

©Copyright 2012

Jighan Jeong

Modeling habitat suitability of *Pinus densiflora* in response to climate change in the Korean peninsula and East Asia

Jighan Jeong

A thesis

Submitted in partial fulfillment of the
Requirements for the degree of

Master of Science

University of Washington

2012

Committee:

Soo-Hyung Kim

Daniel Vogt

Joshua Lawler

Program Authorized to Offer Degree:

Forest Resources

University of Washington

Abstract

Modeling habitat suitability of *Pinus densiflora* in response to climate change in the Korean peninsula and East Asia

Jighan Jeong

Chair of the Supervisory Committee:

Assistant Professor Soo-Hyung Kim and Associate Professor Daniel Vogt

College of the Environment

The current and future climate suitability of *Pinus densiflora* in the Korean peninsula, Japan Isles, and parts of northern China was modeled using Random Forest predictors. *Pinus densiflora* is among the tree species with highest conservation priorities for its historical and cultural importance but distribution of the species is rapidly declining in Korean peninsula. A bioclimate envelope model was built on current spatial presence/absence data of the species from digital forest map of South Korea from Korean Forest Service and high definition digital climate data from 2010. Overall the pattern of predicted current habitat in the region was well matched with existing atlases and literature on *Pinus densiflora*. Area under the Receiver-Operating Characteristic Curve (AUC) value of the prediction was 0.86 and Cohen's Kappa value was 0.51. The future suitable habitats of the species were predicted by the model at two different resolution, 270 m x 270 m and 1 km x 1 km. The finer resolution maps predicted futures in two IPCC climate change scenarios A1b and A2a. The coarser resolution maps predicted future habitat of

the species in a broader extent which covers Northeast Asia with three IPCC climate change scenarios A1b, A2a, and B2a. Overall, thirteen projected climate datasets from six different general circulation models under the three different emission scenarios were used for predicting future habitat suitability in response to climate change in the region. At both resolutions, the model predicted that *Pinus densiflora* will expand its range northward, although the area of distribution will decrease 20 ~ 50 %. This tendency is reinforced with more dramatic predictions of climate change. The predicted climate niches of *Pinus densiflora* suggest that precipitation can affect to the species distribution as well as temperature. This result contrasts past predictions that the distribution of the species is governed only by temperature. A preliminary analysis indicates that biotic factors such as pine wilt disease can also influence the distribution of this species. This result suggests that adaptive pest management strategies are likely to be critical for preserving suitable habitats for *Pinus densiflora* in Korean peninsula in a changing climate.

TABLE OF CONTENTS

| | Page |
|--|------|
| List of Figures | iii |
| List of Tables | iv |
| Introduction | 1 |
| Current Korean Forests | 1 |
| Terrain and History | 2 |
| Human Impact | 4 |
| Climate | 5 |
| Needs for Predicting Future Changes of Forests | 5 |
| Study Objectives | 7 |
| Materials and Methods | 9 |
| Data | 9 |
| Random Forest Model | 11 |
| Environmental Variables | 17 |
| AUC and Kappa Statistics | 17 |
| Result | 19 |
| Fine Resolution Maps | 19 |
| Coarse resolution maps | 29 |
| Discussion | 32 |
| The Model Performance | 32 |
| Variables | 35 |
| The Future of <i>Pinus densiflora</i> | 42 |
| Suggestions for a Future Conservation Policy for Forests | 49 |
| Conclusions | 51 |
| Acknowledgments | 52 |

| | |
|---|----|
| Reference | 53 |
| Appendix A: <i>Pinus koraiensis</i> | 57 |
| Appendix B: <i>Quercus spp.</i> | 59 |

LIST OF FIGURES

| Figure Number | Page |
|--|------|
| 1. Elevational Terrain of the Korean Peninsula | 3 |
| 2. Approximate Known Distribution Range of <i>Pinus densiflora</i> | 14 |
| 3. Current Distribution of <i>Pinus densiflora</i> in South Korea | 15 |
| 4. Relative Importance Table of Environmental Factors in the <i>Pinus densiflora</i> Model | 22 |
| 5. Change of <i>Pinus densiflora</i> Suitable Habitat Areas in the South Korean Region at Different Threshold Values | 25 |
| 6. Model Prediction Map for Current (2010) Suitable Habitat of <i>Pinus densiflora</i> at 270 m x 270 m Resolution | 26 |
| 7. <i>Pinus densiflora</i> A1b Scenario Maps at 270 m x 270 m Resolution | 27 |
| 8. <i>Pinus densiflora</i> A2a Scenario Maps at 270 m x 270 m Resolution | 28 |
| 9. Estimated Future Habitat of <i>Pinus densiflora</i> at 1km x 1 km Resolution | 31 |
| 10. 1980 Prediction Map of Suitable Habitat for <i>Pinus densiflora</i> and Species Distribution Atlas from 1960s | 37 |
| 11. Sample Regression Trees for Variables Used in the <i>PIDE</i> Random Forest Model | 41 |
| 12. Comparing Random Forest Model Prediction and Previous Model Predictions for <i>Pinus densiflora</i> | 47 |
| 13. Pattern of Temperature, Precipitation, and Distribution Change of <i>Pinus densiflora</i> | 48 |
| 14. Pine Wilt Disease Susceptible Area | 50 |

LIST OF TABLES

| Table Number | Page |
|---|------|
| 1. List of GCM by Scenarios for 1km x 1 km Prediction Data | 13 |
| 2. Acronyms and Definitions of Environmental Variables for the Suitable Habitat Model | 16 |
| 3. Land Cover Percentage of Predicted Presence Cells in the South Korean Region | 20 |
| 4. Split Points of Variables that Affect <i>PIDE</i> Habitat Distribution are Presented in the Sample Regression Tree | 42 |

INTRODUCTION

The forests of the Korean peninsula are entering into another age that is created by human influence and climate patterns. Researches has predicted that the conifer-dominated temperate forests will be changed to subtropical mixed forests in near future (Korea Forest Service, 2011). However, only a few papers studied in detail how the distribution of tree species in the region will be changed. This paper models the fundamental spatial niche of the major tree species of this region by current climate and climate change scenarios. *Pinus* are the most diverse and abundant genera of the *Pinaceae* conifer family on the Korean peninsula and four species among them are the major species in the region: Japanese red pine (*Pinus densiflora*), Japanese black pine (*Pinus thunbergii*), Korean pine (*Pinus koraiensis*), and Siberian dwarf pine (*Pinus pumila*) (Kong, 1998; Kong, 2006). In this study, current and future suitable habitat distribution of *Pinus densiflora* (*PIDE*) was predicted by using an environmental envelope since this species is the most abundant and culturally important tree species within the region.

Current Korean Forests

The current forest on the Korean peninsula is a mixed result of terrain, climate, and human activities. Approximately 60 ~ 70 % of the Korean peninsula is forest. Approximately 63.8 % of South Korea is forested land, while 76.5 % of North Korea is estimated as forested land. The forests on the Korean peninsula consist of 41.9 % coniferous forests, 26 % deciduous forests, 29 % mixed forests, and 3.1 % of bamboo or un-stocked forest lands, which have *Pinus* (33 %), *Larix* (7.2 %), and *Quercus* as the main species (Korea Forest Service, 2011). As mentioned previously, it is estimated that 76.5 % of North Korea land is forested (Boo, 2001), however, this amount can't be accurately verified. According to a report presented by the United Nations

Environment Program (UNEP, 2003), at the end of 1996, North Korea had 8,960,000 ha of forested land. However, the Food and Agricultural Organization (FAO) recently estimated that North Korea only had 5,666,000 ha of forested land, which is 46% of the entire land area. Expansion of agricultural lands, collection of excessive fuel wood from the forests, and over-logging were suggested as the reason for this rapid reduction (Korea Forest Service, 2011).

Terrain and History

Korea is a peninsula of hills and mountains. As a result, it contains various micro-climates and slopes even within a small region. Mountains cover the northeast region of the peninsula, while the southeast region has a relatively broad plain area (Fig. 1). This topographical feature of the land made the western part of the peninsula more habitable for humans, so it is densely populated. As a result the western part of the peninsula has felt the deep impact of human modifications. On the other hand, the eastern part remains a more natural area that is less affected by human activities. According to pollen researches, cryophilous evergreen coniferous species, such as *Picea*, *Abies*, *Pinus*, were dominant in this region during the last glacial period. Those species declined from 15,000 to 10,000 years B.P and deciduous broad-leaved trees, such as *Salix*, *Juglans*, *Carpinus*, *Fraxinus*, *Alnus* and etc., increased by the end of glacial period and continues to the present (Choi. Kee, 1998; Kong, 2000b). However, from B.P 6000, conifers which are mainly *Pinus* started to re-dominate Korean forest and became the main species again from B.P 2000. Since the current dominant species *PIDE* is and early successional species which prefers bare ground and strong sunlight for germination and growth, human activities, such as timber harvesting and planting, are generally accepted as the main reason for this reemergence. It has been suggested that the alternation of the sea level and extended dry seasons the Korean

peninsula has experienced are also part of the reason for this change in forest species. However, more evidences is needed to support this theory (Choi. Kee, 1998).

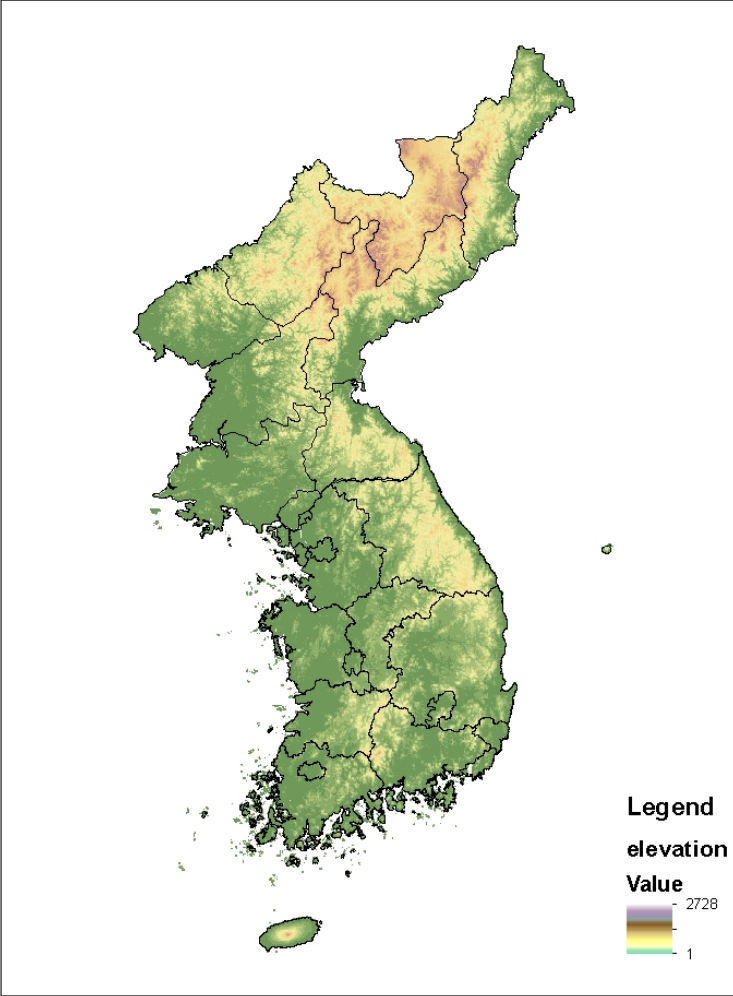


Figure 1. Elevational Terrain of the Korean Peninsula.

Human Impact

Human interventions in the forested lands of the Korean peninsula have a long and documented history. The forests of the Korean peninsula have been impacted greatly by human activities from as early as the Bronze Age. Agriculture started on the Korean peninsula approximately 4000 years ago, and the human population's fuel requirements increased quickly after that. From the three-kingdom period, approximately 2000 years ago, the use of iron tools plus the raising of cattle required the expansion of agricultural fields and further necessitated the reduction of forested lands. Continuous timber felling using advanced tools and the removal of forest litter for fuel and fertilizer made forest soils less fertile and interrupted forest species succession. Various activities, such as government encouraged cultivation, building palaces and ships, mining, woodblock printing, and war, resulted in broad areas of deforestation (Kong, 2000a). From the Wang dynasty (AD 981 ~ 1392) to the Yi dynasty (AD1392 ~ 1910), protecting forests and reforestation which are centered on *Pinus densiflora* were started by the government due to severe deforestation. There are many historical felling prohibition and planting records regarding *PIDE* (Yoon, 2003). During the Japanese colonial era (AD1910 ~ 1945) and the Korean War period (AD1950 ~ 1953) following the Yi dynasty, the Korean peninsula suffered severe deforestation by both the direct and indirect impacts of war. After the war, the average volume of forest was less than 5.0 m³/ha, and approximately 30% of the forested land was either denuded or had a low density of wood stock (Hyun 1956). The South Korean government has adapted strong and successful reforestation policies since 1962. As a consequence, approximately 46 % of the current total areas of forested lands in South Korea are plantations. From the period of 1953 to 1990, approximately 632,264 ha of land have been replanted with exotic conifers and 942,284 ha of land have been planted with exotic and hybrid deciduous species (Kim and Zsuffa,

1994). Such historical records show the deep impact of human activities to the forests on the Korean peninsula. However, intense human disturbance is being dramatically decreased due to the decreased demand for woody fuel in the developed parts of modern society. In the last few decades, the forest policies of South Korea is changed to loose managing rather than reforestation in current couple of decades. It means much less direct human impact will press Korean forest from present. Now, Korean Forest Service of South Korea is managing the forest by protecting and thinning. In particular, focused researches and protecting for *PIDE* is being carried on since the species is one of the main cultural symbols in Korean tradition.

Climate

Korean peninsula has temperate climate with various micro weather and climate conditions due to its mountainous terrain and location between Asia continent and North Pacific Ocean. The latitude of the region is between 33 and 43 degree north, and the longitude is between 124 and 132 degree east. The location results temperate climate which has hot, humid summer with a rainy season from late June to early July and a typhoon season from august to September, and cold, dry winter in effect of westerly and monsoon. Average temperature is 10 ~16 °C and annual precipitation is 1200 mm. This precipitation is almost twice as much as the world average annual precipitation, but most of it is concentrated during summer.

Needs for Predicting Future Changes of Forests

Now both of human activities and climate are changing. Predicting forest response in climate change situation is becoming important more and more since direct human impact on Korean forest is decreasing while current global warming situation is expected to change climate environment remarkably. Wood is not the main source of energy anymore, and it decreased

outflow of materials from current Korean forest significantly. Government forest policy has been changed from reforestation to management and thinning. Still forest fire is on the control, but intensity of human impact decreased dramatically. Climate is also on change with global climate change situation. The average temperature on the earth increased 0.74 °C during last 100 years (IPCC, 2007). The temperature of Korean peninsula showed more stiff change. The annual temperature increased 1.5 °C during the same period, and extreme cold temperature period during winter is decreasing (IPCC, 2007). Annual precipitation is also increasing mainly due to increase of summer precipitation. The annual precipitation of last 10 years was 1,458.7 mm which is approximately 10 % higher than the average, but number of rainy day is decreasing during last 50 years. (Oh, 2008) Overall, the climate of Korean peninsula is moving toward subtropical climate (Kwon *et al.*, 2007). Considering the temperature difference between glacier age and present is only 5 °C, 1.5 °C increase during 100 years is very fast change.

With these changes, predicting future species distribution and composition in the forests of the Korean peninsula is necessary for the people to react properly against undesirable changes of the regional ecosystem. Changing of forest ecosystem is deeply related with core ecosystem functions such as water circulation, soil erosion, nutrient circulation, food web, and so on. Moreover, it is also related with culture and tradition of the people of the region. Therefore, change of the forest can change the regional ecosystem and human life, and thus is in need to be prepared properly. Knowing how the forest will be changed is very important for this point. Chasing estimated change of major tree species by environment envelope modeling can be an appropriate way to satisfy this demand.

However, only a few studies modeled the changes of environmental niche in entire peninsula scale. There were several papers modeled future habitat of various species by statistical

regression with Warmth Index, Minimum Temperature of Coldest month Index (MTCI), annual average temperature, annual precipitation (Choi *et al.*, 2010; Lee *et al.*, 2006; Lee, 2011; Park *et al.*, 2010; Yoon, 2003). These papers modeled changing of species habitat by few core environmental factors separately and analyzed simple relationship between individual environmental factor and species distribution mostly by linear regression. Therefore, more comprehensive relationship analyzing between species distribution and various environmental factors in entire peninsula scale is asked to predict future change of suitable habitat of forest species since multiple factors affect to species distribution synthetically.

Study Objectives

The objective of this paper is to examine current suitable habitat of *PIDE* in Korean peninsula at fine resolution, and to explore change of suitable habitat distribution of the species under different projected future climate scenarios. Projected habitat maps of *PIDE* by different climate scenarios in and around Korean peninsula will be shown. *PIDE* is one of the most important tree species as a cultural and historical symbol in Korea. *PIDE* was one of the most abundant species in the history of Korean peninsula. Since the species has broad utility as material of building and ship, past dynasties in Korea had protected and planted the species. The species also has strong relationship with regional culture and art for its abundance and utility. Confucianism considers the species as a symbol of fidelity, constancy, principle, belief, and strong vitality (Yoon 2003). *PIDE* comes up even in the national anthem of South Korea. Now *PIDE* has more value as an ornamental tree species and the source of matsutake mushroom rather than as a construction material.

The future of *PIDE* is one of the most concerned issues among the topics about Korean forest since the species are declining or showing signs of decline recently. Conservation of natural *PIDE* forest is becoming a national issue for its fast declining. Tree pest such as pine gall midge and pine wilt disease (PWD), heat and drought stress from climate, forest fire, landslide, and increase of ornamental uses are pointed for the reason of decline (Korea Forest Service, 2011). Many studies expected *QUSP* will be the successional descendant species of *PIDE* in Korean forest (; Dolezal *et al.*, 2009; Kim, 1992; Lee and You, 2009; Lee *et al.*, 2006; Song, 1997). Climate change is suspected preferentially for the reason of these changes since it can affect to both of climate stress itself and spreading of pests that inducing disease. At this point, modeling suitable habitat of *PIDE* can help to see the reason of the changes in the region. Declining in climatically suitable habitat would be the result of other reasons such as pest, while declining in unsuitable habitat would be affected by climate change itself mainly. Conservation policies for the species also can be more efficient with future suitable habitat prediction. The result of this study will help management planning by showing expected future suitable habitat of *PIDE*.

MATERIALS AND METHODS

Data

A digital forest map of South Korea from the Korea Forest Service of South Korea was used to determine the distribution of *PIDE* (http://fgis.forest.go.kr/fgis/Thematic/thematic_01.aspx). The map was based on remote sensed data and revised by field survey at the resolution of 30m x 30m. Thus, digital forest maps can show upper canopy species precisely, but lower canopy species do not adequately appear at this resolution. For instance, *Quercus spp.* occupies 27 % among the total number of trees in South Korea, but the distribution area in the digital forest map shows only 0.4 % since most of them are growing in the lower canopies of *PIDE* dominated forests (Korea Forest Service, 2011). (Korea Forest Service, 2011). Another atlas that shows approximated distributions for *PIDE* was also used for verifying model prediction (Fig. 2). Environmental factor data for building the model was from High Definition Digital Climate Maps (HD-DCMs). Fine resolution climate data from 1980, 1990, 2000, and 2010 for the entire Korean peninsula was modeled in the maps. The maps were made by creating a regional climate model that was based on ECHAM5, a general circulation model (GCM) from Max-Planck Institute. The climate model was revised and downscaled by observed local climate data during 1971-2000 and 1981-2008 (Yun, 2010). Various local micro-climate models were applied such as monthly average of maximum daily temperature model (Chung *et al.*, 2009; Régnière *et al.*, 1995), heat island effect model (Choi *et al.*, 2003), minimum daily temperature (Chung *et al.*, 2006; Yun *et al.*, 2006), and the radiation (Yun, 2009) of the Korean peninsula with 30 m x 30 m resolution land surface data for predicting local climate condition at fine resolution.

Precipitation is also expressed but has 270 m x 270 m resolution due to the technical limitation

of the climate model (Daly *et al.*, 1994; Shin *et al.*, 2008). Various secondary products, such as average temperature, frost-free days, growing season precipitation, growing degree days, and chill unit (Table 2, acronym CU) were calculated from the primary products. The future microclimate for the entire Korean peninsula was also predicted by using the climate model based on Intergovernmental Panel on Climate Change's (IPCC) climate change scenario A1b and A2a. Both of these scenarios predict warmer climate than presently seen on the Korean peninsula but the A2a scenario assumes a more extreme change. Presence/absence data and environmental data were all processed by ArcGIS 10 (ESRI, 2011). The average aspect and slope of each grid cell were also calculated by the same program, which is based on a digital elevation map. All data were re-organized into a 270 m x 270 m resolution and arrayed to a row for each grid cell by a spatial joining function. In this fine resolution, more than a million grid cells were processed for South Korea. The suitable habitat model was built on the presence/absence data from the South Korean region, and this was due to the limitation of distribution data for the North Korean region. Only Korean peninsula region was predicted for the suitable habitat of tree species for this finer resolution. The years 2010, 2040, and 2070, which have thirty years of time interval, were used as the time points for the prediction. The 1980 map was made as well, but it was used only for checking the model prediction for the current habitat distribution patterns.

Coarse resolution maps were made by using 1 km x 1 km resolution GCM data from the International Centre for Tropical Agriculture (CIAT) climate change's downscaled data, which are developed by Decision and Policy Analysis (DAPA) program. The data are provided from the homepage of Climate Change Agriculture and Food Security (CCAFS, <http://www.ccafs-climate.org/data/>). The original data provide monthly maximum temperatures, minimum temperatures, average temperatures, and precipitation data that are predicted by various GCMs

(Table 1). A total of 13 different datasets, which include three scenarios and six GCMs, were used for future predictions. The resulting values in each grid cell from the future predictions were averaged to verify suitable habitats commonly agreed upon by the models. Coarse resolution maps cover not only the Korean peninsula but also the Manchurian region of China and Japan where there are major land areas containing the *PIDE* species. Predicted suitable habitats can be compared with a broader distribution area found in a schematic atlas (Fig. 2), but the finest resolution data for this research was found at 1 km x 1km. Model predictions showed difference in this resolution and extent. Since the source climate data were limited, climate data predicted for the years 2050 and 2080 were used for future predictions.

Random Forest Model

Random Forest (Breiman, 2001), a statistical model that is derived from Regression Tree analysis, was used to build a niche-based species distribution model of *PIDE* to predict presence/absence of the tree species in each grid cell. Regression Tree is a nonparametric statistical model based on binary recursive partitioning response variables by explanatory factors (Breiman, 1984; De'ath and Fabricius, 2000). It constructs a set of decision rules that divide the response dataset step by step, and builds a tree-like structure. Each split is chosen when it can maximize the distance between two divided groups. Since Random Forest is used for predicting the presence/absence of each point by using environmental information with the current presence/absence data, it is a niche-based modeling method that predicts the appropriate fundamental niche of the species without considering ecological interactions such as succession, competition, or disease. The model averages many numbers of regression trees that are grown with a randomized subset of predictors. Each regression tree is constructed by accumulated and binary split of presence/absence data set. Random Forest uses only a partial sample data set to

build each regression tree and repeatedly creates new regression tree by resampling without pruning (bootstrapping). To reduce bias errors in regression the randomization of predictors, averaging trees, and bootstrapping are all used. Averaging trees also enables prediction in gradient that is poorly expressed in single regression tree analysis. In this paper, Random Forest function in statistical program R was used to build the suitable habitat model. (Liaw and Wiener, 2002; R Development Core Team, 2010) The function bootstraps 64 % of the total observation. To build model for *PIDE*, half of the presence data and almost same number of absence data were randomly sampled with environmental data from the spot. 150 random trees were generated and averaged for each model. A none-sampled data set was used for calculating the AUC of Receiver-Operating Characteristic Curve (ROC) and also used was the Kappa value to check the model's reliability. Normally, Random Forest can perform the best when the numbers of presence and absence cells that are sampled are almost equal since inequality of the classes can make over- or under-estimated prediction (Breiman, 2001; Ledig *et al.*, 2010). Therefore, the number of samples for model building is regulated to have approximately the same number of presence or absence cells. However, no presence/absence portion regulating was used in this study since 1:1 presence/absence sampled model couldn't showed any better model performance. The portion of *PIDE* was big enough (36%) to evade an under-estimated prediction for presence. Conservative threshold 0.5 was used for determining presence/absence initially, and other possible thresholds that optimize sensitivity or specificity were analyzed by ROC plot as well. The model was calculated twenty times and the results are averaged. AUC and Kappa were obtained by averaging the results.

Table 1. List of GCM by Scenarios for 1km x 1 km Prediction Data.

The GCM scenarios for predictions are based on different Special Report on Emissions Scenarios (IPCC, 2000). A1 assumes maximum energy requirement but the sub-scenarios are based on different dependence on fuel sources on the base assumption. A1b assumes balanced dependence across fossil and non-fossil sources. Climate changing pattern of A1b is similar with current tendency. A2 scenario also assumes high energy requirement in more divided world. A2a scenario expects more emission and warmer change than A1b scenario. B2 scenario assumes lower energy requirement in more ecology-friendly world. B2a scenario predicts less climate change in the future.

GCM is a mathematical model of the general circulation of atmosphere and ocean that is used for making climate model with additional features such as terrain and human activities. Each GCM chases circulation dynamics by their own way and has different features. In this paper, The Random Forest model for suitable habitat ran individually for each climate dataset from each GCM. The results are combined in a map to see the overall agreement of GCMs and scenarios about suitability of the species at the time. More detailed information about GCMs can be found in the references.

| A1b | A2a | B2a |
|---|--------------|--------------|
| CCCMA-CGCM3.1 (Scinocca <i>et al.</i> , 2008) | CCCMA-CGCM31 | CCCMA-CGCM31 |
| CSIRO-MK3.0 (Gordon and Research, 2002) | CSIRO-MK30 | CSIRO-MK30 |
| NCAR CCSM3.0 (Collins <i>et al.</i> , 2006) | NIES99 | NIES99 |
| UKMO HADCM3 (Gordon and Research, 2002) | UKMO HADCM3 | UKMO HADCM3 |
| UKMO HADGEM1 (Johns <i>et al.</i> , 2006) | | |

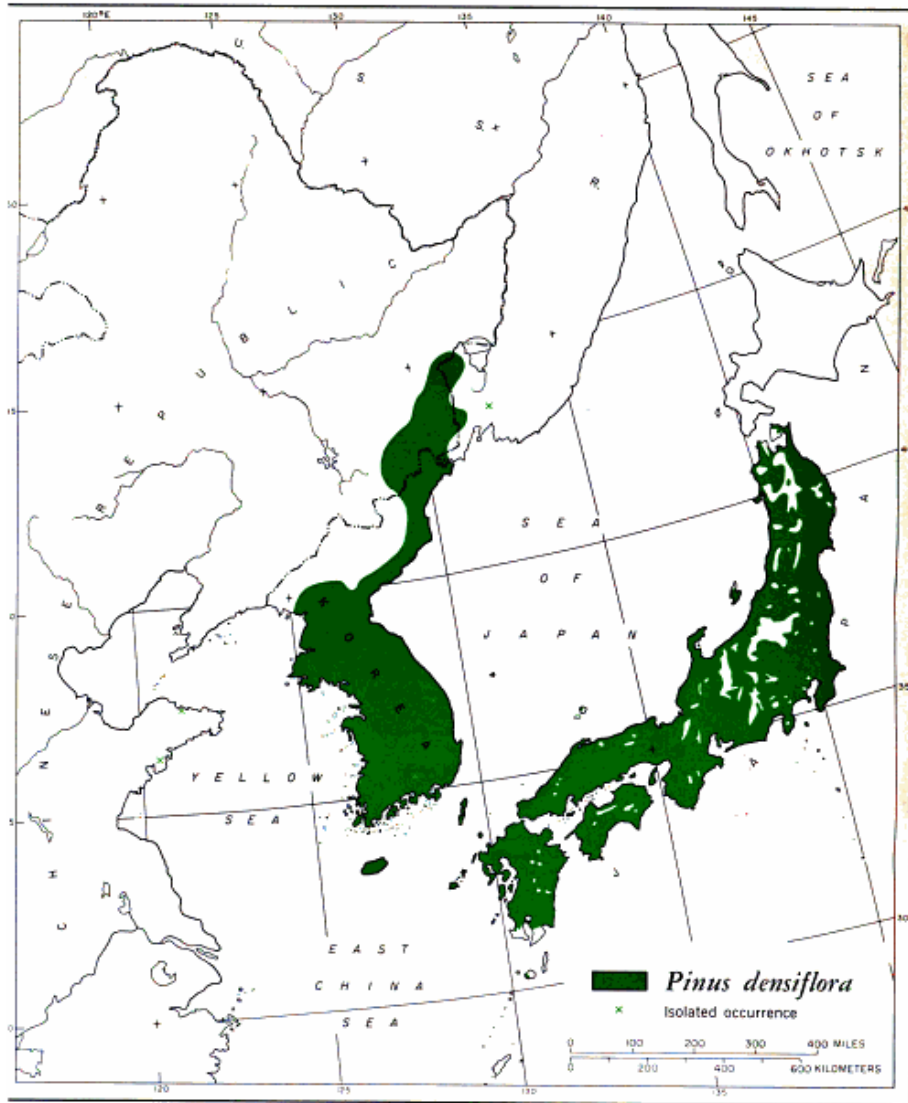


Figure 2. Approximate Known Distribution Range of *Pinus densiflora* (Critchfield and Little, 1966).

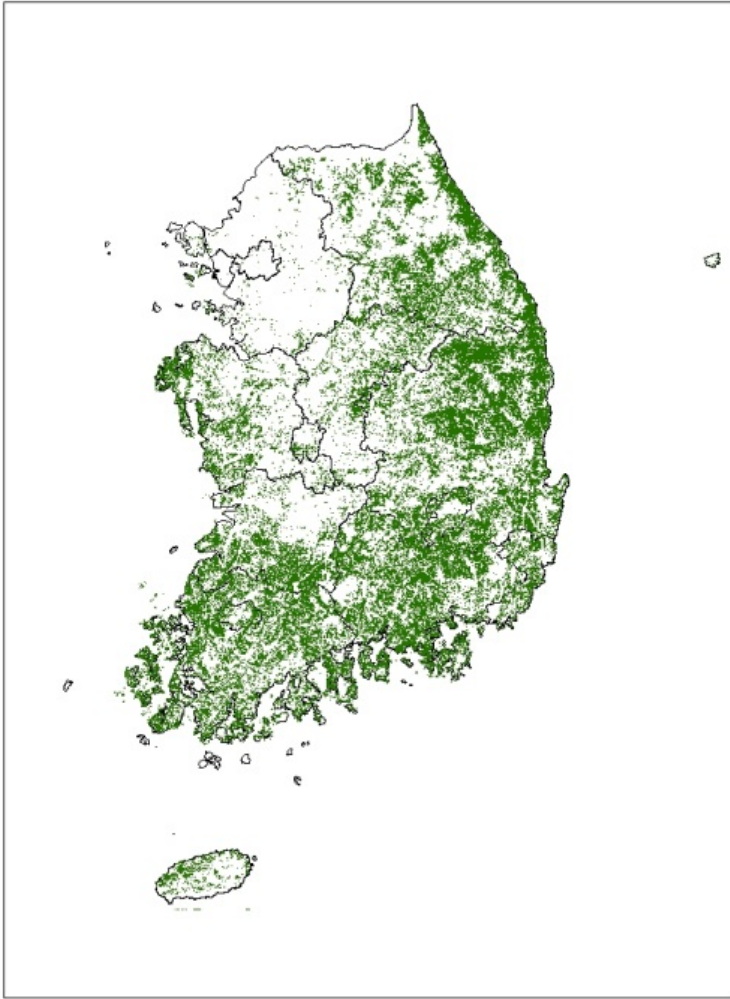


Figure 3. Current Distribution of *Pinus densiflora* in South Korea.

This forest map is in 270 m x 270 m resolution. Green spots show presence cells of the species. The black line is an administrative borderline. The west part of South Korea is a plain and has dense population with broad urban and agricultural areas while the eastern parts have broader forested areas but much less populations due to its mountainous terrain. *PIDE* is the dominant tree species in the mountainous areas.

Table 2) Acronyms and Definitions of Environmental Variables for the Suitable Habitat Model

| Acronyms | Definitions |
|----------|--|
| ADI | Annual dryness index: $(\text{GDD5})^{0.5}/\text{mean annual precipitation}$ (Ledig <i>et al.</i> 2010) |
| ASPECT | Average aspect of the cell by azimuth |
| AV1 | Average temperature of the coldest month (January) |
| AV8 | Average temperature of the warmest month (August) |
| AVT | Annual average temperature |
| 81TMP | Difference between average temperature of warmest month and coldest month |
| CU | Degree-days $< 0^\circ \text{C}$ based on the minimum temperature |
| GDD0 | Degree-days $> 0^\circ \text{C}$ based on the average temperature |
| GDD5 | Degree-days $> 5^\circ \text{C}$ based on the average temperature |
| GSGDD5 | Degree-days $> 5^\circ \text{C}$ based on the average temperature during growing season (April to September) |
| NFROST | Days in a year without frost |
| PRECI | Annual accumulative precipitation |
| 49PRECI | Growing season accumulative precipitation |
| RAD | Daily average radiation (Excluded in the coarse resolution models) |
| SLOPE | The angle of inclination by degree from 0 to 90 |
| TN1 | Minimum temperature of the coldest month |
| TX8 | Maximum temperature of the warmest month |

Environmental Variables

17 environmental factors were used for the fine resolution model variables. Table 2 shows acronyms and definitions for the variables. The coarse resolution model used 16 factors which does not contain radiation (RAD) because of lack of availability of these data. Elevation was often used as an important topographical variable for predicting current suitable habitat, but was not adopted into this study since the distribution range of altitude is a subordinate feature of climatic factors in most of cases, and elevation shift of the appropriate habitat by climate change was expected.

AUC and Kappa Statistics

The area under the receiver operating characteristic (ROC) curve (AUC) and Kappa statistic value were used to check the model's reliability. ROC is a two-dimensional graph that shows the relationship between sensitivity and specificity of the model's prediction which is broadly used to check the accuracy of diagnostic tests especially in medical field. Sensitivity means the ability of the model for picking the true value from true value while specificity means the ability for picking false from false. Accuracy of the prediction can be expressed by area under the curve of ROC. The value of AUC of ROC can be from 0.5 to 1 and 1 means perfect prediction for true and false values while 0.5 means no correct prediction. Above 0.7 means accurate prediction (Swets, 1988). Although Lobo *et al* 2008 pointed its limitation in use, AUC is widely being used statistic value for evaluating accuracy of species distribution models that are built from presence/absence data.

Cohen's Kappa coefficient (often simply called Kappa) is a statistical method that measures the degree of agreement between two raters. It is often used for evaluating agreement between actual distribution in digital forest maps and prediction in species suitable habitat modeling. It is

calculated by the formula $Kappa = (P_a - P_e) / (1 - P_e)$ (P_a = possibility of two rater agree, P_e = overall probability of random agreement). Kappa shows how two data sets agree well by gradient of value from 0 (total none agreement) to 1 (perfect agreement). A number between 0 and 1 can be interpreted as follows: below 0.2: poor agreement; 0.2~0.4: fair agreement; 0.4~0.6: moderate agreement; 0.6~0.8: good agreement; 0.8~1: very good agreement (Fleiss *et al.*, 1969).

RESULT

Fine Resolution Maps

The suitable habitat model for *PIDE* was made and AUC and Kappa of the model was calculated. A Threshold of 0.5 was used to judge presence or absence since the RF model in the R program results in continuous possibilities of presence from 0 to 1. 0.5 is a conservative threshold for species distribution model that uses ratio of presence/absence data close to 1:1. The area of current suitable habitat of *PIDE* was predicted with a 5 % smaller than the presence-detected area in the digital forest maps at this threshold. Specificity-sensitivity analysis revealed 0.46 is the threshold that maximizes similarity between observed presence and predicted presence with current environmental data. Predicted area with this new threshold didn't show big difference with 0.5 threshold result in current (2010) climate, but showed much bigger difference in future predictions (Table 3). Figures 2 and 3 show the digital forest maps of South Korea and the predicted suitable habitat maps of the entire Korean peninsula for *PIDE* at 270 m x 270 m resolution. The habitat model showed very high AUC, up to 0.96 in a few model runs among mostly 0.85 to 0.87 and moderate Kappa value from 0.51 to 0.52.

Slope, 81TMP, AV1, and 49PRECI were important environmental factors for distribution of *PIDE* (Fig. 4). Slope was the most important variable that determines *PIDE* distribution. This result is different from those of Kim *et al.* 2008 which analyzed the impact of topographical factors for *PIDE* distribution (Kim *et al.*, 2008). They showed aspect is a more important factor than slope in resolution of 100 m x 100 m. However, the maximum resolution of their research was 200 m x 200 m and their research site was a much smaller area than South Korea. The mismatch of results shows the importance of resolution in this kind of modeling research. Slope

was a more important factor than aspect for the suitable habitat in the resolution of 270 m x 270m and used successfully in this resolution of modeling. More *PIDE* was distributed in mid-slope mountainous areas than either the plains or in the steep high mountain areas. *PIDE* was distributed more in the mountainous eastern region of the Korean peninsula rather than the flat western plains. Even besides *PIDE*, almost no broad-leaf tree stand exists on the plain that has a low slope due to a high urban and agricultural development rate of these land surfaces on the Korean peninsula. In addition, *PIDE* does not exist on the steep high mountain regions of the Korean peninsula. Lee *et al.* statistically analyzed the spatial distribution of *PIDE* and *Quercus spp.*, and suggested that 15° as the most abundant presence slope for *PIDE* at the 100 m x 100 m resolution, and that is what his study also found (Lee *et al.*, 2006).

Table 3. Land Cover Percentage of Predicted Presence Cells in the South Korean Region.

This table shows the percentage of the predicted suitable habitat area of the tree species in South Korea in two climate change situation. A1b assumes current tendency and A2a assumes more warm and humid changes than A1b scenario.

Presence/absence of the species was determined at the threshold level of 0.5 (middle point of presence value 1 and absence value 0), but this table shows the percentage of the predicted habitat area using the threshold of 0.46 as well. This new threshold is the optimal value that predicted prevalence becomes equal with actual observed prevalence. The new threshold shows the decrease in *PIDE* suitable habitat areas in the climate change situation can be much milder than found for the conservative threshold 0.5, despite a small change of 0.04 for the threshold value.

PIDE (threshold = 0.5 / 0.46)

| 2010 | Scenario | 2040 | 2070 |
|--|----------|----------------|---------------|
| 35.7% / 37.6 % | A1b | 18.3% / 27.1% | 20.2% / 29.7% |
| | A2a | 22.6% / 31.0 % | 17.5% / 24.4% |
| <ul style="list-style-type: none"> • Distribution area in the Digital Forest Map: 37.5% | | | |

Other factors, such as temperature and precipitation can restrict the distribution of *PIDE*. Temperature factors that are related with winter, such as the difference of the average temperature between the coldest month and the hottest month and average temperature of January were both the next most important factors affecting the distribution of *PIDE*. Precipitation related variables were not that important, but growing season precipitation was the 4th most important factor. Lee *et al.* also showed that the relationship between the distribution of *PIDE* and the annual precipitation is not clear, but April's temperature had a clear impact pattern (Lee *et al.*, 2006). These results are well matched with the result of the random forest model that shows temperature has a bigger impact than precipitation on the distribution of *PIDE*.

Ranks of moderate important variables were changed often by model running for their similar relative importance. TX8, NFROST, TN1, ADI, and RAD were ranked 5th ~ 9th in relative importance table as a group. Although *PIDE* is very weak shade tolerant species, relative importance of solar radiation was moderate. Phonological variables such as GDD0, GDD5, GSGDD5, and CU were ranked with less importance. Annual precipitation was ranked as the least important variable and it shows the strong drought tolerance of *PIDE*.

These low ranked variables didn't largely determine the distribution of *PIDE* suitable habitats, but still affected marginal habitat cells. Removing one of these less-important variables didn't decrease AUC or Kappa values much, however removing more than a couple of less important variables steadily decreased the model's performance.

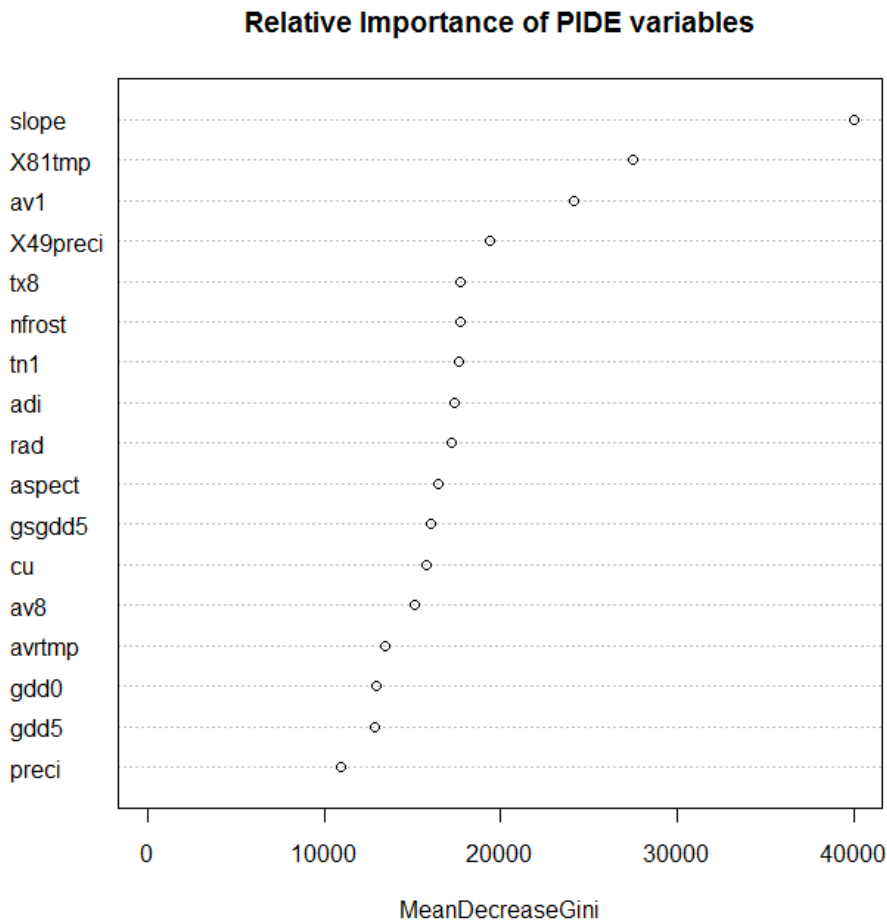


Figure 4. Relative Importance Table of Environmental Factors in the *Pinus densiflora* Models.

X axis shows importance score (raw error) of the variables. It is calculated: $a - b = \text{raw error}$. (a = the number of times a variable was correctly classified (based on the out of bag verification); b = the number of times a variable was correctly classified after classifications have been randomly assigned to the out of bag data set; raw error = the difference between the two). Therefore, the number of model regression trees or classified cell number affect the magnitude of the importance score, and thus it is only useful as a comparison to see the relative importance between the variables. SLOPE, 81TMP, and AV1 were revealed as the most powerful factors for all the three models.

The results of the presence/absence values determined by calculations of the Random Forest model were from 0 to 1 and the predicted area of presence was a bit less than the actual area. The number of cells which represent habitat area that have the value above the presence/absence

threshold of 0.5 was 5 % smaller than the number of cells of the actual forest map and had almost the same at the threshold of 0.46.

The current prediction map of 270 m x 270 m resolution for *PIDE* predicts correctly its real distribution pattern not only in the South Korean region but also in the North Korean region. *PIDE* is distributed almost everywhere except in the mid-western plains of South Korea. In contrast, the 1960 distribution map shows limited distribution of *PIDE* except eastern coastal area of North Korea (Fig. 2). The prediction map is only based on the species distribution data from South Korea but replicates this pattern almost exactly for both of the regions. However, *PIDE*'s presence in the mid-western plains was under-predicted from what the actual distribution of the species currently exhibits (Fig. 2, 3, 6).

Future prediction maps showed a quick decline of *PIDE* in the inland region and a shift of its distribution to the north, but a wide stable suitable habitat remaining (Fig. 7, 8, 9, Table 3). The *PIDE* predictions showed that current habitat areas will decrease but maintained coastal habitat stably and the extent of the overall habitat will increase northward. However, the shift in *PIDE* distribution pattern is not clearly observed before 2070 while its decline in the inland region was suggested as quick and obvious. These tendencies are accelerated in more extreme change situations, such as A2a scenario or for the 2070-2080 yearly maps, than shown for the milder change situations such as A1b or for the 2040-2050 maps. Although the result of the *PIDE* model for future scenarios show a low maximum possibility for presence and the area of suitable habitat is smaller than current area (Fig. 5), 20 ~30 % of South Korea was maintained as the suitable habitat for *PIDE* in all of the future scenarios (Table 3). This tendency is clearer with presence/absence threshold set at 0.46 than normal threshold value of 0.5.

In the A1b scenario which assumes a milder change than the A2a scenario, the area of an appropriate niche for *PIDE* decreased in 2040 but partially recovered in the warmer climates of 2070 especially in the inland region (Fig. 7). The niche for *PIDE* will be decreased in the eastern inland region, but maintained along the eastern and southern coast. The extent of *PIDE* habitat in the western plains will increase northward with time in this scenario. Gaema highland, which is located in mid-northern North Korea, clearly shows an increase of habitat range to the north in 2070.

The A2a scenario assumes a warmer climate change, the habitat expands northward more quickly but the area of the *PIDE* habitat keeps decreasing until 2070 (Fig. 8). At the presence/absence threshold of 0.5 in the scenarios, the area of habitat will decrease 30 ~50 % through time, but the intensity of the habitat area decreased was much milder in the optimal threshold values of 0.46 (Table 3). Almost the entire Korean peninsula including the mid-western plains becomes appropriate habitat for *PIDE* until 2070, and the appropriate habitat keeps expanding northward through time in this scenario. However, the major suitable habitat is still the region that is near the eastern and the southern coasts. The distribution patterns in these warmer scenarios were affected more by the pattern of precipitation changes.

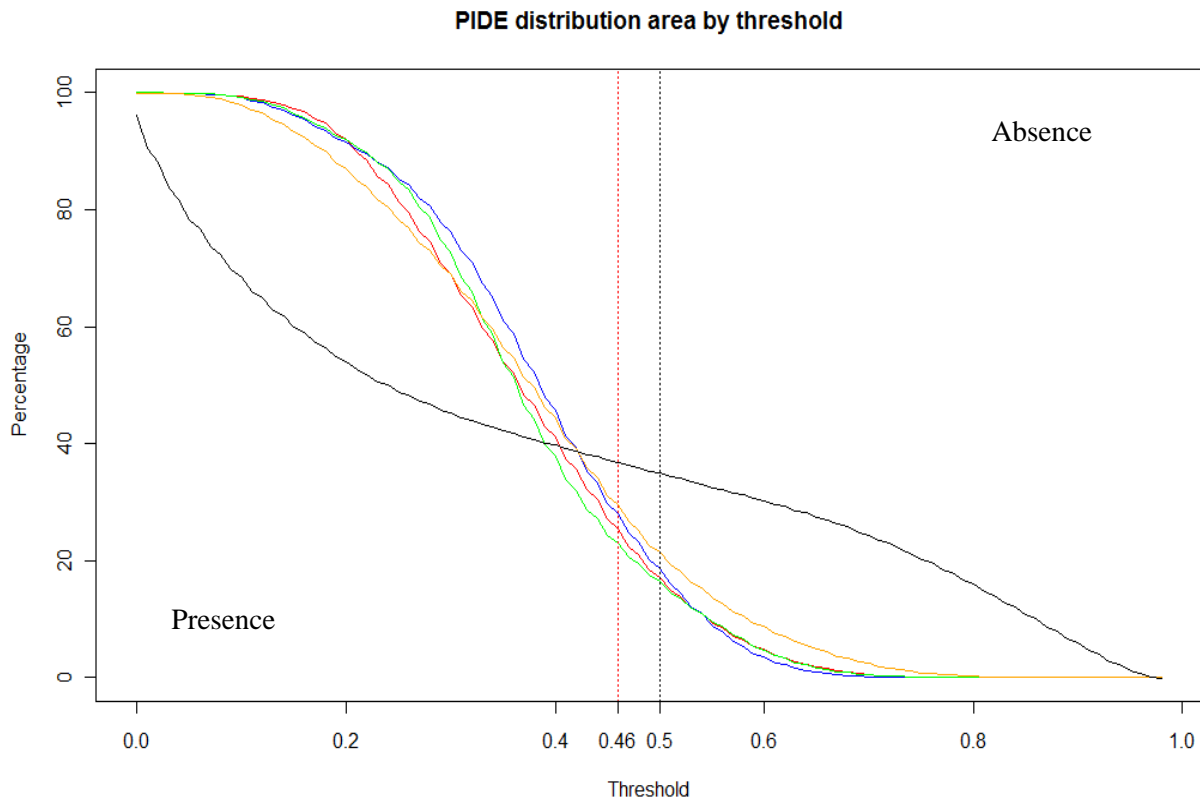


Figure 5. Change of *Pinus densiflora* Suitable Habitat Areas in the South Korean Region at Different Threshold Values.

Black line: prediction for current (2010), Red line: A1b 2040, Blue line: A1b 2070, Green line: A2a 2040, Orange line: A2a 2070

Area under the lines means percentage of predicted *PIDE* suitable habitat area in the scenario that the line represents. All of the future predictions show very steep slopes from 0.4 to 0.6. It means most of the region in the scenario has medium possibility of presence of *PIDE*. All of the future scenario predictions showed less variance of possibility of presence than current prediction. Most of the model result values are concentrated between 0.3 and 0.6.

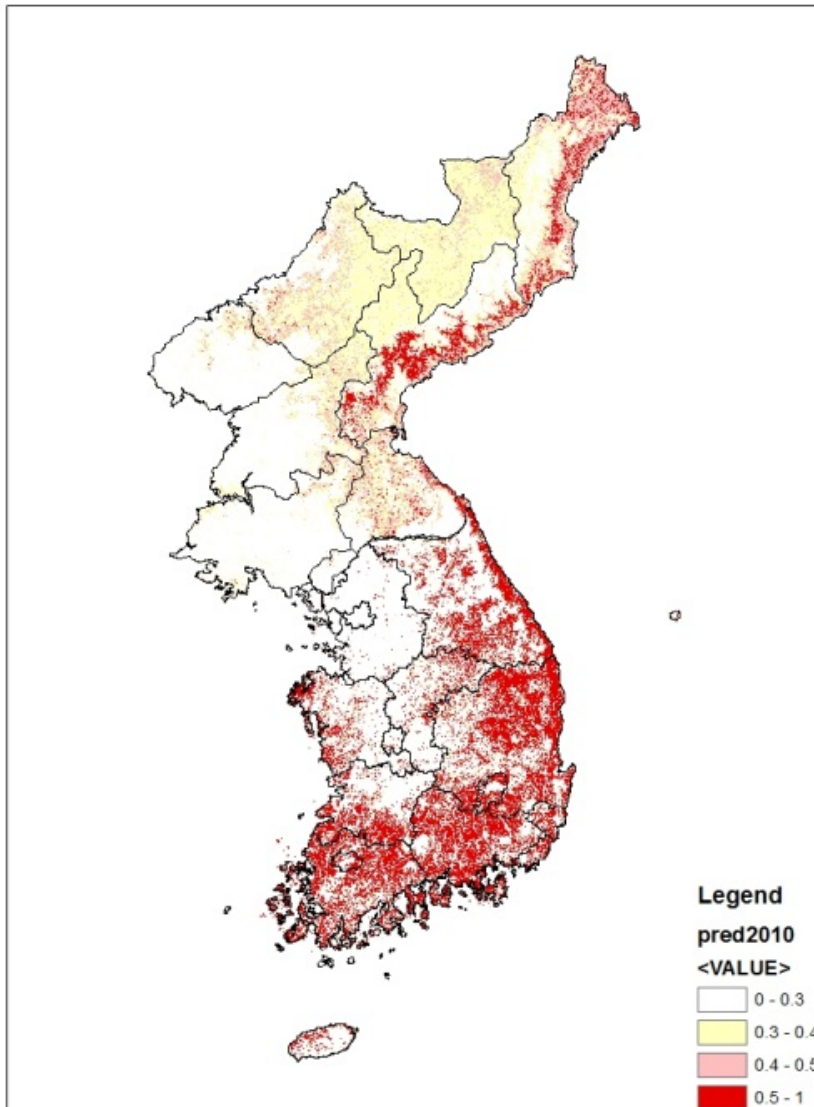


Figure 6. Model Prediction Maps for Current (2010) Suitable Habitat of *PIDE* at 270 m x 270 m Resolution.

Red shows the area was predicted as it has presence possibility bigger than the threshold (0.5) while pink (0.4) and yellow (0.3) show less likely presence predicted area. No color means area have possibility below 0.3. Random forest results consecutive value from 0 to 1 which means not likely present to totally likely.

Predicted area reproduced source data map very well. Mid-western region showed difference with atlas (Fig. 1), but also showed overall tendency of the species presence along south-eastern coast line.

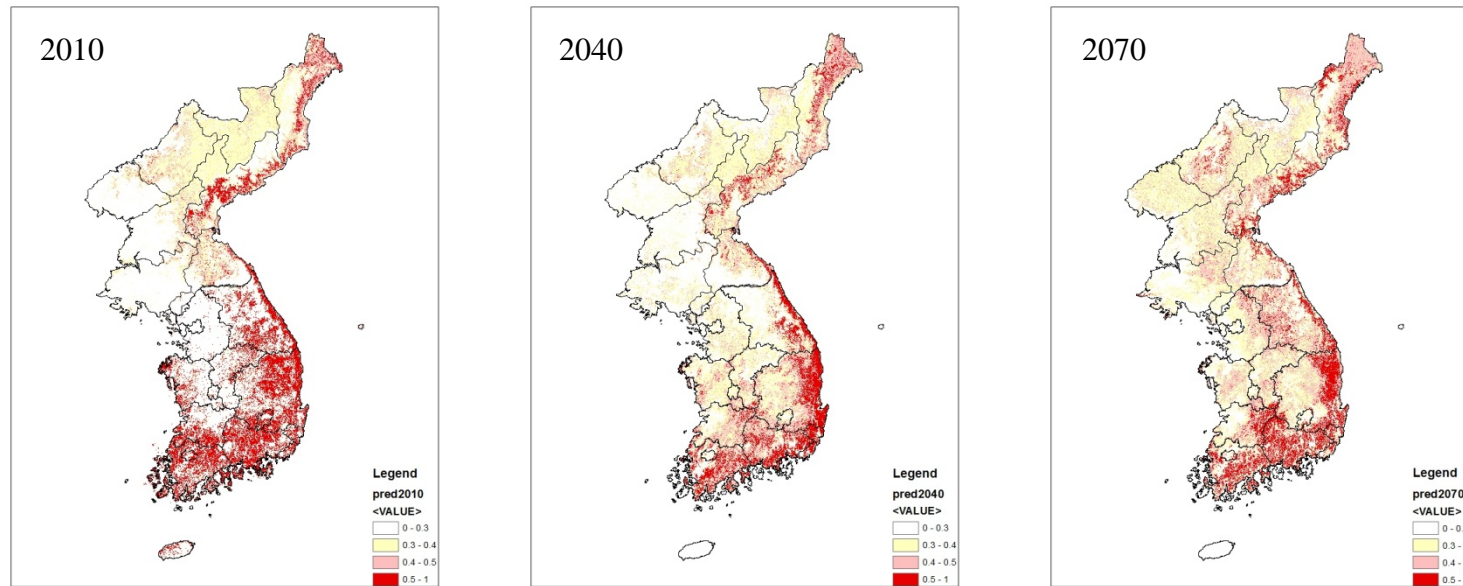


Figure 7. *Pinus densiflora* A1b Scenario Maps at 270 m x 270 m Resolution.

These maps are showing change of *PIDE* suitable habitat by IPCC climate change scenario A1b that assumes moderate change by current (2010, same map with Fig. 6) climate change tendency. While south-eastern coastal region maintains *PIDE* habitat, inland region suitability will decrease at 2040 and recover at 2070. Habitat area will decrease in South Korea region but increase in North Korea region. Islands are not included for future prediction maps due to data limitation.

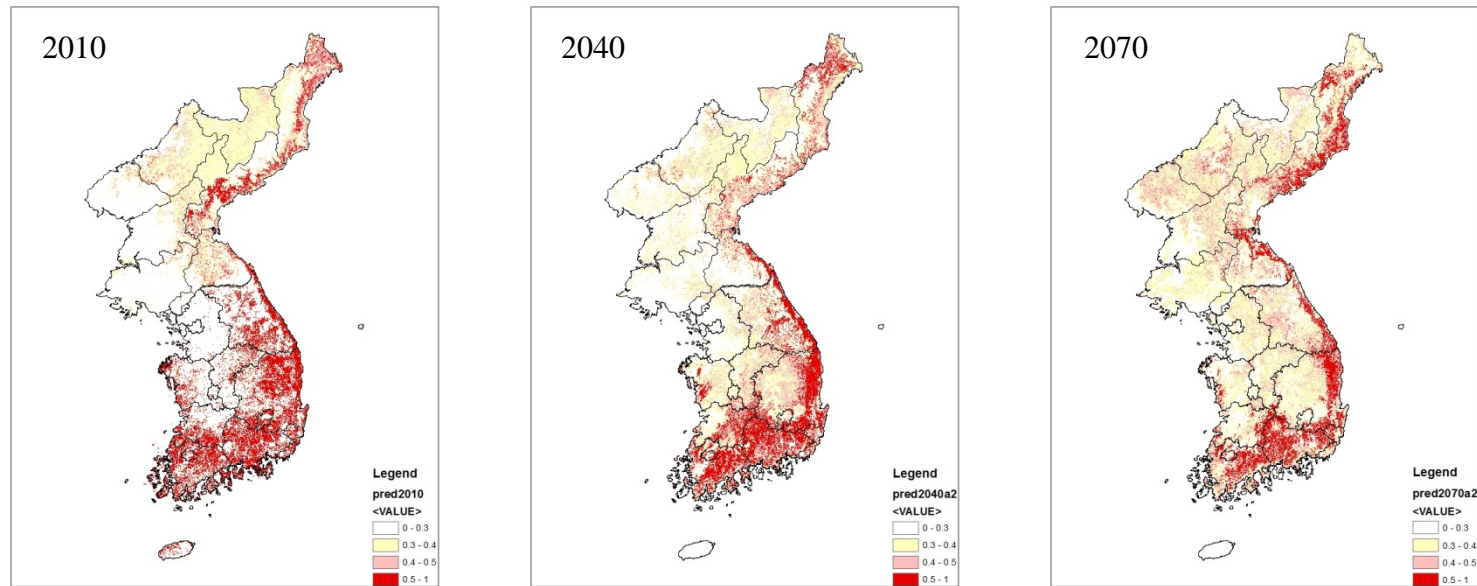


Figure 8. *Pinus densiflora* A2a Scenario Maps at 270 m x 270 m Resolution.

2010 maps is a same map with Fig.6. These maps are showing change of *PIDE* suitable habitat by IPCC climate change scenario A2a which assumes more extreme changes than current climate change tendency. *PIDE* suitable habitat will be changed similarly with A1b, but broader southern habitat will be maintained. Northern sparse suitable area also has broader extent than A1b scenario prediction. Overall, distribution extent will become broader to northward but suitable habitat area will decrease. Islands are not included for future prediction maps due to data limitation.

Coarse Resolution Maps

The suitable habitat model of *PIDE* was remade for making coarse resolution maps in the same way that fine resolution maps are modeled except for RAD, the solar radiation variable is not included due to the absence of the data. The remade 16 variable model has 0.02 less AUC value and 0.01 less Kappa than the 17 variable model that was used in building fine resolution maps. Ranks of importance between the variables were not changed after removing the RAD variable.

Model prediction for current suitable habitat of *PIDE* in 1 km x 1 km resolution and broader extent was well matched with the known distribution range map (Fig. 2, 9 - 2010). The model successfully predicted current known *PIDE* distribution not only for the Korean peninsula but also for the island regions of Japan and even the small habitat located on the Shandong peninsula in east China. The *PIDE* habitat distribution pattern of the Korean peninsula in this coarse resolution didn't show a big difference when compared with the fine resolution map. The eastern and southern mid-elevation regions were predicted as the major habitat for *PIDE*. The mid-western plains were still underestimated when compared with the known distribution range. Slope plays an important role in this phenomenon. The appropriate slope of a small hill area surrounded by plains can disappear more easily in this coarser resolution.

Model prediction agreement distribution showed similar pattern with the A1b scenario prediction of the fine resolution map that assumes moderate warming. Future predictions of the climate data at the 1 km x 1 km resolution were made for the 2050 and 2080 decades time period instead of the 2040 and 2070 decades due to the limitation of source data, but the pattern of change was similar. Although the predicted *PIDE* area decreased, the range extended northward.

The appropriate habitat for *PIDE* in South Korea and in regions of Japan region has not changed much, however the distribution area extends inland and to the north.

The difference between scenarios was also similar using a fine resolution map. Warmer A2a scenarios show broader areas and an extent for the *PIDE* range in particular in the South Korean region than moderate A1b or much less warm B2a scenarios.

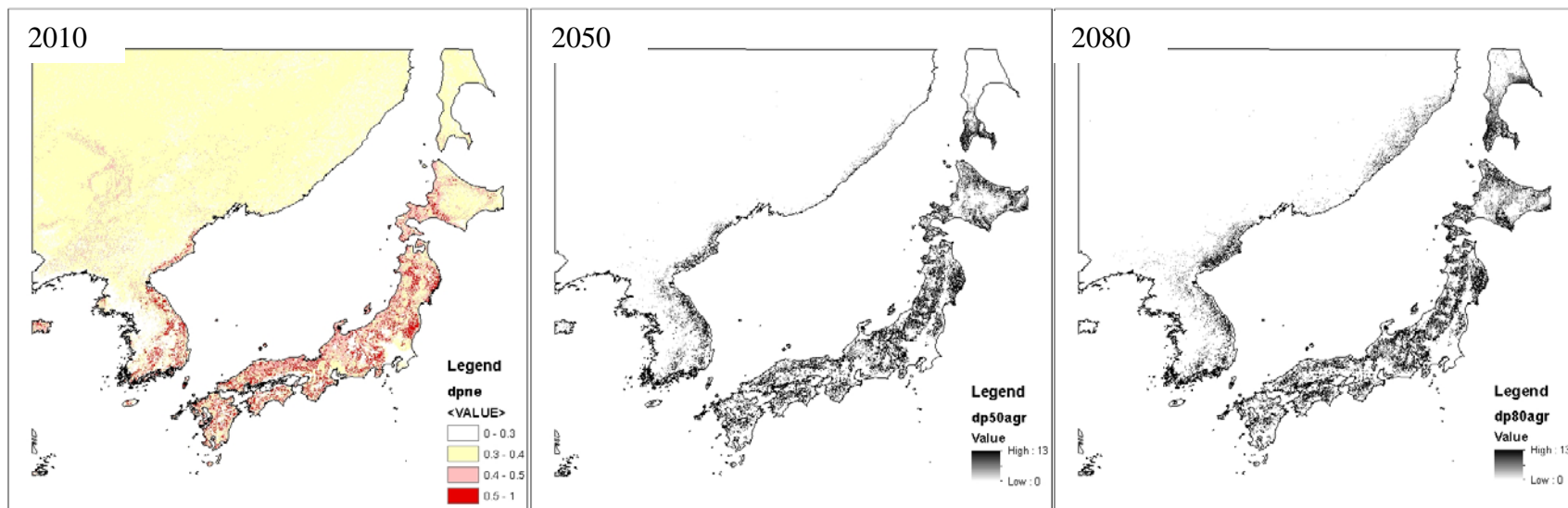


Figure 9. Estimated Future Habitat of *Pinus densiflora* at 1km x 1 km Resolution

2010 map shows a result of one current climate data based prediction. Red spots mean suitable habitat that over threshold 0.5. 2050 and 2080 maps are showing agreement of total 13 different maps from 3 different climate scenarios and 6 different GCMs. The darker colors indicate better model agreement. In this resolution, *PIDE* shows almost no change for its suitable habitat in the future except expanding to northward slightly. The yellow to red colored current map shows a result value from one dataset about current climate.

DISCUSSION

The Model Performance

The Random Forest suitable habitat models for *PIDE* successfully predicted the current extent of the species. The AUC value of the model was high enough as it is in the range from 0.85 to 0.96 by model running. However, the Kappa value which indicates how well the prediction value for total cells in the spatial range is matched with the actual data was from 0.51 to 0.52 which is just moderate. Since Kappa evaluates compatibility of all cells in the range, its value has a more direct relationship with the similarity of area of distribution and predicted habitat. When the Kappa value approaches 1, the predicted area also closely matches to the actual area. If the model prediction shows any difference with reference distribution, Kappa responds more sensitively than AUC, since Kappa counts one change twice while AUC does not. If a false prediction is made by the model, AUC evaluates it just once for a false presence or a false absence. However, in Kappa calculations, a false prediction affects the value twice since the false prediction changes the result in an increase of false absence and a decrease of true presence in the contingency table.

The combination of a high AUC and a moderate Kappa value can be interpreted in two ways. First, it can mean some areas have appropriate environmental conditions for *PIDE* habitat but there is no tree species present due to some other reasons. Although the impacts of urban or agricultural areas are accounted for slope, human activities can still generate this type of error. Successional substitution by late successional species also can explain the high AUC and low Kappa. It means some fundamental niche of *PIDE* is not realized and the model prediction reveals this fundamental niche. Second, non-environmental factors existing in the past can result

in the presence of a species outside of its potential environmental habitat today and therefore facilitate the model's overestimating the environmental habitats to which a species can be adapted. The Random Forest model calculations are only based on the given distribution of the species and environmental factors, so the impacts of other influential factors can distort the model's predictions. For example, human management can facilitate the survival and growth of a species even in what would ordinarily be considered an inappropriate habitat for that species. In that sense, because *PIDE* has been protected historically by the government, perhaps some aspects of its predicted habitat distribution can be partially explained by this potential error. However, the model did show overall good prediction for both of the extent of the *PIDE* habitat and its area which is perhaps explained by the fact that it was not planted or intensively managed for the last hundred years.

Reliable performances of the model to predict the extent of the suitable *PIDE* habitat reflect appropriateness of spatial ranges in source species distribution data. Appropriateness of spatial ranges is reinforced in the modeling by ensuring that the entire input ranges reflect the diversity of expected habitat and environmental indices to decrease the probability of potential over- or under- estimates of species distribution. This is especially important to consider for those area that could be considered as marginal areas of the distribution pattern (Thuiller *et al.*, 2004).

Mostly *PIDE* is distributed in Korean peninsula and on the islands of Japan (Fig 2). The extreme southern distribution limit is found on Yaku Island of Japan, but *PIDE*'s general southly distribution limit is found around Kyushu Island of Japan. The climate on Kyushu Island is similar to the climate found on Jeju Island of South Korea which was incorporated in *PIDE* suitable habitat model of this study. Mountainous regions can reflect a northern environmental

PIDE habitat similar to that exhibited by northerly latitudes and these environmental parameters were used in the source *PIDE* dataset (Fig. 3).

However, other limitations of niche-based species distribution models still should be considered when the result of *PIDE* suitable habitat model is interpreted. Niche-based models may be inaccurate in the long-term predictions (Loehle and LeBlanc, 1996) due to the limitation such as below. First, niche-based models are weak to predict species distribution at the outside of the source distribution data since they are not prepared to new environmental settings that that is not incorporated in the source data. This feature makes incorporating entire environmental range of the species into the source data very important. Thuiller *et al* (2004) showed this limitation very well by comparing niche-based species distribution model performances in Europe. Second, niche-based species distribution models could miss important environmental factor by impacts of non-environmental factors that are hard to express spatially. Niche-based models don't assess the impact of 'ultimate' factors, such as phylogenic, evolutionary, or geological histories of the species (Morin *et al.*, 2007). Moreover, ecological interactions, such as competitions, predations, or symbioses, are also hard to be incorporated in niche-based models. However, these non-environmental factors form realized niche that are shown in the actual species distribution data, and niche-based models predict fundamental niche from the realized niche data (Leibold, 1995). Thus niche-based models could miss appropriate environmental ranges that are currently unavailable for the species by non-environmental factors. Third, Niche-based models are not based on the physiological processes of the species and thus do not reflect phenotypic plasticity and local adaptations (Morin and Thuiller, 2009). For example, appropriate range of humidity or temperature for a plant species can be changed by level of CO₂ in the atmosphere. Therefore,

niche-based species distribution models should be interpreted with considerations about historical and physiological features of the objective species.

Variables

Among the environmental factors that were used for the suitable habitat model, slope and temperature differences between summer and winter seasons were the most important environmental variables. Slope was the most important factor for explaining *PIDE* habitat distribution. Strong positive correlations of slope with the presence of *PIDE* habitats explained why there was an absence of the species predicted in the western plains area of the Korean peninsula (Fig. 6) when in fact the species is found there today in non-urban settings (Fig. 3, 11). However, this minor slope anomaly didn't diminish the overall performance of the model. For example, the model prediction of suitable *PIDE* habitat using the 1980 climate dataset from HD-DCMs shows a very close match between the 1980 prediction and that of the 1960 atlas in the western plains region (Fig. 10). The explained correlation between predicted *PIDE* habitat and slope in the model is assumed to be a result of the intensive human activity in the plains relative to the upslope regions. Dense human populations assert an enormous amount of pressure to develop much of the plains areas in the Korean peninsula into urban or agricultural areas. Even the small plains areas on the slope of the mountains have also been developed as farmlands. However, small hilly areas on broad plains could support forest stands. This can even be seen around graves located on hills or mountains, especially where cultural tradition has maintained *PIDE* stands. This social context has resulted in a strong contrast with predicted tree species distribution on slopes in Korean peninsula. Without human intervention, even other topographical factors, such as aspect, can influence the natural distribution patterns of tree species (Kim *et al.*, 2008). However, the effects of aspect on tree distribution patterns are also

covered and explained by the effects of temperature and soil moisture, and therefore don't add any more to the power of modeling and predicting tree distribution in South Korea.

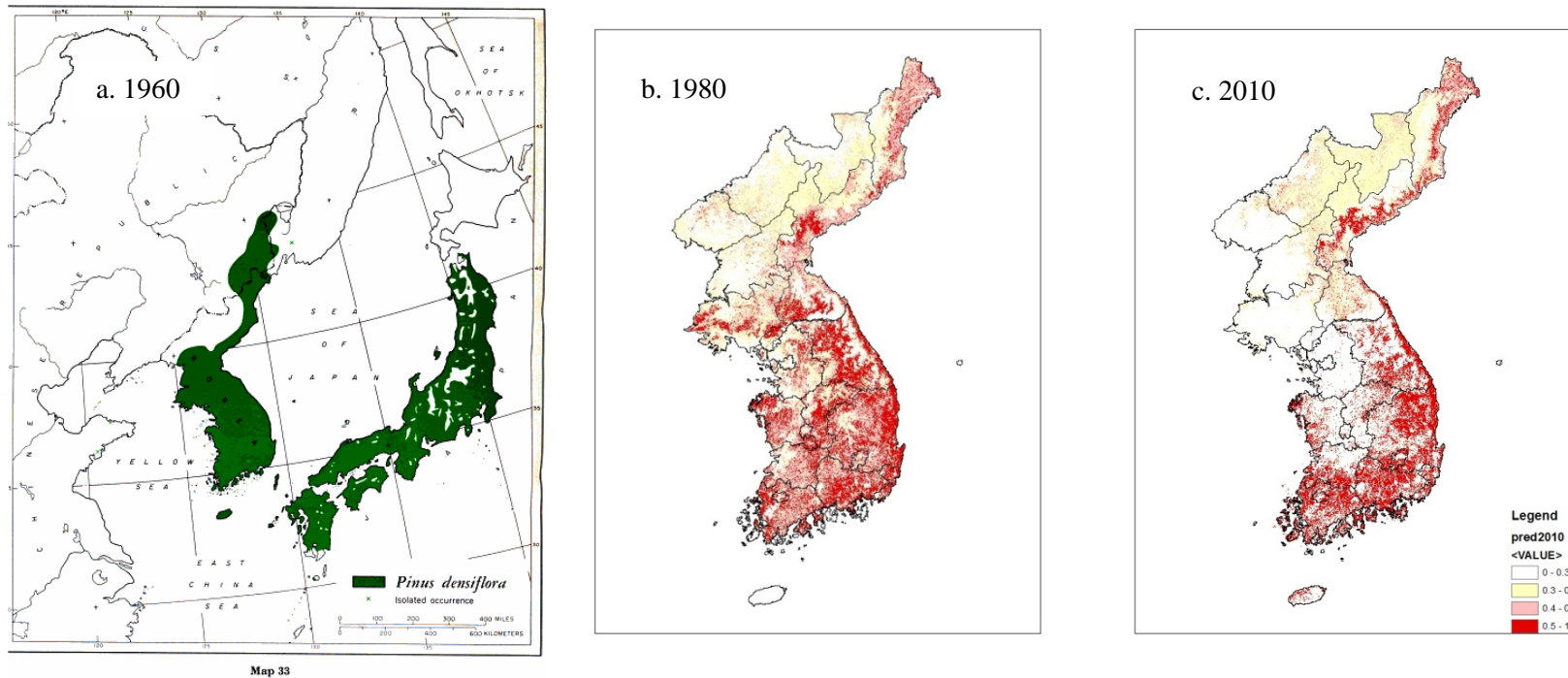


Figure 10. 1980 Prediction Map of Suitable Habitat for *Pinus densiflora* and Species Distribution Atlas from 1960s.

1980 prediction map (b) shows better agreement with distribution atlas from 1960s (a) about the extent of *PIDE* than current habitat prediction map (c). Western plain region was also a part of suitable habitat for *PIDE* at the time point. The *PIDE* model can predict low slope area as its suitable habitat if the other conditions are suitable as this 1980 prediction map shows. Change of suitable habitat distribution shows decline of habitat suitability in mid-western region. Islands are not predicted in 1980 map due to data limitation.

Resolution can play an important role in performance of environmental variables in the model. The *PIDE* habitat model showed some small suitable habitats in the mid-western plains region at the 270 m x 270 m resolution, but the small habitats almost disappeared at the 1 km x 1 km resolution. This disappearance of the small habitat is perhaps best explained by the fact that the slope impact is apparent at the finer resolution and depicts small hilly areas that contribute to tree habitats. But at the coarser resolution, only the plains are apparent and thus the slope impact disappears which decreases the model's prediction of *PIDE* habitats in that area. Aspect can be important in depicting *PIDE* habitats when using finer resolutions, but again at the coarser resolutions the aspects are averaged over the larger area decreasing its impact on habitat mapping. Aspect was not among the most important factors in modeling tree habitat patterns in this study, however at the 100 m x 100 m resolution Kim *et al.* in 2008 showed that aspect was an important factor for mapping *PIDE* distribution.

The second most important factor correlating with the habitat distribution patterns of *PIDE* is 81TMP (i.e., the difference between average temperature of warmest month and coldest month) and this reflects the physiological influence of temperature on the species. This influential factor shows the adaptability of the species to large fluctuations of seasonal temperatures. This species is found in the temperate forests that has four distinct seasons and temperature changes. This physiological characteristic also appears in different forms of the winter/summer monthly temperatures such as AV1 (i.e., average temperature of the coldest month-January), AV8 (i.e., average temperature of the warmest month-August), TN1 (i.e., minimum temperature of the coldest month), and TX8 (i.e., maximum temperature of the warmest month) in the relative variable importance table (Fig. 4). The relative importance of AV1, which is located at 3rd, shows importance of seasonal temperature. Certainly as a conifer species, *PIDE* has advantages

for its ability to grow during the winter, but on the other hand, the average temperature of the hottest month (i.e., AV8), was one of the least important factors. Importance of minimum and maximum monthly temperatures (TN1 and TX8, respectively) didn't exhibit any significant features and their importance was just considered moderate and similar with each other in this study.

Annual precipitation was the least important factor while growing season precipitation is ranked as the 4th most important factor in the model (Fig. 4). Following are some facts that explain this. First, annual precipitation doesn't affect much to mature *PIDE* since it has very strong drought tolerance (Cho, 2007). Moreover, the fact that seas surround Korea contributing to a stable and sufficient amount of annual precipitation *PIDE* grows well in most of the Korean peninsula. The annual precipitation is dominated by rain clouds from the ocean and not different in that narrow regions while factors such as altitude, aspect, and the terrain of the Korean peninsula can contribute to and create the large difference seen in temperatures even in the narrow regions. In other words, annual precipitation does not differ much by region at the extent and location of the Korean peninsula. These same factors also moderate the variations of annual precipitation found on the islands of Japan which is another main distribution region of the species. However, variations in the temperatures do exist because of the complexity of the terrain which affects factors such as the humidity and the complexity even creates arid conditions for short periods especially during the warmer growing seasons. For example, some physiological studies point out that appropriate amount of precipitation during the spring growing season is very important for the growth of *PIDE* seedlings. Park revealed that the main factor limiting for growing of *PIDE* in South Korea is the soil water content during the spring growing season rather than by temperature itself (Park, 2000). Lee *et al.* also showed the importance of spring rainfall for *PIDE*

seed germination (Lee et al., 2004). These spatial and temporal differences in environmental parameters can contribute to the mosaic patterns of habitat distribution. In this study, the modeling shows differences between importance value of annual precipitation and the growing season precipitation and the correlation between growing season precipitation. These results reflects physiological feature of the species.

Regression Tree analysis for the same data set can show this pattern clearly. Regression Tree analysis for the same subset of the Random Forest model can show a ‘tree’ in the ‘forest’. The tree doesn’t show the total structure of the forest, but can be a clue for analyzing the forest. Two Regression Trees for the same dataset derived from the *PIDE* Random Forest habitat model created splits of data at levels of significant parameter influence on *PIDE* habitat (Fig. 10). The orders in which specific variables significantly split the data in the Regression Tree closely matched the orders of variables derived from the importance table from the Random Forest model (Fig. 4, 11). Table 4 shows the Regression Tree split points of important variables that were revealed in the sample Regression Trees. Important model behaviors can be interpreted from these results. For example, the Regression Tree from the *PIDE* model determined that environmental conditions such as mid-slope, low summer-winter temperature differences, relatively mild maximum summer temperature, and mild average winter temperature are all good features for *PIDE* presence.

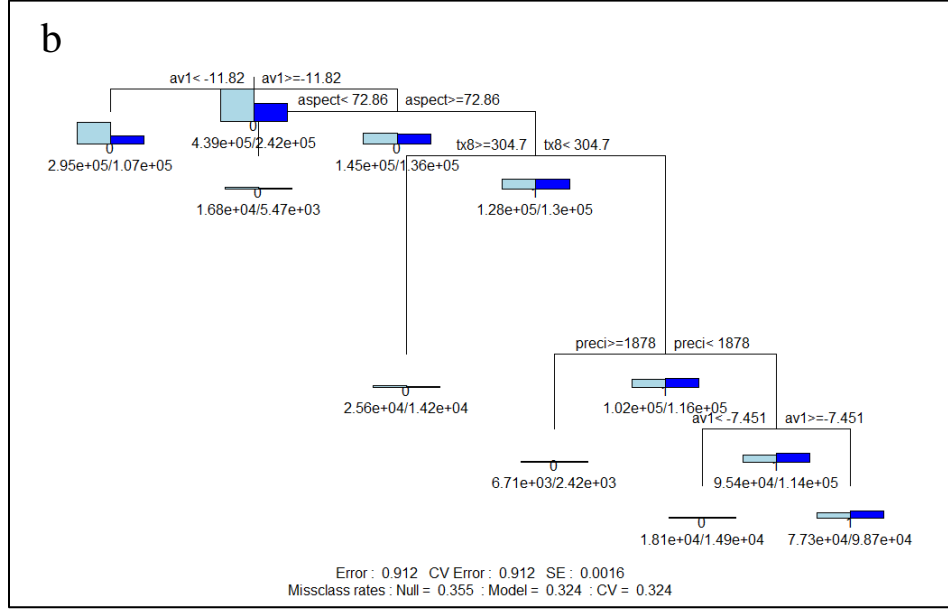
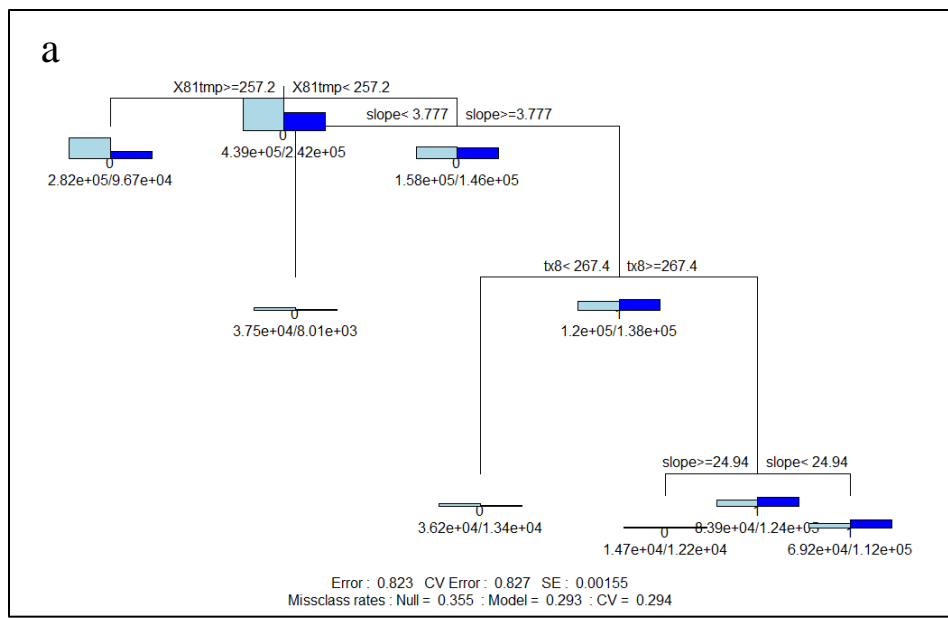


Figure 11. Sample Regression Trees for Variables Used in the *Pinus densiflora* Random Forest Model.

This figures show example Regression Trees for *PIDE* Random Forest model dataset. The Random Forest variable importance rank was slope – 81tmp – av1 – 49preci – tx8 – nfrost – tn1 – adi – rad – aspect – gsgdd5 – cu – av8 – avrtmp - gdd0 – gdd5 – preci. This example Regression Tree shows the importance of slope and 81tmp in predicting *PIDE* habitat in the model.

- a. The first example Regression Tree for *PIDE* model with full dataset.
- b. The second example Regression Tree for *PIDE* model without slope and 81tmp

Table 4) Split Points of Variables that Affect *Pinus densiflora* Habitat Distribution are Presented in the Sample Regression Trees

Random Forest model averages Regression Trees from random subset of source data and votes the best split point for each variable (Breiman, 2001). This Regression Tree split point table is derived from entire dataset rather than subsets (Fig. 11). Therefore, this table shows approximate examples of split points rather than actual split points of *PIDE* Random Forest model.

| Full variables (all parameter included) | 1 st split | 2 nd split | 3 rd split | 4 th split | 5 th split |
|--|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 81TMP (°C) | > 25.7 | | | | |
| SLOPE (degree) | | > 3.8 | | < 24.9 | |
| TX8 (°C) | | | > 26.7 | | |
| Partial variables (without 81TMP and SLOPE) | 1 st split | 2 nd split | 3 rd split | 4 th split | 5 th split |
| AV1 (°C) | > 11.8 | | | | > 7.5 |
| ASPECT (degree) | | > 72.9 | | | |
| TX8 (°C) | | | < 30.5 | | |
| PRECI (mm) | | | | < 1878 | |

The Future of Pinus densiflora

Interesting results were observed when the information mentioned above was included in the model to predict the future changes in the habitat distribution for *PIDE*. Predictions of changes in forest habitat distribution in the future for the Korean peninsula mostly have concentrated on successional processes, heat stress and pest and anticipated replacement of *PIDE* stand by *Quercus spp.* (Kato and Hayashi, 2006; Kato, 2003; Kim, 1992; Kim, 2003; Song, 1997; Toyohara and Fujihara, 1998). A report in 2001 from the Korean Forest Research Institution

even suggested that *PIDE* could totally disappear in South Korea by 2060 under current rate of changes (Jeon, 2001; Jeon *et al.*, 2002). Moreover, the report mentioned if global warming occurred in combination with natural successional processes (converting *PIDE* to deciduous species) and increasing pests, *PIDE* could even disappear from the entire Korean peninsula by 2010. Choung and Hong observed changing of *PIDE* dominated forests to coniferous/deciduous mixed forest composition that is dominated by *PIDE* and *Quercus mongolica* at higher altitudes or latitudes and by *PIDE* - *Quercus serrate* at lower altitudes or latitudes (Choung and Hong, 2006). Other studies also reported the dominance of *Quercus spp.* in younger generations of *PIDE* dominated forest (Kim *et al.*, 2000; Kim, 2003; Song, 1997). These studies agree with the well-known prediction that *Quercus spp.* will succeed *PIDE* habitats in the Korean forests. These kinds of predictions have forecast a decrease of *PIDE* stands and an increase of *Quercus spp.* stands on the Korean peninsula for the future (Kim and Kil, 1996).

However, the Random Forest modeling of the future to include predicted climate change showed a different future for *PIDE* habitat distribution patterns. The future predictions of the model showed relatively stable areas and an increasing extent northward for the suitable habitat of *PIDE*. At the same time, the future of *Quercus spp.* in the Korean peninsula is not that bright. Although *Quercus spp.* is one of the most dominated tree species in Korean peninsula, it is not dominant deciduous tree species in the southern coast and islands since it is a cool-temperate zonal species (Song, 2008). According to the climate change scenarios, the climate of the Korean peninsula will be changed into sub-tropical climate which exists in the southern coast and islands currently. It means *Quercus spp.* would lose broad suitable habitats in the Korean peninsula by the warmer climate in the future. Random Forest modeling for *Quercus spp.* which is limited in the Korean peninsula showed a result corresponds with this estimation (See Appendix 2). In

other words, the model prediction suggests a new possibility that if the climate will change to an extreme, *Quercus spp.* could decline and the *PIDE* stands would therefore be able to have more time against the successional pressures until other sub-tropical, shade tolerant, late successional tree species arrive and mature. Kim 2003 anticipated *Q. monglica* stands could be succeeded by *Abies holophylla*, *Acer momo*, *Fraxinus rhynchophylla*, *Tilia amurensis*, and *Acer pseudosieboldianum*, all of which have better shade tolerance. With this in mind, Park predicted the spread of evergreen, broad-leaved species northward under a climate-warming scenario (Park *et al.*, 2010). In the climate change scenario some of these late successional species just discussed and southern evergreen species are expected to join to the current *PIDE* - *Quercus* dominated forests. *PIDE* stands could be sustained for a while until new species become established and mature.

There are some studies and reports that correspond with the Random Forest model result and interpretation. Kim contended that with global warming the *PIDE* distribution range will be expanded, rather than shifted, to the north in preference of the tree species for the warmer-drier climate (Kim, 2005). Recent reports about the increase of *PIDE* population and expansion of habitat areas of the species toward higher elevations on the warm island regions, such as Ulleungdo and Jeju, would lend support to these expectations. The 2011 report of the Korean Forest Research Institution shows four folds increase of *PIDE* habitat at mid-elevations on Mt. Hanra of Jeju which is the biggest island that is located at the south of South Korea where a sub-tropical climate exists. The wide distribution range of the *PIDE* habitat on the islands of Japan stretches as far south as part of Yaku Island (42°40' N) and also confirms that the species can successfully grow in very warm climates (Hayashi, 1952).

In the future increases in winter temperatures and changes in the growing season precipitation can strongly facilitate the expansion of the current *PIDE* habitat ranges into regions that are considered unsuitable currently. Yoon analyzed and predicted the impacts of topography, temperature, precipitation, and soil on the current and future spatial distribution of *PIDE* using statistical regressions (Yoon, 2003). Yoon suggested the optimal environmental ranges for growing *PIDE* were on slopes between $0^{\circ} \sim 26^{\circ}$, annual average temperatures between $10^{\circ}\text{C} \sim 14^{\circ}\text{C}$, and annual precipitations between 1,100 mm \sim 1,500. He argued that when factors were separately considered, temperature was a more important factor than precipitation when examining predictions of *PIDE* habitat. The results of the Random Forest modeling in this study matched Yoon's results that are mentioned above. The *PIDE* distribution extent and decrease in area predicted by Yoon for future climate change scenarios used only temperature. Choi *et al.* also predicted shrinkage of *PIDE* and *Pinus koraiensis* habitat in the South Korean region by analyzing the Warmth Index and the Minimum Temperature of the Coldest Month Index (Choi *et al.*, 2010). However, the Random Forest modeling uses the factors together to predict suitable habitat. The Random Forest model showed results that were similar with decreases in the distribution pattern of *PIDE* related to just temperature for most of the future climate change scenarios. However, this study found different habitat distribution patterns for *PIDE* that were affected by spatial patterns of precipitation (Fig. 12, 13). It appeared that *PIDE* maintained most of its southern habitat areas except in the dry inland, mid-eastern mountain regions where it is affected by the down-slope Foehn wind occurring on the dry rain-shadow sides of the mountain ranges during the spring growing season. These future prediction results are interpreted that precipitation becomes a more powerful factor affecting species distributions in a relatively warmer situation. This could be explained by the fast increase of the air temperature during the

growing season could induce a fast increase in evapotranspiration, and thus limit the growth of *PIDE*. Climate change is predicted to bring 5% more precipitation in the East Asia region (Min *et al.*, 2004). Considering that *PIDE* is a strong drought tolerant species, even small increases in precipitation during the growing season by climate change could benefit the species (Cho, 2007; Lee and Lee, 2003). In contrast, *Quercus spp.* wouldn't see large benefits in this scenario because they are not as affected by small increase in precipitation due to its ability to adapt to a larger moisture range (Lee and You, 2009). It means that a warmer and more humid climate change scenario could present a physiological advantage to *PIDE* rather than the *Quercus spp.* across a broad region of the Korean peninsula. However, different environmental conditions that are not incorporated into the *PIDE* model, such as predicted changes in the concentration of atmospheric CO₂, can alter the physiological response and competitiveness of the tree species in the future. Thus, acquiring more physiological information is required for more precise future predictions.

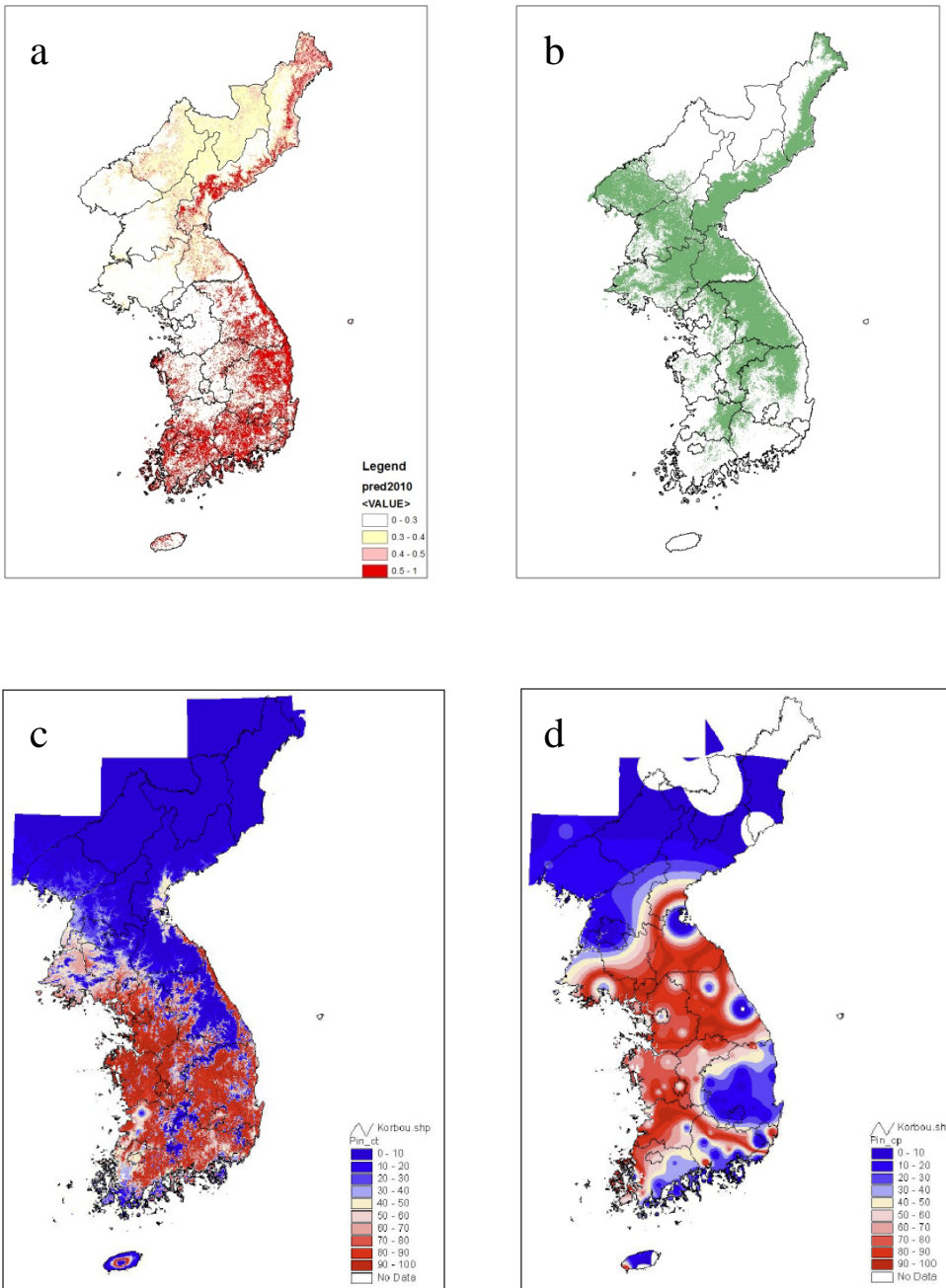


Figure 12. Comparing Random Forest Model Prediction and Previous Model Predictions for *Pinus densiflora*.

a: Random Forest model prediction for current climate. It shows better prediction ability for the southern end habitat of the species.

b: WI and MTCI based prediction for current suitable habitat of *PIDE*. Green area means suitable habitat. WI: 60.0 ~ 95.0, MTCI: -104.2 ~ -47.5 (Choi *et al.*, 2010)

c: Annual average temperature regression model prediction for current suitable habitat of *PIDE*. (Yoon, 2003)

d: Annual precipitation regression model prediction for current suitable habitat of *PIDE*. (Yoon, 2003)

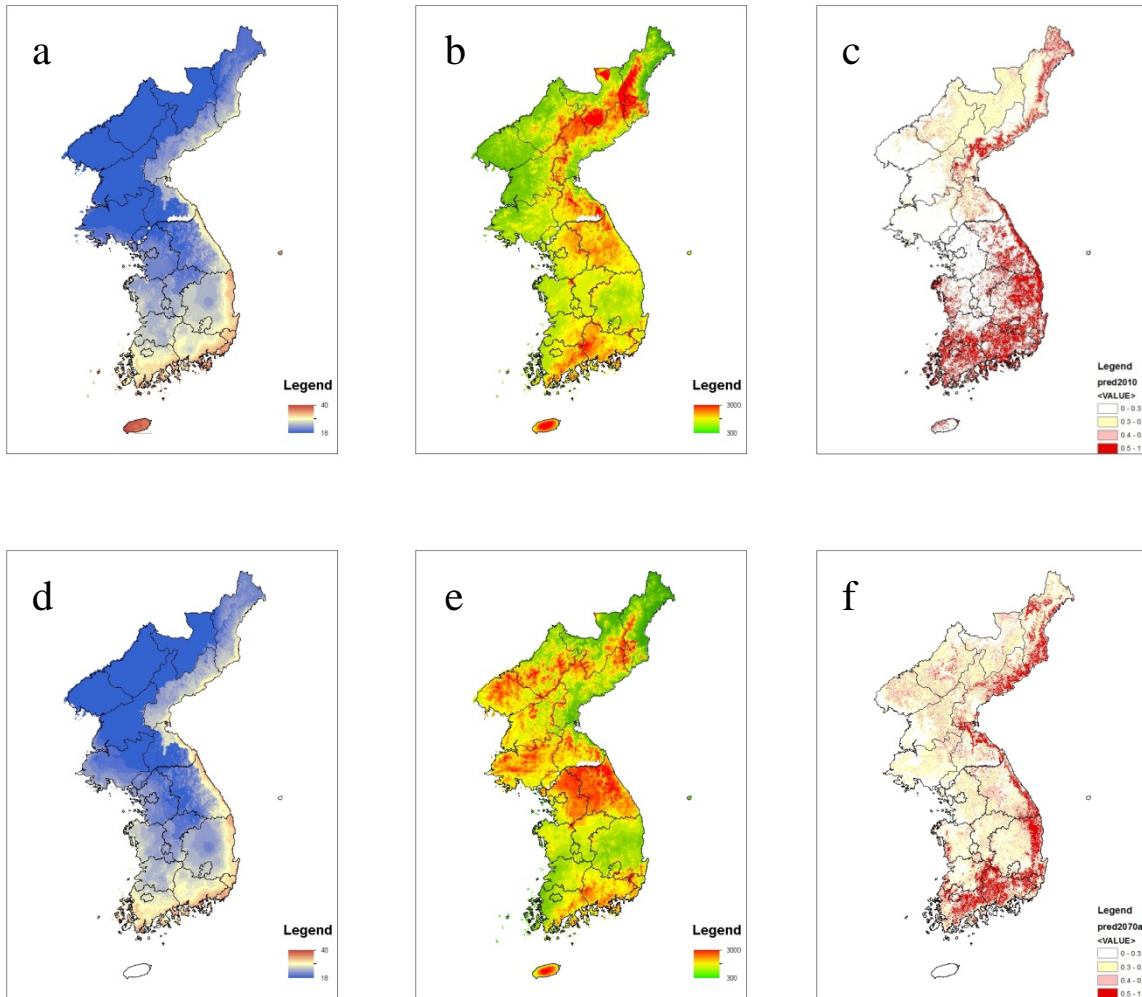


Figure 13. Pattern of Temperature, Precipitation, and Distribution Change of *Pinus densiflora*.

a, d: Summer/winter temperature difference map for 2010 (a) and 2070 of A2a scenario (d)

b, e: Growing season precipitation map for 2010 (b) and 2070 of A2a scenario (e)

c, f: *PIDE* suitable habitat prediction map for 2010 (e) and 2070 of A2a scenario (f)

Suggestions for Future Conservation Policy for Forests

The results discussed in the previous sections suggest that forest pest management could be more important in trying to preserve the current *PIDE* dominated Korean forests than was previously assumed. *PIDE* is among the species that have the greatest demand for conservation efforts due to its important role in the regional ecosystems and Korean culture. *PIDE* is the species that occupies the biggest portion of current mature forest canopy. Moreover, it is among the most important symbols of Korean history and people and even appears appear in the lyrics of the Korean national anthem. Thus there is a big concern about the future of *PIDE* stands especially with the gloomy predictions that suggest the disappearance of that species within 100 years due to climate change and the spreading of pest. In this study, the Random Forest modeling suggested that *PIDE* can retain their habitat in the Korean peninsula even with a severe climate change scenario. The predictions for suitable habitat of *PIDE* for 1980 and 2010 showed the suitable habitats decreased dramatically in the western plains and the eastern inland regions of the Korean peninsula. Simple projections from these results could prompt under-estimates of the ability of *PIDE* to adapt future warmer climates. In fact the predicted decreases of suitable habitats mostly occurred in the densely populated regions stimulating the increase in concerns and under-estimates of the future of the *PIDE* habitat.

However, pests such as the pine wilt disease (PWD) nematode and the pine gall midge still remain as big threats as they have already destroyed large areas of *PIDE* forests in Japan over the last 100 years (Sakamoto *et al.*, 2003; Toyohara and Fujihara, 1998). These pests could spread even more lively under warming climate (Dwinell, 1997; Jikumaru and Togashi, 2000). For example, PWD could quickly spread in the areas that have average August temperatures greater than 20 °C and the vector beetle of the disease could live in the areas that have average

temperatures greater than 21.3°C during July and -10°C during January (Robinet *et al.*, 2009).

The areas that coincide with these conditions already overlap well with the current *PIDE* distribution areas and will continue to increase with a warmer climate change (Fig. 14).

Therefore, pest control centered on conservation efforts for *PIDE* stands would be effective to incorporate in future climate change scenarios. Pest control management is helpful for other *Pinus* species, such as *Pinus koraiensis*, since these pests can also attack them as well. There was a report of PWD infection of *PIKO* at Kwangju, Kyongkido, South Korea (2007).

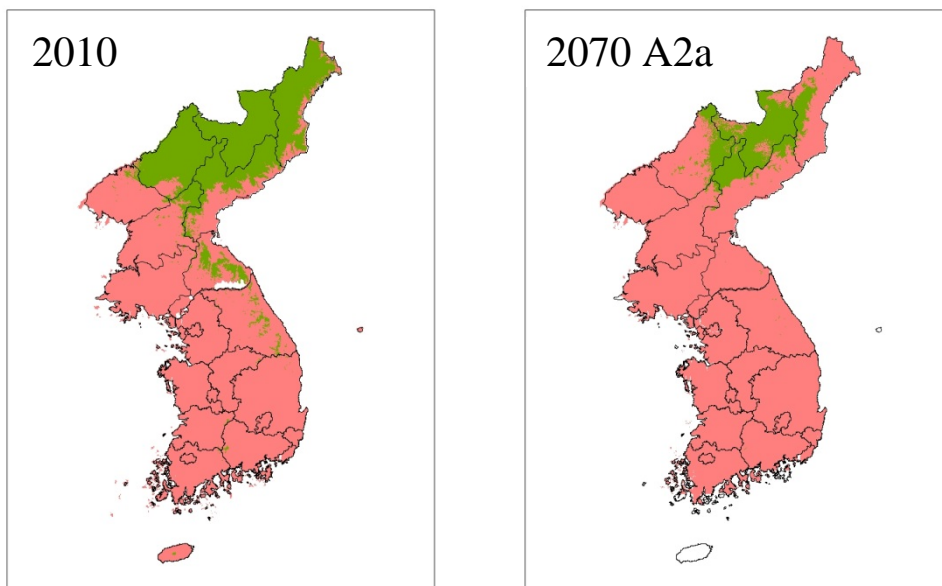


Figure 14. Pine Wilt Disease Susceptible Area.

Red area indicates PWD susceptible area from Robinet *et al.* 2009 (Average temperature over 21.3°C in July and -10°C in January). Almost everywhere of South Korea and majority of North Korea region will be susceptible for the tree disease until 2070.

CONCLUSION

This study modeled the suitable habitat of a major tree species, *Pinus densiflora* (*PIDE*), in the Korean peninsula using the Random Forest model and environmental data obtained from HD-DCMs at two different resolutions. The following is a summary of the major results found in this study.

1. Current distribution of *PIDE* in the Korean peninsula was successfully simulated by using the Random Forest model and 17 different factors synthetically. Previous forest modeling studies have generally used the statistical regression method and mostly only one factor at a time. In fact when statistical methods were used in this study with multiple factors, simulation of the actual distribution patterns were improved over just using one factor at a time.

2. Suitable habitat of *PIDE* predicted for the future using IPCC's climate change scenarios. Although the suitable *PIDE* habitat areas of the species decreased by almost 50 % in a scenario when the model's was set at 0.5, the *PIDE* habitats along the southern and eastern coasts were still mostly retained. Moreover, its extent even distinguishably expanded northward when using the most extreme A2a scenario and a more realistic threshold value of 0.46 was set. These results show the importance of being able to predict the impact of temperature and precipitation together on the habitat distribution patterns, and thus reinforce the utility of using multiple environmental factors in model simulations. This type of modeling exercise is especially necessary to help manage our environment and establish policies to prepare for and/or mitigate predicted climate change effects

3. This study predicted that *PIDE* will maintain its environmental niche in Korean peninsula even when the most extreme climate change scenarios are given, while other current competitor tree species such as *Quercus spp.* may not. The predicted increase in precipitation in Korean peninsula during the growing season may help *PIDE* to survive and even reproduce in the extreme warming influence. However, with the predicted climate change occurring, conditions will improve for not only *PIDE* but also for tree diseases. Therefore, development of appropriate pest management programs will be very important to help conserve the future of Korea's *PIDE* forest and at the same time support one of their cultural icons.

ACKNOWLEDGEMENTS

I appreciate my Master's committee members, Soo-Hyung Kim, Daniel Vogt, and Joshua Lawler for their great guidance and patience. I thank friends in the Kim lab and Vogt lab for their help as well. Rania Chung and Stephan Gmur gave a lot of help in using the GIS program. Lloyd Nackely helped editing the manuscript. This work was supported in part by National Center for Agro-Meteorology (NCAM), and by the Korea Meteorological Administration Research and Development Program, Republic of Korea.

REFERENCE

- Boo K-S. The agriculture of North Korea. (2001).
- Breiman L. Classification and regression trees. (1984) Belmont, Calif.: Wadsworth International Group.
- Breiman L. Random forests. *Mach Learn* (2001) 45:5-32.
- Cho YCS, H.C., Kim, S.S. ; Lee, C.S. . Dynamics and Conservation of the Gwangneung National Forest in Central Korea: A National Model for Forest Restoration. *Journal of Plant Biology* (2007) 50:615-625.
- Choi J, Chung U, Yun Ji. Urban-Effect Correction to Improve Accuracy of Spatially Interpolated Temperature Estimates in Korea. *J Appl Meteorol* (2003) 42:1711-1719.
- Choi S, Lee W-K, Son Y, Yoo S, Lim J-H. Changes in the distribution of South Korean forest vegetation simulated using thermal gradient indices. *SCIENCE CHINA Life Sciences* (2010) 53:784-797.
- Choi. Kee R. The Post-glacial Vegetation History of the Lowland in Korean Peninsula. 韓半島 後氷期の低地帯 植生史 (1998) 21:169-174.
- Choung HL, Hong SK. Distribution patterns, floristic differentiation and succession of *Pinus densiflora* forest in South Korea: A perspective at nation-wide scale. *Phytocoenologia* (2006) 36:213-229.
- Chung U, et al. Extrapolation of Daily Maximum Temperature in a Mountainous Terrain. *Asia-Pac J Atmos Sci* (2009) 45:473-482.
- Chung U, et al. Minimum temperature mapping over complex terrain by estimating cold air accumulation potential. *Agr Forest Meteorol* (2006) 137:15-24.
- Collins WD, et al. The Community Climate System Model Version 3 (CCSM3). *Journal of Climate* (2006) 19:2122-2143.
- Critchfield WB, Little EL. Geographic distribution of the pines of the world. (1966) Washington, D.C.: U.S. Dept. of Agriculture, Forest Service.
- Daly C, Neilson RP, Phillips DL. A Statistical Topographic Model for Mapping Climatological Precipitation over Mountainous Terrain. *J Appl Meteorol* (1994) 33:140-158.
- De'ath G, Fabricius KE. CLASSIFICATION AND REGRESSION TREES: A POWERFUL YET SIMPLE TECHNIQUE FOR ECOLOGICAL DATA ANALYSIS. *Ecology* (2000) 81:3178-3192.
- Dolezal J, et al. Tree growth and competition in a post-logging & Quercus mongolica forest on Mt. Sobaek, South Korea. *Ecological Research* (2009) 24:281-290.
- Dwinell LD. The pinewood nematode: Regulation and mitigation. *Annual Review of Phytopathology* (1997) 35:153-166.
- ESRI. ArcGIS Desktop: Release 10. (2011).
- Fleiss JL, Cohen J, Everitt BS. Large Sample Standard Errors of Kappa and Weighted Kappa. *Psychol Bull* (1969) 72:323-&.
- Gordon HB, Research CDoA. The CSIRO Mk3 Climate System Model (2002) (Type of Medium).
- Hayashi Y. The natural distribution of important trees, indigenous to Japan. *Conifers Report 2: Bull Gov Forest Exp St, Tokyo* (1952): 241-242.
- IPCC. Special Report on Emissions Scenarios : a special report of Working Group III of the Intergovernmental Panel on Climate Change. (2000).
- IPCC. Climate change. 2007 : the physical science basis : contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. (2007) New York, NY ; Cambridge: Cambridge University Press.
- Jeon SP, Yongha. Jeong, Huicheol. Harasawa, Hideo. Takahashi, Kiyoshi. Impact Assessment of Ecosystem Influenced by Changing Global Climate and its National Management Practices II :

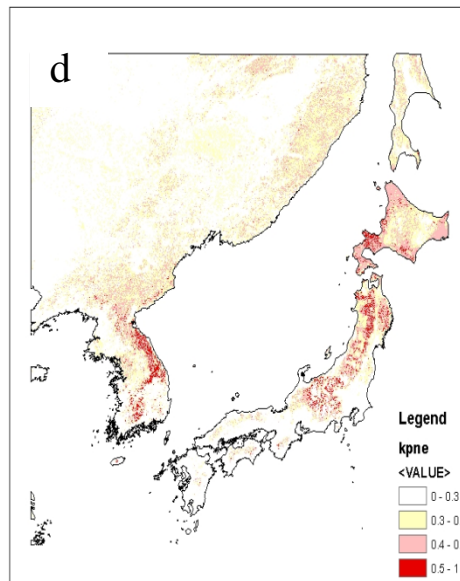
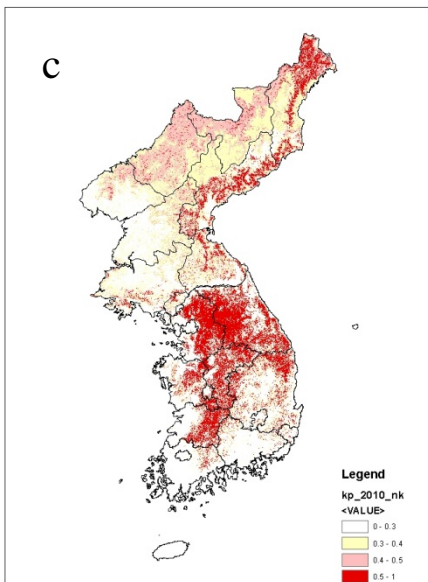
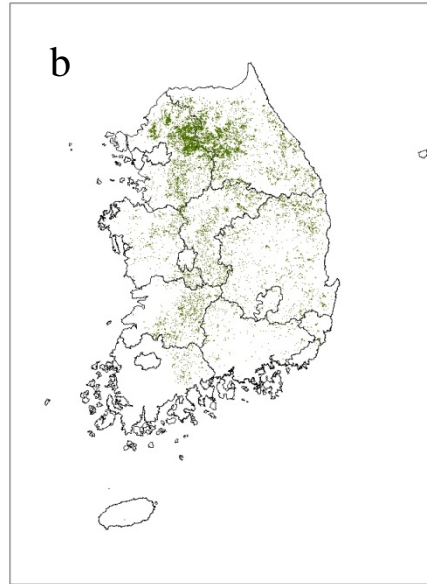
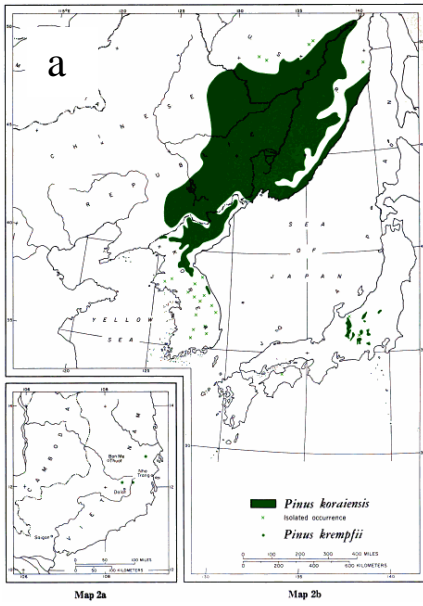
- Focusing on the bioclimatic model development of the forest. Korea Environment Institute (2001) RE-08.
- Jeon SW, Park YH, Jung HC, Lee DK, Song K, Yang WT. Impact Assessment of Ecosystem Influenced by Changing Global Climate and its National Management Practices III. (2002) Seoul: Korea Environment Institute
- Jikumaru S, Togashi K. Temperature effects on the transmission of *Bursaphelenchus xylophilus* (Nemata : Aphelenchoididae) by *Monochamus alternatus* (Coleoptera : Cerambycidae). *Journal of Nematology* (2000) 32:110-116.
- Johns TC, et al. The New Hadley Centre Climate Model (HadGEM1): Evaluation of Coupled Simulations. *Journal of Climate* (2006) 19:1327-1353.
- Kato J, Hayashi I. Quantitative analysis of a stand of *Pinus densiflora* undergoing succession to *Quercus mongolica* ssp. *crispula*: 1. A 31-year record of growth and population dynamics of the canopy trees. *Ecological Research* (2006) 21:503-509.
- Kato J. The determination and prediction of pine to oak forest succession in Sugadaira, Central Japan. *Korean J. Ecology* (2003) 26:155-163.
- Kim CH. (A)study on the structure of forest vegetation and the secondary succession in Togyusan national park, Korea (1992) Irie: WonKwang University.
- Kim HE, Kwon KC, Jung TS. Disturbance and Regeneration Process of the *Pinus densiflora* Forest in Mt. Worak Forest *Bioenergy* (2000) 19:79-85.
- Kim J-H. The Analysis of Forest Successional Trend by Species Replacement Model in the Natural Forest. *Forest Bioenergy* (2003) 22:1-10.
- Kim J-U, Kil B-S. Estimation for Changes of Net Primary Productivity and Potential Natural Vegetation in the Korean Peninsula by the Global Warming. *The Korean journal of ecology* (1996) 19:1-7.
- Kim JW. *Pinus densiflora* of Kora in the view of Pine Wilt Disease nematode and forest fire in East coastal region, What is the problem. *Korean J. Ecology* (2005) 28:113-120.
- Kim KH, Zsuffa L. Reforestation of South-Korea - the History and Analysis of a Unique Case in Forest Tree Improvement and Forestry. *Forest Chron* (1994) 70:58-64.
- Kim T-M, Lee W-K, Jung S-E, Kwak H-B. Analysis of relationship between spatial distribution of *Pinus densiflora* and topographical factors. *Journal of Geographic Information System Association of Korea* (2008) 16:359-372.
- Kong W-S. The alpine and subalpine geoecology of the Korean Peninsula. *Korean J. Ecology* (1998) 21:383-387.
- Kong W-S. Vegetation Disturbance of Korea during the Pre-Chosun Dynasty Period. *The Korean journal of quaternary research* (2000a) 14:33-48.
- Kong W-S. Vegetational history of the Korean Peninsula. *Global Ecology and Biogeography* (2000b) 9:391-402.
- Kong W-S. Biogeography of native Korean Pinaceae. *J. Korean Geogr. Soc.* (2006) 41:73-93.
- Korea Forest Service. Statistical yearbook of Forestry. (2011).
- Kwon Y-A, Kwon W-T, Boo K-O, Choi Y. Future Projections on Subtropical Climate Regions over South Korea Using SRES A1B Data. *Journal of the Korean Geographical Society* (2007) 42:355-367.
- Ledig FT, Rehfeldt GE, Saenz-Romero C, Flores-Lopez C. Projections of Suitable Habitat for Rare Species under Global Warming Scenarios. *Am J Bot* (2010) 97:970-987.
- Lee CS, Kim JH, Yi H, You YH. Seedling establishment and regeneration of Korean red pine (*Pinus densiflora* S. et Z.) forests in Korea in relation to soil moisture. *Forest ecology and management* (2004) 199:423-432.

- Lee CS, Lee AN. Ecological Importance of Water Budget and Synergistic Effects of Water Stress of Plants due to Air Pollution and Soil Acidification in Korea. *The Korean journal of ecology* (2003) 26:143-150.
- Lee HJ, You YH. Ecological Niche Breadth of *Q. mongolica* and Overlap with *Q. acutissima* and *Q. variabilis* along with Three Environment Gradients. *Korean journal of environmental biology* (2009) 27:191-197.
- Lee JS, Lee WK, Yoon J, Song JC. Distribution Pattern of *Pinus densiflora* and *Quercus* Spp. Stand in Korea Using Spatial Statistics and GIS. *Korean J. Forestry* (2006) 95:663-671.
- Lee SH. Projection of Climate Change Effects on the Potential Distribution of Endemic Genera in Korea (2011) Daejeon: Chungnam National University.
- Leibold MA. The Niche Concept Revisited: Mechanistic Models and Community Context. *Ecology* (1995) 76:1371-1382.
- Liaw A, Wiener M. Classification and regression by Random Forest. *R news* 2 (2002):18-22.
- Loehle C, LeBlanc D. Model-based assessments of climate change effects on forests: A critical review. *Ecological modelling*. (1996) 90:1.
- Min S-K, Park EH, Kwon W-T. Future Projections of East Asian Climate Change from Multi-AOGCM Ensembles of IPCC SRES Scenario Simulations. *Journal of the Meteorological Society of Japan* (2004) 82:1187-1211.
- Morin X, Augspurger C, Chuine I. PROCESS-BASED MODELING OF SPECIES' DISTRIBUTIONS: WHAT LIMITS TEMPERATE TREE SPECIES' RANGE BOUNDARIES? *Ecology* (2007) 88.
- Morin X, Thuiller W. Comparing niche- and process-based models to reduce prediction uncertainty in species range shifts under climate change. *Ecology* (2009) 90:1301-1313.
- Oh SN. Global warming and climate change in Korean peninsula (2008): *The Korean Society For Bio-Environment Control*. 15-29.
- Park J-c, Yang K-c, Jang D-h. The Movement of Evergreen Broad-Leaved Forest Zone in the Warm Temperate Region Due to Climate Change in South Korea. *Journal of climate research* (2010) 5:29-41.
- Park W. Monitoring Cambial Activity and Tree Growth, and Simulation Modeling for Predicting Growth Decline due to Global Warming and Air Pollution. KISTI (2000).
- R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing (Austria, Eds.) (2010).
- Régnière J, Bergeron V, Cooke B, Canadian Forest Service Q, Laurentian Forestry C. BioSIM a computer-based decision support tool for seasonal planning of pest management activities : user's manual. (1995) Sainte-Foy, Quebec: Canadian Forest Service - Quebec.
- Robinet C, et al. Role of Human-Mediated Dispersal in the Spread of the Pinewood Nematode in China. *Plos One* (2009) 4.
- Sakamoto K, Miki N, Tsuzuki T, Nishimoto T, Yoshikawa K. Comparison of stand dynamics after dieback caused by pine wilt disease among pine forests with different management regimes in western Japan. *Journal of Forest Research* (2003) 8:303-309.
- Scinocca JF, McFarlane NA, Lazare M, Li J, Plummer D. The CCCma third generation AGCM and its extension into the middle atmosphere. *Atmos. Chem. Phys. Discuss. Atmospheric Chemistry and Physics Discussions* (2008) 8:7883-7930.
- Shin S-C, et al. Estimation of High Resolution Gridded Precipitation Using GIS and PRISM. *Atmosphere* (2008) 18:71-81.
- Song H-K, K-K Jang. Study on the DBH analysis and forest succession of *Pinus densiflora* and *Quercus mongolica* forests. *Jour. Korean For. Soc.* (1997) 86:223-232.

- Song MS. Analysis of distribution and association structure on the Sawtooth Oak (*Quercus acutissima*) Forest in Korea (2008): Changwon University.
- Swets JA. Measuring the Accuracy of Diagnostic Systems. *Science* (1988) 240:1285-1293.
- Thuiller W, Brotons L, Araújo MB, Lavorel S. Effects of restricting environmental range of data to project current and future species distributions. *Ecography* (2004) 27:165-172.
- Toyohara G, Fujihara M. Succession of Secondary Forests Affected by Pine Wilt Disease in Western Japan Followed on Vegetation Maps. *Applied Vegetation Science* (1998) 1:259-266.
- UNEP. DPR Korea: State of environment 2003. United Nations Environment Programme (2003).
- Yoon J. Characteristics and change prediction of spatial distribution of *Pinus densiflora* stands in Korea : Emphasis on impacts of topography, climate, and soil factors (2003) Seoul: Korea University.
- Yun J-I. A Simple Method Using a Topography Correction Coefficient for Estimating Daily Distribution of Solar Irradiance in Complex Terrain. *Korean Journal of Agricultural and Forest Meteorology* (2009) 11:13-18.
- Yun JI. Agroclimatic Maps Augmented by a GIS Technology. *Korean j. Agricultural and Forest Meteorology* (2010) 12:63-73.
- Yun JI, et al. Minimum temperature mapping over complex terrain by estimating cold air accumulation potential. *Agr Forest Meteorol* (2006) 137:15-24.

APPENDIX A

Failure in selecting appropriate source species distribution data – *Pinus koraiensis* case.



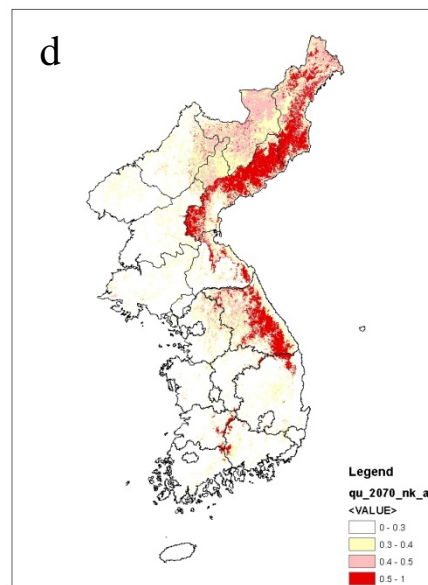
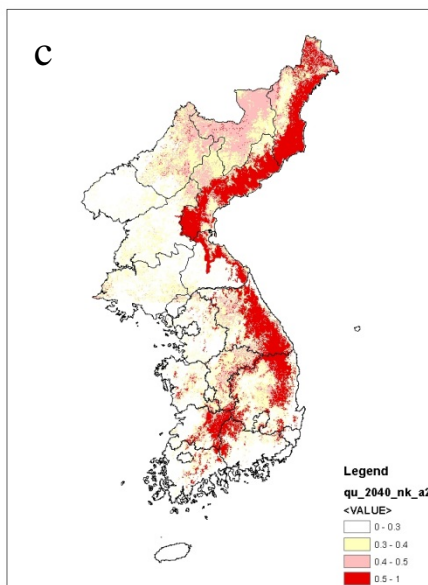
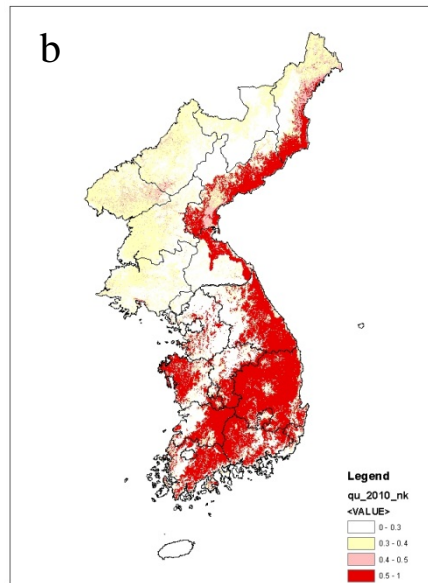
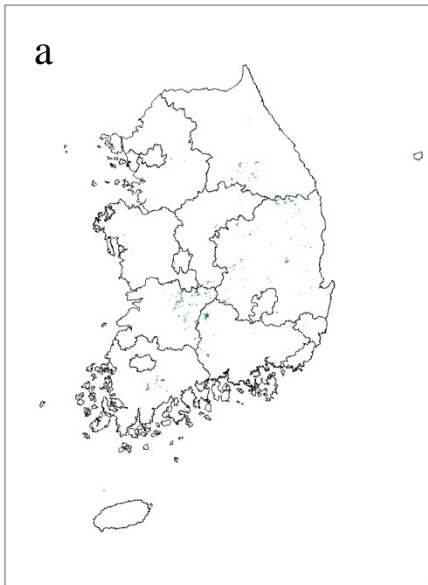
- a. Distribution atlas for *Pinus koraiensis* from 1960 (Critchfield and Little, 1966)
- b. Current *Pinus koraiensis* stands from digital forest map of South Korea, 2006
- c. Random Forest model prediction at 270 m X 270 m resolution and Korean peninsula extent
- d. Random Forest model prediction at 1 km X 1km resolution and East Asia extent

An example of why it is important to determine the initial input source of distribution data can be observed when using the Random Forest model to predict the suitable habitat patterns for another important pine species found in the Korean peninsula, i.e., *Pinus koraiensis*. The distribution and environmental were from digital forest map and HD-DCMs for only South Korea region (b).

However, the main distribution regions of *Pinus koraiensis* are in the mountainous highlands in North Korea and Manchurian region of northeast China. The model for *Pinus koraiensis* failed to accurately predict its known main distribution regions outside of South Korea whereas the *PIDE* model did predict the habitats of objective species found in Japan and China's Shandong peninsula very well. The differences in accurately predicting current suitable habitats for *PIDE* versus *Pinus koraiensis* by the Random Forest model, especially in regions outside of South Korea, was attributed to the actual distribution coverage of those two species initially appropriated in the source distribution data.

APPENDIX B

Random Forest modeling for future of *Quercus spp* in the Korean peninsula.



a: *Quercus* presence map in digital forest map. This distribution map underestimated actual distribution of the tree family since the late succession trees occupy lower canopy which is hard to detect with remote sensing.

b: Predicted *Quercus spp.* suitable habitat distribution by current climate condition. Current predicted distribution pattern of *Quercuss spp.* is overlapped well with *PIDE* habitat distribution.

c: Predicted *Quercus spp.* suitable habitat distribution in 2040 of A2a climate change scenario. Suitable habitat along south coastline is disappearing quickly

d: Predicted *Quercus spp.* suitable habitat distribution in 2070 of A2a climate change scenario. Suitable habitat in South Korea region almost disappeared except narrow high mountain region.

Random Forest model for *Quercus spp.* was made from the same data source with the *PIDE* model. Although AUC of the model was much higher than the *PIDE* model (0.96), the model has limitation from the source species distribution data which doesn't cover all of the possible environmental distribution range of the species. However, the model prediction still has some points to look at. The model predicted a different change of habitat distribution pattern of *Quercuss spp.* with *PIDE* in future climate warming scenarios despite of its similarity of current suitable habitat distribution with *PIDE*. According to A2a scenario, the suitable habitat of *Quercus spp.* will move to the northern regions quickly and almost disappear in South Korea. Although the source distribution data didn't cover entire distribution range of the tree family, its decrease is plausible if we focus on only in Korean peninsula region since the major oak species in Korean peninsula, *Quercus mongolica* is cold tolerant species that has broad northern distribution until Mongolia and Siberia region. This result means a future warmer climate can

give a more physiological advantage to *PIDE* and increase the competitiveness of *PIDE* against any future successional pressures.