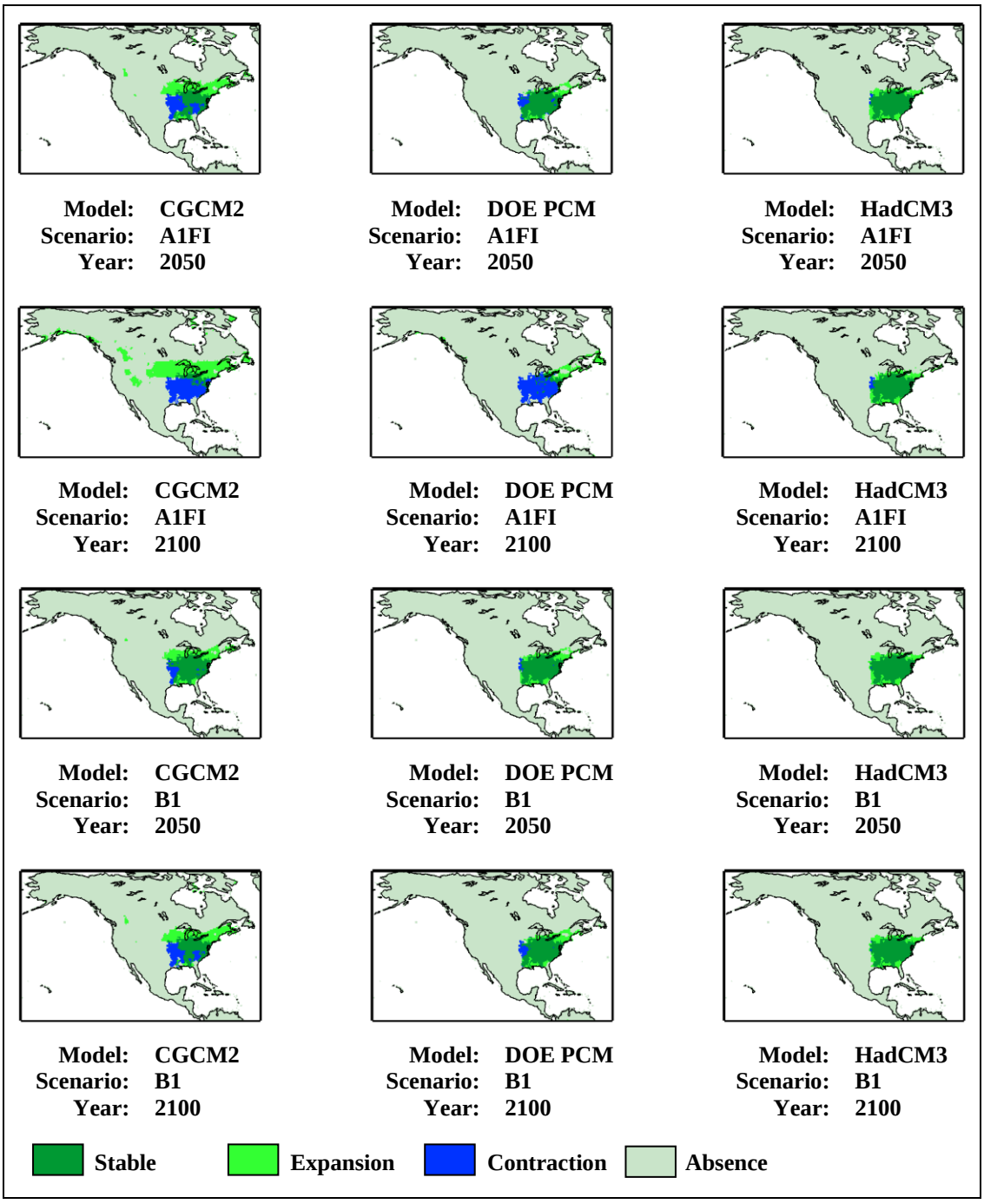


Figure C.2: Future distribution projections of *Asimina triloba* with different climate models and emission scenarios.

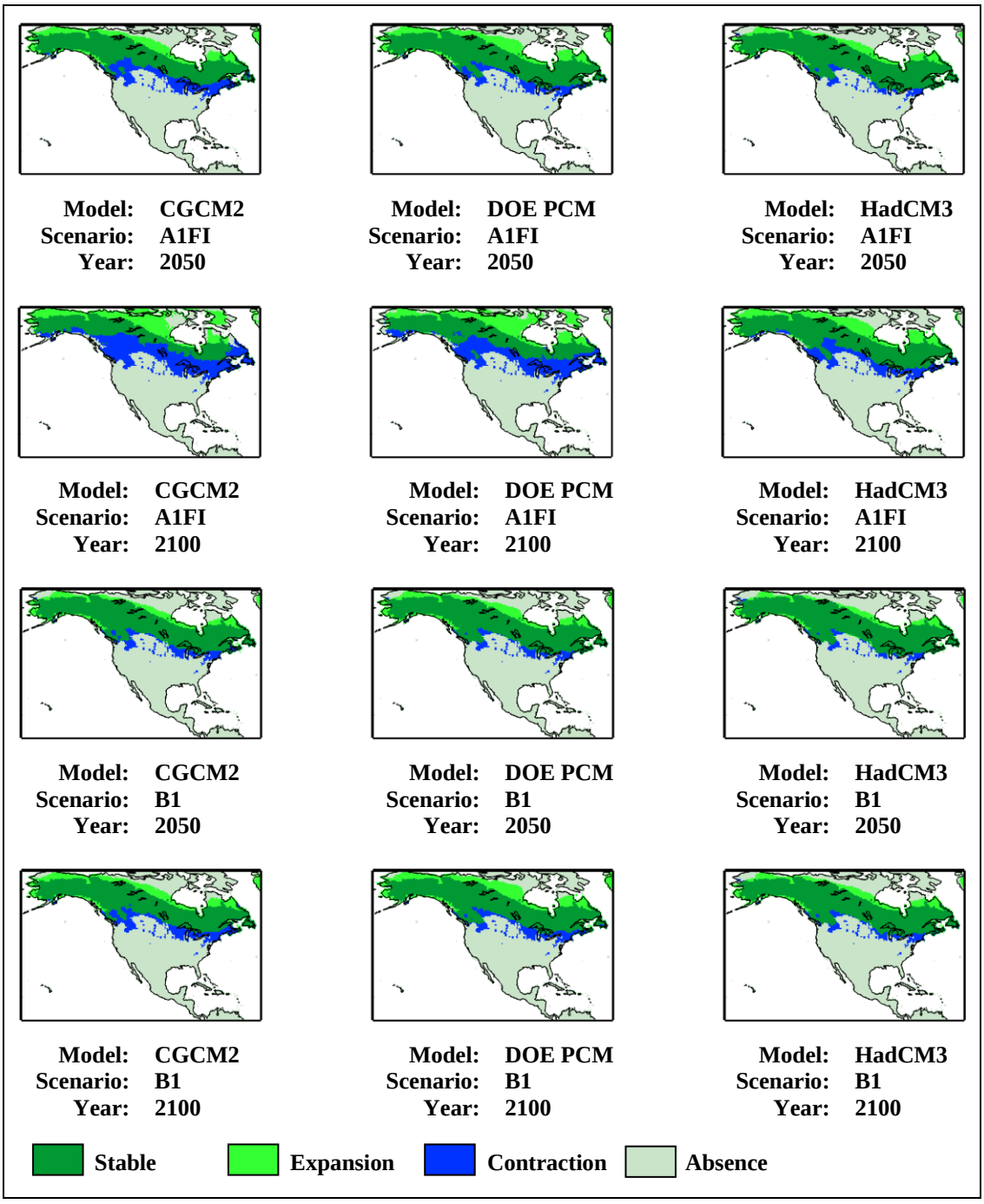


Figure C.3: Future distribution projections of *Betula papyrifera* with different climate models and emission scenarios.

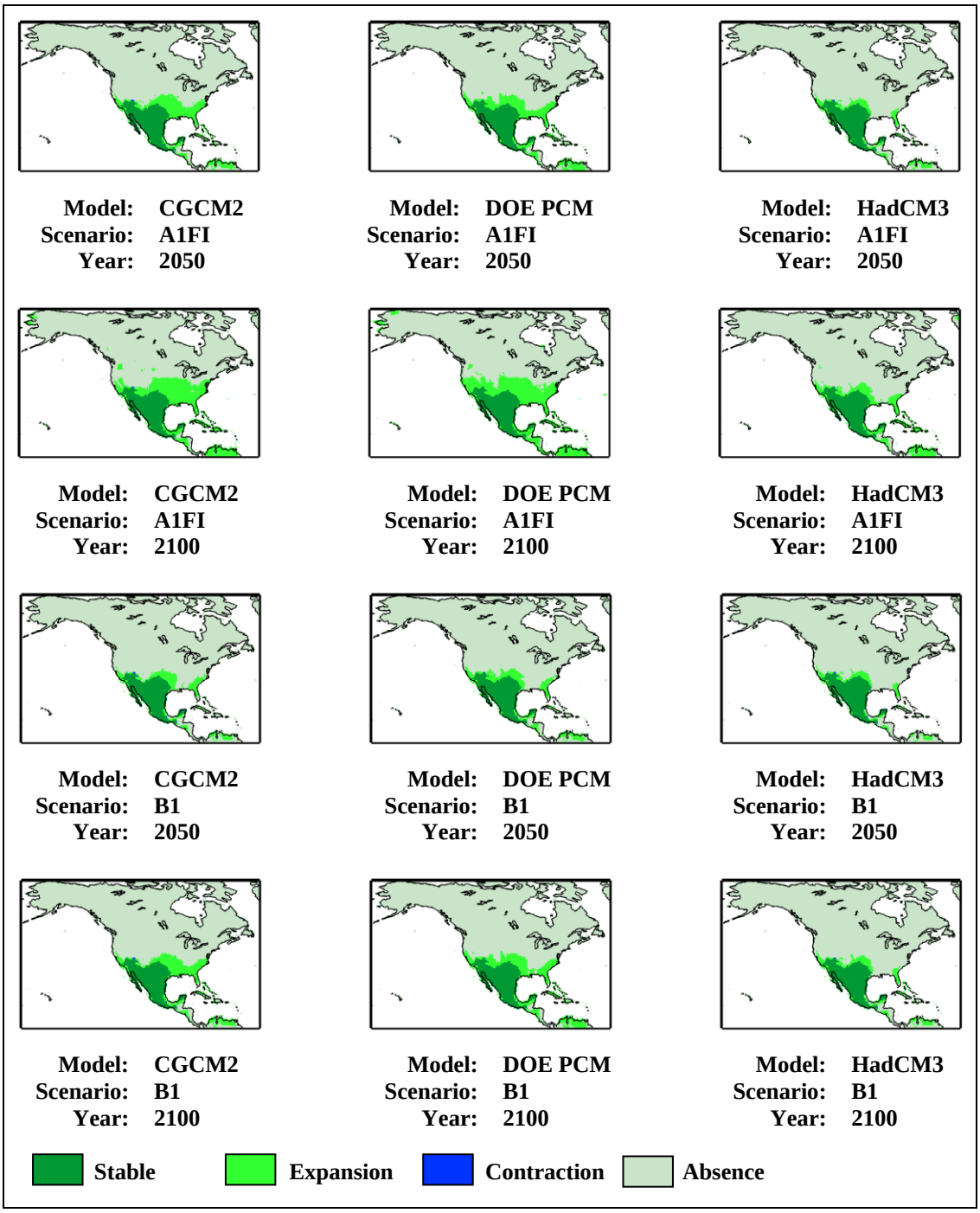


Figure C.4: Future distribution projections of *Prosopis juliflora* with different climate models and emission scenarios.

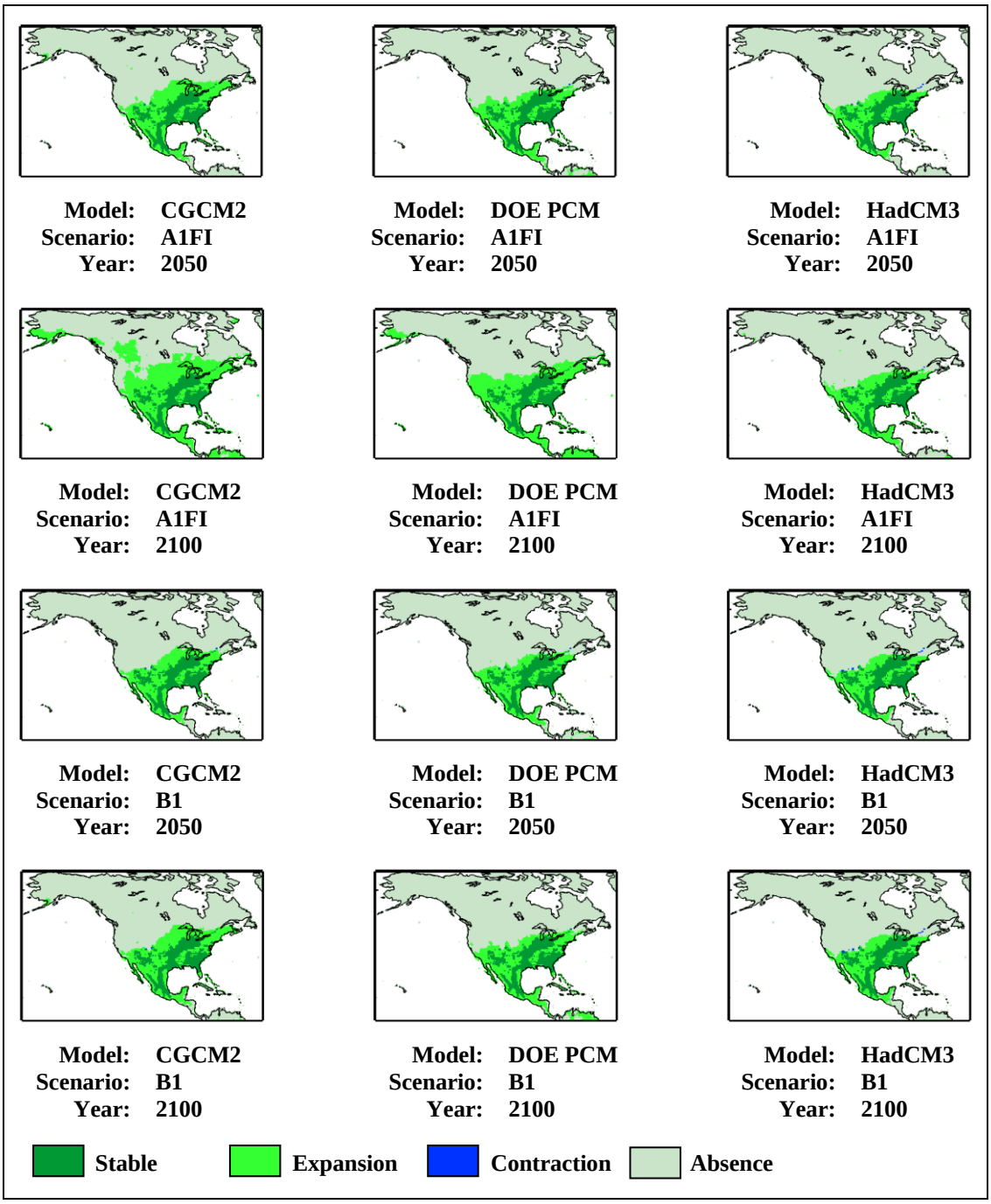


Figure C.5: Future distribution projections of *Ptelea trifoliate* with different climate models and emission scenarios.

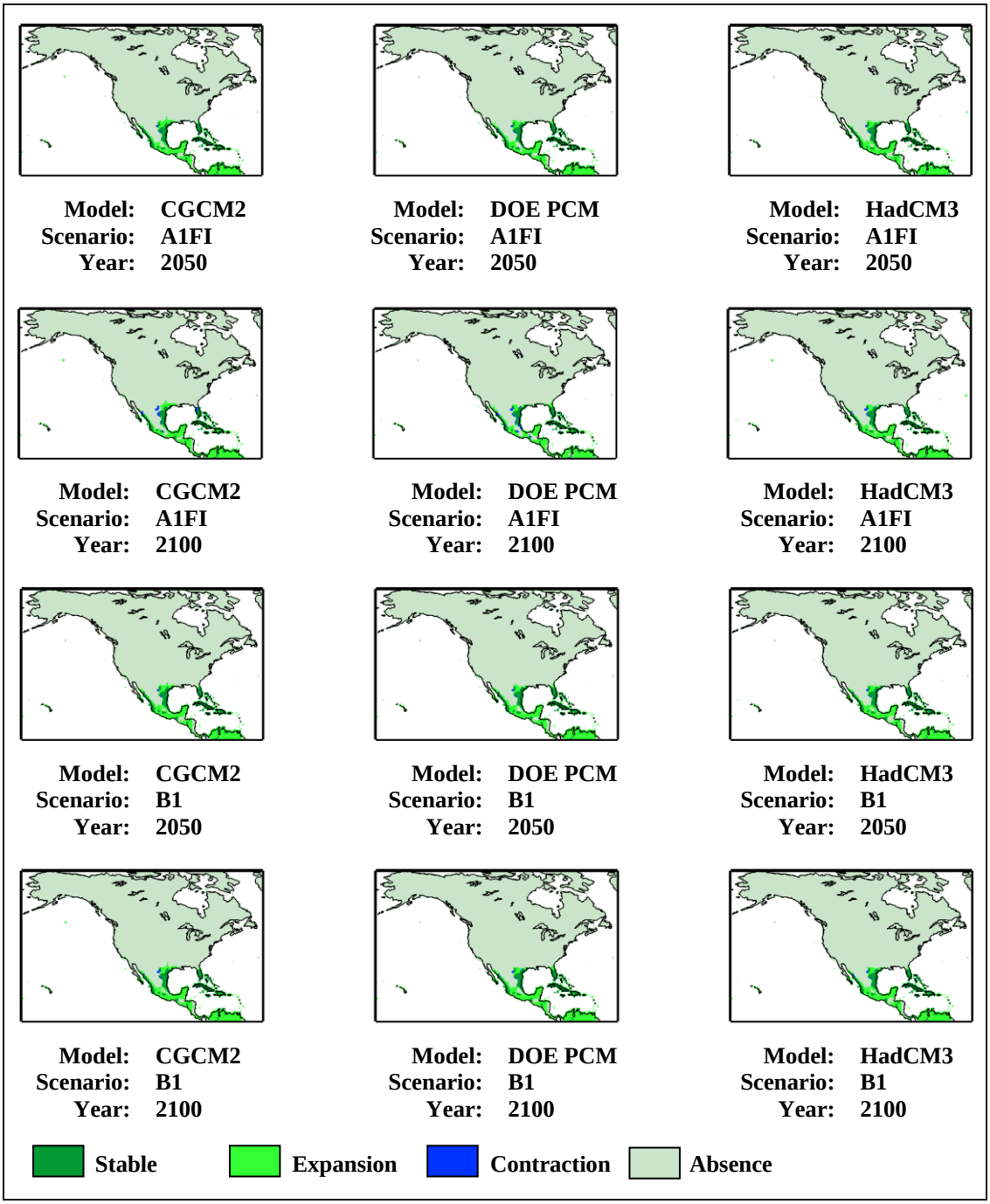


Figure C.6: Future distribution projections of *Zanthoxylum fagara* with different climate models and emission scenarios.

BIOGRAPHY

He was born in 1981 in Istanbul. He graduated from Haydarpaşa High School in 1998 and received his B.Sc. degree from Mathematics Engineering Department in Istanbul Technical University in 2003. He has been a M.Sc. candidate in the Computational Science and Engineering Department of Informatics Institute in Istanbul Technical University since October, 2004.