

Course e-Notebook

Course e-Notebook

Tongli Wang

Contents

Introduction	1
Tongli Wang	

Module I. Introduction to Climate Change and Its Impacts on Forest Ecosystems

Topic 1.1. Introduction to Global Climate Change	5
Tongli Wang	
<i>Video recordings:</i>	5
<i>Contents</i>	7
<i>1. Weather and climate</i>	8
<i>2. Extreme events</i>	10
<i>3. Climate change and climate variability</i>	11
<i>4. Observed climate change</i>	17
<i>5. Causes of climate change</i>	22
<i>6. Future projections of climate change</i>	26
<i>Summary</i>	31

Topic 1.2. Impacts of Climate Change on Forest Ecosystems	33
Tongli Wang	
<i>Contents</i>	33
<i>2. Impacts of Climate Change on Forest growth and productivity</i>	39
<i>3. Impacts of Climate Change on the vulnerability of forest ecosystems to pests and diseases</i>	43
<i>4. Impacts of Climate Change on forest fires</i>	45
<i>5. Impacts of Climate Change on biodiversity and sustainability of forest ecosystems</i>	46
<i>6. Impacts of Climate Change on Forest Ecosystems' services</i>	49
<i>7. Mitigation of climate change by forest ecosystems</i>	53
<i>Summary</i>	57

Module II. Introduction to Climatic and Ecological Models

Topic 2.1. Introduction to Climate Data and Climate Models	61
Tongli Wang	
<i>Contents</i>	62
<i>1. Why climate data are important?</i>	63
<i>2. How to obtain climate data?</i>	67
<i>3. Climate variables and time intervals</i>	71
<i>4. Major sources of climate data</i>	72
<i>Summary</i>	78
Topic 2.2. Introduction to Ecological Models	81
Tongli Wang	
<i>Contents</i>	82
<i>1 Overview</i>	83
<i>2. Types of models</i>	85
<i>3. Building process-based models</i>	90
<i>4. Widely used process-based models</i>	93

Module III. Scale-free Climate Models and Their Applications

Topic 3.1. Introduction to ClimateNA and ClimateAP	107
ClimateNA and ClimateAP	
Tongli Wang	107
<i>Contents</i>	108
<i>2. Data sources</i>	111
<i>3. Dynamic local downscaling</i>	113
<i>4. Development of additional climate variables</i>	117
<i>5. Main features</i>	119
<i>6. Web versions</i>	124
<i>7. Spatial data</i>	127
<i>8. ClimateNAr package</i>	128

Topic 3.2. Using ClimateNA/AP to Generate Point and Spatial Climate Data Tongli Wang	131
	131
<i>Contents</i>	131
1. <i>Download and installation</i>	132
2. <i>Using the interactive mode</i>	134
3. <i>Using multi-location mode</i>	139
4. <i>Multiple GCMs and scenarios processing</i>	142
5. <i>Generating spatial climate data</i>	144
6. <i>Using time-series functions</i>	150
7. <i>Using the web version</i>	152
8. <i>Using pre-generated spatial data</i>	157
9. <i>Using the R version of ClimateNA</i>	159

Module IV. Introduction to Niche-based
Models

Topic 4.1. Introduction to Niche-based ecological Models Tongli Wang	165
	165
<i>Contents</i>	166
1. <i>Adaptation of plant species to climate</i>	167
2. <i>Ecological niche</i>	169
3. <i>Impacts of climate change on ecological niche</i>	174

Topic 4.2. Introduction to Genecology functions	181
Tongli Wang	
	181
<i>Contents</i>	182
1. <i>Overview</i>	183
2. <i>Definition of genecology</i>	184
3. <i>Transfer functions</i>	187
4. <i>General transfer functions</i>	191
5. <i>Genecology functions</i>	196
Topic 4.3. Introduction to Climate Response Functions	203
Tongli Wang	
	203
<i>Contents</i>	204
1. <i>Overview</i>	205
2. <i>Definition</i>	206
3. <i>Response functions with a single climate variable</i>	207
4. <i>Response functions with multiple climatic variables</i>	210
5. <i>Multiple response functions</i>	213
6. <i>Universal response functions</i>	216

Module V. Model Applications

Topic 5.1. Model Interpretation and Applications	225
Tongli Wang	
	225
<i>Contents</i>	225
<i>1. Model interpretation</i>	225
Topic 5.2. Case Studies with Forest Tree Species	239
Tongli Wang	
	239
<i>Contents</i>	239
	239
<i>1 Douglas-fir</i>	239
<i>2. Chinese fir</i>	243
<i>3. Bluegum</i>	247
Topic 5.3. Case Studies with Forest ecosystems	251
Tongli Wang	
	251
<i>Contents</i>	251
<i>1. Biogeoclimatic ecosystems in British Columbia</i>	252
<i>2. Forest ecosystem in China</i>	259

Module VI. Building Ecological Models in R

Topic 6.1 Building niche-based models - 1/3	267
Tongli Wang	
<i>Contents</i>	267
1. <i>Overview</i>	268
2. <i>Some basics of R</i>	268
3. <i>Species occurrence data preparation</i>	274
4. <i>Climate variables generation</i>	284
4.1 <i>Generating climate variables for specific sample locations</i>	284
5. <i>To obtain species occurrence data</i>	289
Topic 6.2 Building niche-based models - 2/3	295
Tongli Wang	
<i>Contents</i>	295
5. <i>Random Forest Model</i>	295
6. <i>MaxEnt model</i>	314

Topic 6.3 Building transfer functions and response functions - 3/3	325
Tongli Wang	
<i>Contents</i>	325
<i>1. Overview</i>	325
<i>2. Dataset preparation</i>	326
<i>3. Transfer functions</i>	338
<i>4. Genealogy functions</i>	353
<i>5. Response functions</i>	355
<i>6. Universal response functions</i>	358
Appendix	365

Introduction

Tongli Wang

2 Tongli Wang

Module I. Introduction to Climate Change and Its Impacts on Forest Ecosystems

Topic 1.1. Introduction to Global Climate Change

Tongli Wang

Video recordings:

1. Introduction:



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://pressbooks.bccampus.ca/fode010notebook/?p=5#video-5-1>

2. Weather and climate:



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://pressbooks.bccampus.ca/fode010notebook/?p=5#video-5-2>

3. Climate variability and climate change:



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://pressbooks.bccampus.ca/fode010notebook/?p=5#video-5-3>

4. Observed climate change:



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://pressbooks.bccampus.ca/fode010notebook/?p=5#video-5-4>

5. Causes of climate change:



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://pressbooks.bccampus.ca/fode010notebook/?p=5#video-5-5>

6. Climate change projections:



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://pressbooks.bccampus.ca/fode010notebook/?p=5#video-5-6>

7. Mitigation and adaptation:



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://pressbooks.bccampus.ca/fode010notebook/?p=5#video-5-7>

Contents

[1. Weather and climate](#)

[2. Extreme events](#)

[3. Climate change and climate variability](#)

[3.1. Climate change](#)

[3.2. Climate variability](#)

[3.3. El Niño Southern Oscillation \(ENSO\)](#)

[3.4. Possible changes in mean and variance](#)

[4. Observed climate change](#)

[4.1. Observed changes in average temperature](#)

[4.2. Observed changes in mean and variance of temperature](#)

[4.3. Observed changes in ice caps in the arctic](#)

[4.4. Observed changes in sea level](#)

[5. Causes of climate change](#)

[5.1. Greenhouse gases](#)

[5.2. The role of human activity](#)

[5.3. Solar irradiance](#)

[6. Future projections of climate change](#)

[6.1. Projections of fossil fuel emission](#)

[6.2. Projections of global temperature](#)

[7. Mitigation and adaptation to climate change](#)

[7.1. Mitigation to climate change](#)

[7.2. Adaptation to climate change](#)

[Summary](#)

1. Weather and climate

Before we talk about climate change, we first need to be clear about the definition of climate and the difference between climate and weather.

Weather is the state of the atmosphere at a place and time as regards temperature, precipitation, humidity, cloud, wind, etc. Weather is different in different parts of regions (as illustrated in Figure 1) and the world and changes over minutes, hours, days, and weeks.



Figure 1. A map of the weather forecast in the US. Images from “DILO-Winter Solstice Dec 21 2018” by tommaync, licensed under CC BY-NC-SA 2.0

Climate, on the other hand, is about the statistics of weather over long periods of time including averages, extremes, timing, the spatial distribution of weather attributes as mentioned above. The period is typically 30 years – a normal period, as shown in Figure 2. We refer to these 30-year averages of weather observations as **climate normals**. The reference normal period is 1961-1990 defined by the World Meteorological Organization (WMO).

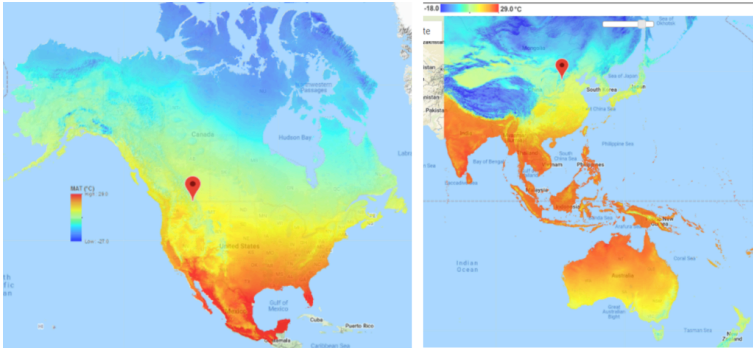


Figure 2. The patterns of mean annual temperature in North America (left) and Asia (right) for the 1961-1990 normal period. Images by Tongli Wang under CC BY-SA.

The difference between climate and weather is best summarized by the popular phrase: “the **C**limate is what you expect; the **w**eather is what you get” – Mark Twain

2. Extreme events

Weather and climate extremes such as hurricanes, tornadoes, heavy downpours, heat waves, and droughts affect all sectors of the economy and the environment, impacting people where they live and work. Extreme events also include unexpected, unusual, unpredictable, severe or unseasonal weather, or weather events at the extremes of the historical distributions as illustrated in Figure 3. Extreme weather events are considered as climate components.

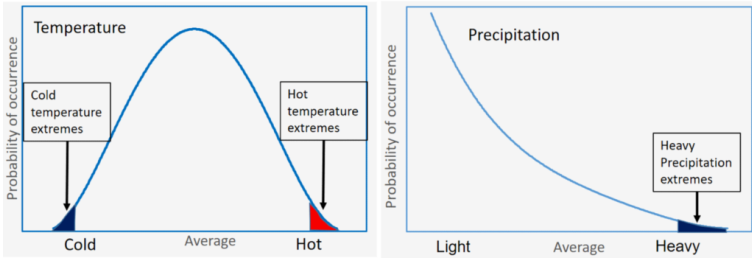


Figure 3. Illustration of temperature and precipitation extremes. Images by Tongli Wang under CC BY-SA.

3. Climate change and climate variability

3.1. Climate change

For a given location, we expect the weather to change a lot from day to day, month to month, and year to year (Figure 4, left). In contrast, we expect the climate to remain relatively constant. If the climate doesn't remain constant, as shown in Figure 4 (right), we call it **climate change**. It is important to distinguish between short-term variability and long-term climate change.

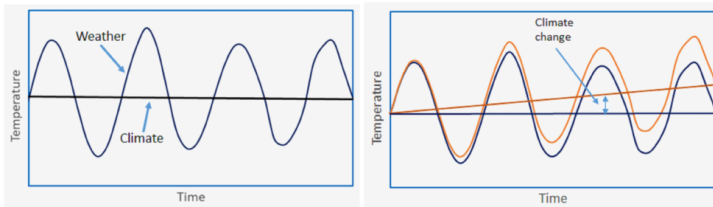


Figure 4. Illustration of weather, climate, and climate change. Images by Tongli Wang under CC BY-SA.

3.2. Climate variability

Climate variability is defined as variations in the mean state and other statistics of the climate on all temporal and spatial scales, beyond individual weather events. The term “climate variability” is often used to denote deviations of climatic statistics over a given period of time (e.g. a month, season or year) when compared to long-term statistics for the same calendar period, as shown in Figure 5. Climate variability is measured by these deviations, which are usually termed **anomalies**.

Climate variability may be due to natural internal processes within the climate system (internal variability), or to variations in natural or anthropogenic external factors (external variability). Climate variability and climate change can occur independently, but they can also occur at the same time. The main driver of climate variability in Pacific region is El Niño Southern Oscillation (ENSO).

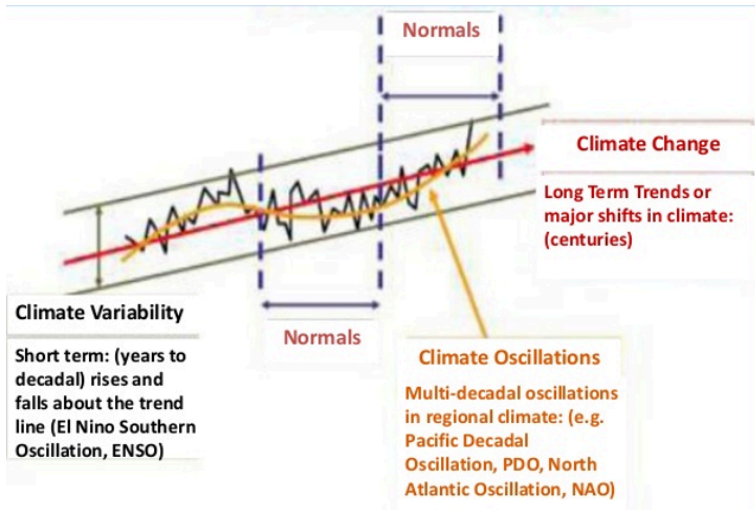


Figure 5. Illustration of climate variability and climate change.
Source: <https://www.slideshare.net/polylsgiedx/lecture-7-hk-climate-its-long-term-trend-and-variability-52280400/4> (in public domain)

3.3 El Nino Southern Oscillation (ENSO)

ENSO is an irregularly periodic variation in winds and sea surface temperatures over the tropical eastern Pacific Ocean, as shown in Figure 6. It affects the climate of much of the tropics and subtropics. The warming phase of the sea temperature is known as **El Niño** and the cooling phase as **La Niña**.

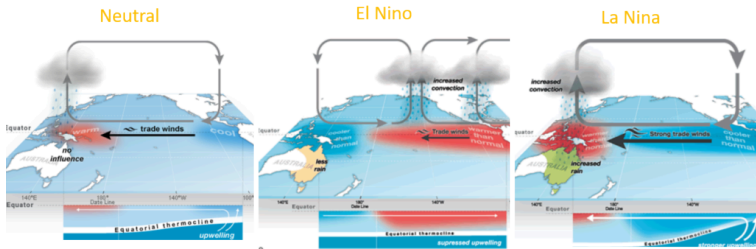


Figure 6. El Niño Southern Oscillation (ENSO). Source: <http://www.bom.gov.au/climate/enso/history/In-2010-12/three-phases-of-ENSO.shtml> (in public domain).

The El Niño and La Niña phenomenons are well explained in these two short videos:



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://pressbooks.bccampus.ca/fode010notebook/?p=5#oembed-1>

A Youtube video from [usoceangov](https://www.youtube.com/user/usoceangov)



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://pressbooks.bccampus.ca/fode010notebook/?p=5#oembed-2>

A Youtube video from [Met Office – Weather](#)

3.4 Possible changes in mean and variance

For temperature, climate change may result in an increase in mean, variance, or both mean and variance as illustrated in Figure 7. If climate change results in an increase in mean temperature only, the distribution of the probability of occurrence of temperatures would shift to the right without changing the shape of the distribution (Figure 7a).

If climate change results in an increase in variation of the temperature only, the temperature distribution would not shift, but the shape of the distribution would become flatter and wider, resulting in greater probabilities of extreme events (Figure 7b).

If climate change results in an increase in both mean and variation of the temperature, the temperature distribution would shift right and the shape of the distribution would become flatter and wider, resulting in greater probabilities of extreme events in the hot end (Figure 7c). This is likely the case of contemporary climate change.

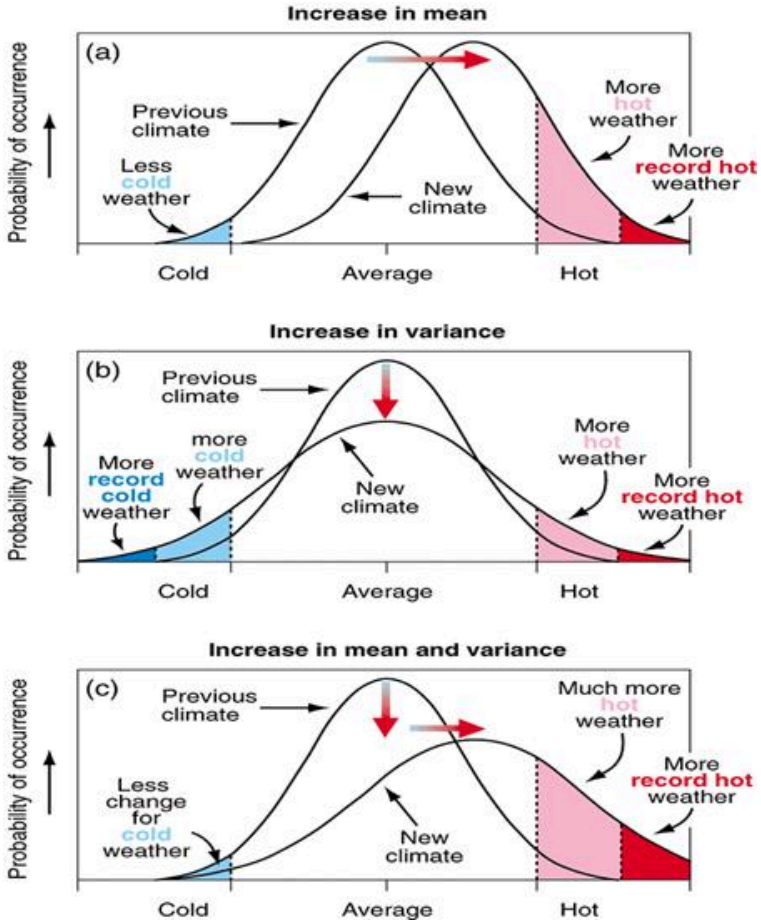


Figure 7. IPCC (2001) graph illustrating how a shift and/or widening of a probability distribution of temperatures affects the probability of extremes. Source: <https://www.climatecommunication.org/new/features/heat-waves-and-climate-change/heat-waves-the-details/>

For precipitation, the change may be more obvious for variance than for mean, resulting in less probability of occurrence in light precipitation and more in heavy precipitation (Figure 8).

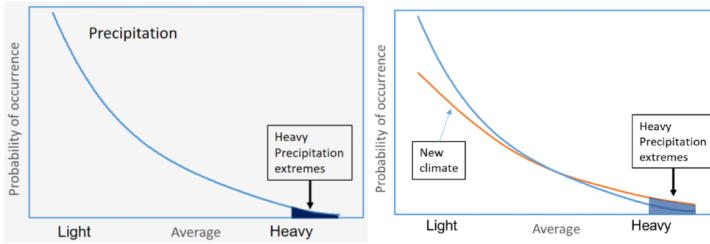


Figure 8. Illustration of changes in mean and variance affecting the probability of extremes. Images by Tongli Wang CC BY-SA.

4. Observed climate change

4.1. Observed changes in average temperature

Climate change has been observed over the last three decades. Due to annual fluctuation (Figure 9, upper), the trend of climate change is more obvious when organized in decadal averages (Figure 9, lower).

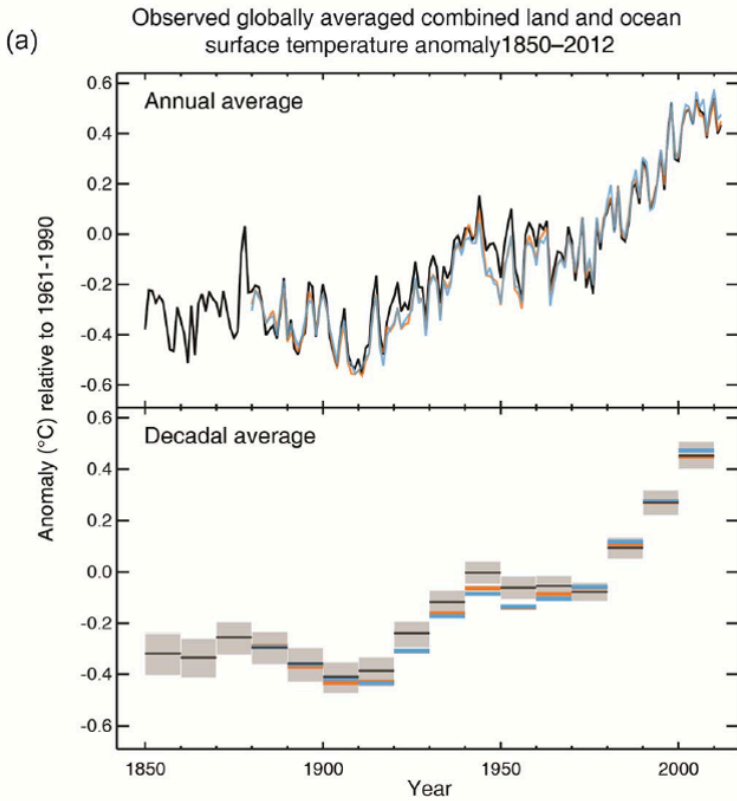


Figure 9. Observed changes in global average temperature. Open source: IPCC Fifth Report (2011).

The change in temperature is even more evident in the current decade of 2011 – 2018 (Figure 1.1.10), which is not included in Figure 1.1.9.

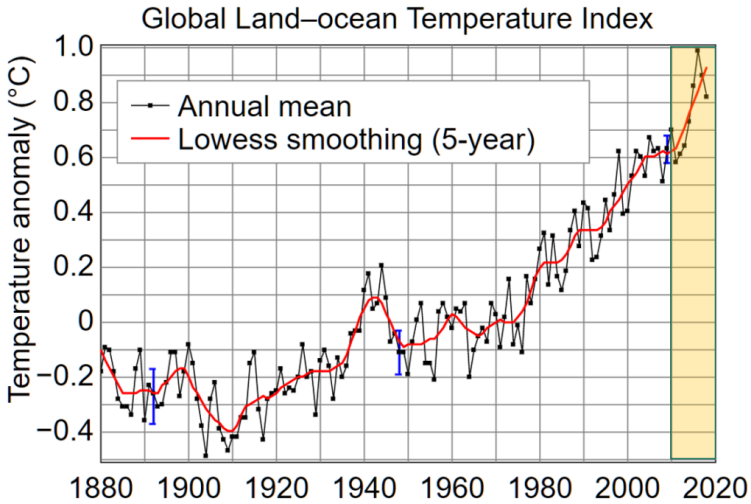


Figure 10. Observed changes in global average temperature. Open source: IPCC Report (2011).

4.2. Observed changes in mean and variance of temperature

Observations of the temperature of Northern and Southern Hemisphere land areas (James Hansen and Makiko Sato, 2016) show that distributions of temperature both shifted to the right (increased in mean) and flattened (increased in variance) over the last decades. The increases in variance in the summer months are greater than in the winter months. The increase in both mean and variance suggests an increase in the extremely hot temperatures.

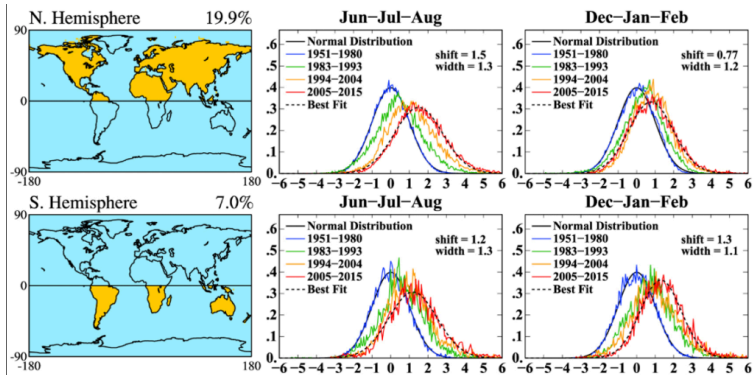


Figure 11. Frequency of occurrence of local temperature anomalies (relative to 1951–1980 mean) divided by the local standard deviation (horizontal axis) for land areas shown on the map. The area under each curve is unity. The numbers above the maps are the percent of the globe covered by the selected region. ‘Shift’ and ‘width’ refer to the dashed curve fit to 2005–2015 data and are relative to the 1951–1980 base period. Source (open access): <https://iopscience.iop.org/article/10.1088/1748-9326/11/3/034009>

4.3. Observed changes in ice caps in the arctic

The effects of climate change in the Arctic include rising air and water temperatures, loss of sea ice. Sea ice is currently in decline in area, extent, and volume. Sea ice area refers to the total area covered by ice, whereas sea ice extent is the area of the ocean with at least 15% sea ice, while the volume is the total amount of ice in the Arctic.

In 1984 the sea ice extent (Figure 12, lower) was roughly the average of the minimum from 1979 to 2000, and so was a typical year. The minimum sea ice extent in 2012 was roughly half of that average (Figure 12, upper).

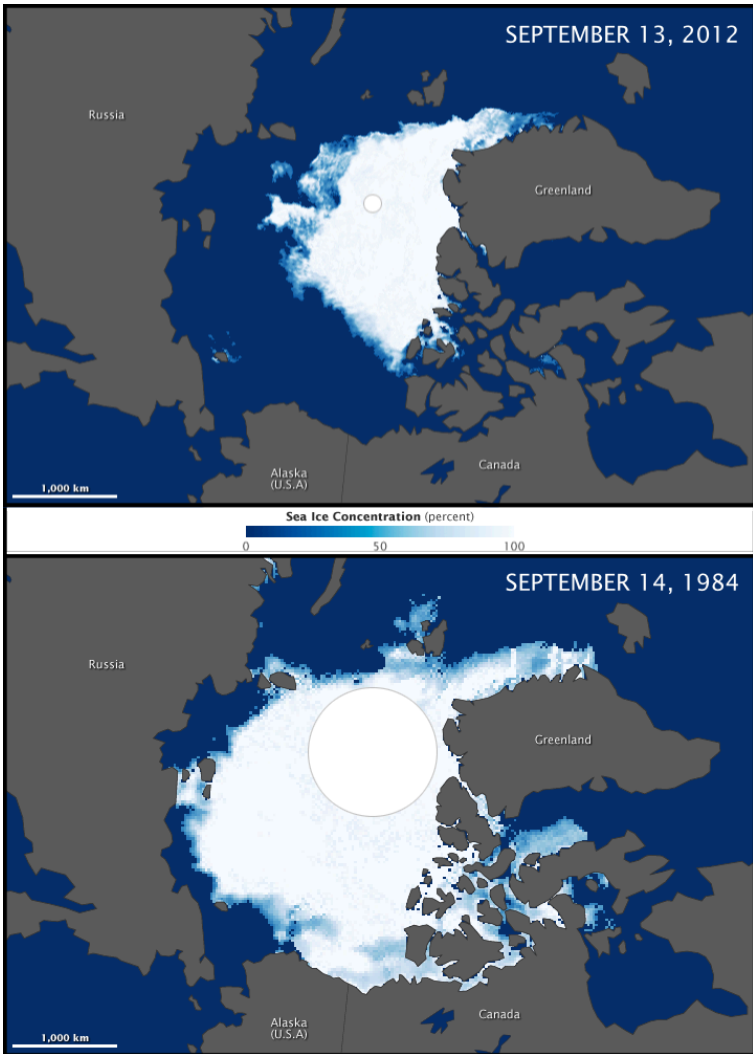


Figure 12. The maps above compare the Arctic ice minimum extents from 2012 (top) and 1984 (bottom). NASA Earth Observatory images by Jesse Allen (in the public domain).

4.4. Observed changes in sea level

Climate change is causing a continuous increase in sea level. Sea level rise is caused primarily by two factors related to global warming: the added water from melting ice sheets and glaciers and the expansion of seawater as it warms. The graph in Figure 13 shows the change in sea level since 1993 as observed by satellites.

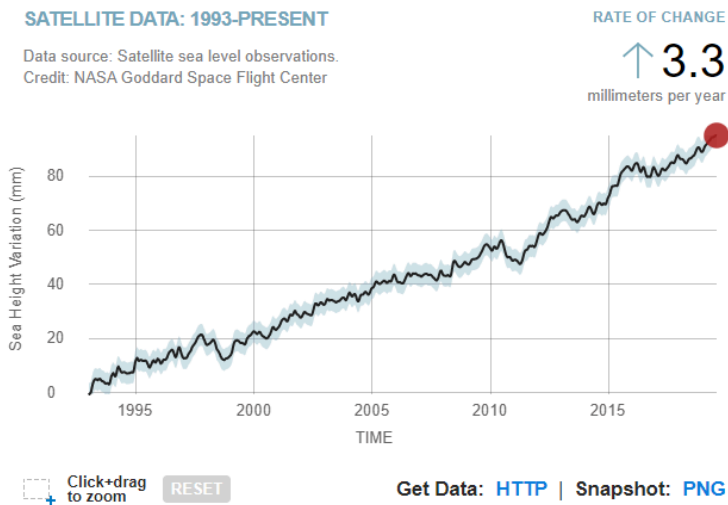


Figure 13. The change in sea level since 1993 as observed by satellites. Source (in the public domain): <https://climate.nasa.gov/vital-signs/sea-level/>

5. Causes of climate change

5.1. Greenhouse gases

Certain gases in the atmosphere block heat from escaping, causing “greenhouse effect”, thus named as greenhouse

gases. Without naturally occurring greenhouse gases, Earth's average temperature would be about -18°C instead of 15°C for the present. Long-lived gases that remain semi-permanently in the atmosphere and do not respond physically or chemically to changes in temperature are described as “forcing” of climate change (i.e., climate change forcing factors). Gases, such as water vapor, which respond physically or chemically to changes in temperature are seen as “feedbacks”. More information about climate forcing and feedback is available [Here](#).



Figure 14. A simplified animation of the greenhouse effect. Credit: NASA/JPL-Caltech (in public domain). Source: <https://climate.nasa.gov/causes/>

Gases that contribute to the greenhouse effect include:

Water vapor. The most abundant greenhouse gas, but importantly, it acts as a feedback to the climate. Water vapor increases as the Earth's atmosphere warms, but so does the possibility of clouds and precipitation, making these some of the most important feedback mechanisms to the greenhouse effect.

Carbon dioxide (CO₂). A minor but very important component of the atmosphere, carbon dioxide is released

through natural processes such as respiration and volcano eruptions and through human activities such as deforestation, land-use changes, and burning fossil fuels.

Methane. A hydrocarbon gas produced both through natural sources and human activities, including the decomposition of wastes in landfills, agriculture, and especially rice cultivation, as well as ruminant digestion and manure management associated with domestic livestock. Methane is a far more active greenhouse gas than carbon dioxide, but it is much less abundant in the atmosphere.

Nitrous oxide. A powerful greenhouse gas produced by soil cultivation practices, especially the use of commercial and organic fertilizers, fossil fuel combustion, nitric acid production, and biomass burning.

5.2. The role of human activity

Humans have increased atmospheric CO₂ concentration by more than a third since the Industrial Revolution began. This is the most important long-lived “forcing” of climate change. The carbon dioxide data, measured as the mole fraction in dry air, show a rapid rise since 1958 and hit the record high in 2019 (Figure 15).

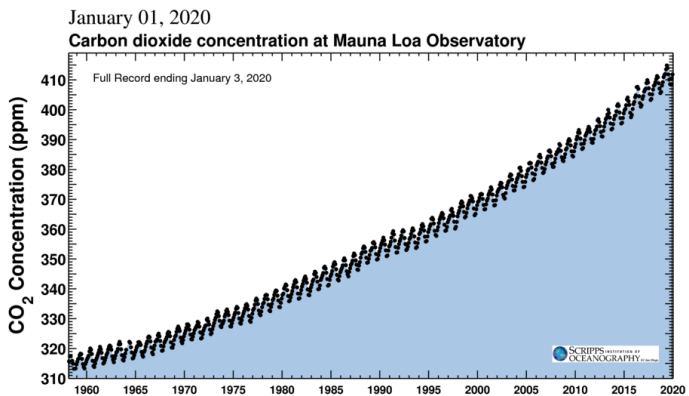


Figure 15. Changes in CO₂ level since 1958. Credit: Scripps Institution of Oceanography (in public domain). Source: <https://scripps.ucsd.edu/programs/keelingcurve/co2-graphs/>

In its Fifth Assessment Report, the Intergovernmental Panel on Climate Change (IPCC), a group of 1,300 independent scientific experts from countries all over the world under the auspices of the United Nations, concluded there's a more than 95 percent probability that human activities over the past 50 years have warmed our planet.

5.3. Solar Irradiance

As shown in Figure 16, the amount of solar irradiance did not increase over the past half-century. Thus, it is extremely unlikely that the sun has caused the observed global temperature warming trend over this period.

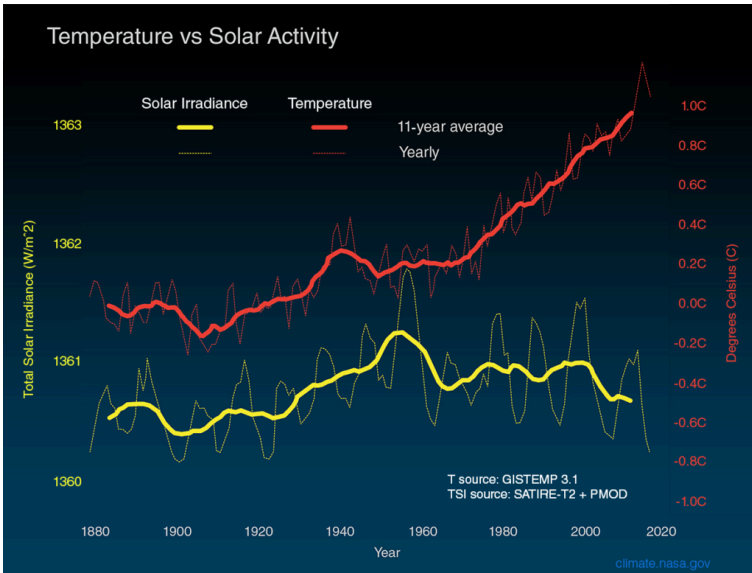


Figure 16. Global surface temperature changes (red line) and the Sun's energy that Earth receives (yellow line) in watts (units of energy) per square meter since 1880. The lighter/thinner lines show the yearly levels while the heavier/thicker lines show the 11-year average trends. Credit: NASA/JPL-Caltech (in public domain). Source: <https://climate.nasa.gov/causes/>

6. Future projections of climate change

6.1. Projections of fossil fuel emission

These are four scenarios produced by the IPCC to span a range of emissions trajectories for this century (Figure 17). These scenarios are called Representative Concentration Pathways (RCPs). A Representative Concentration Pathway (RCP) is a greenhouse gas concentration (not emissions) trajectory adopted by the IPCC for its fifth Assessment Report (AR5) in 2014. Four pathways have

been selected for climate modeling and research. They describe different climate futures, all of which are considered possible depending on the volume of greenhouse gases (GHG) emitted in the years to come. The four RCPs, namely RCP2.6, RCP4.5, RCP6, and RCP8.5, are labeled after a possible range of radiative forcing values (i.e., incoming energy – outgoing energy) in the year 2100 (2.6, 4.5, 6.0, and 8.5 W/m², respectively). RCP8.5 is often referred to as a ‘business as usual pathway, while RCP2.6 requires immediate and drastic action to start reducing global CO₂ emissions before 2020.

The IPCC for its sixth Assessment Report (AR5) was released in 2013. Shared Socioeconomic Pathways (SSPs), replacing the RCPs, are scenarios of projected socioeconomic global changes up to 2100 (Figure 17, right). The scenarios are:

SSP1: Sustainability (Taking the Green Road)

SSP2: Middle of the Road

SSP3: Regional Rivalry (A Rocky Road)

SSP4: Inequality (A Road divided)

SSP5: Fossil-fueled Development (Taking the Highway)

The scenario SSP5-8.5 in the AR6 is corresponding to RCP 8.5 in the AR5, and SSP2-4.5 is corresponding to RCP 4.5. The level of greenhouse gas emission in the AR6 is slightly higher than that in the AR5.

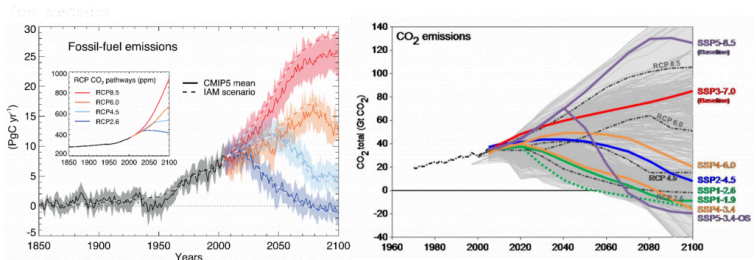


Figure 17. Projections of fossil fuel emission for four different scenarios – Left: the Representative Concentration Pathways (RCPs) from IPCC AR5 (Image Source: from IPCC Fifth Assessment Report, Technical Summary). Right: Shared Socioeconomic Pathways (SSPs) from IPCC AR6 (Image source: O’Neill et al. (2016), open source at doi:10.5194/gmd-9-3461-2016)

6.2. Projections of global temperature

For the “business as usual” emission scenario RCP 8.5, the projected global temperature increases by 2100 will be about 6°C on average (Figure 18, left). Even with the aggressive mitigation scenario RCP2.6 requiring immediate and drastic action, the projected increase will be about 1.5°C by the end of the century, which is the target of the [Paris Agreement](#). The projected global temperature increases are even higher in the IPCC AR6 (Figure 18, right)

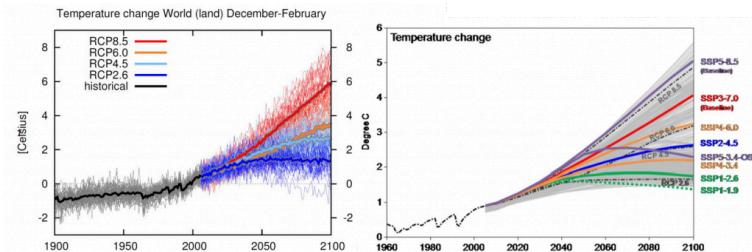


Figure 18. Projections of future temperature increase based on different scenarios. Left: based on RCPs from the IPCC Fifth Assessment Report. Right: based on SSPs from the IPCC Sixth Assessment Report (O’Neill et al. (2016), open source at doi:10.5194/gmd-9-3461-2016).

The following maps in Figure 19 show the spatial pattern

of the temperature increase by 1.5°C and 4.0°C on average based on IPCC AR6.

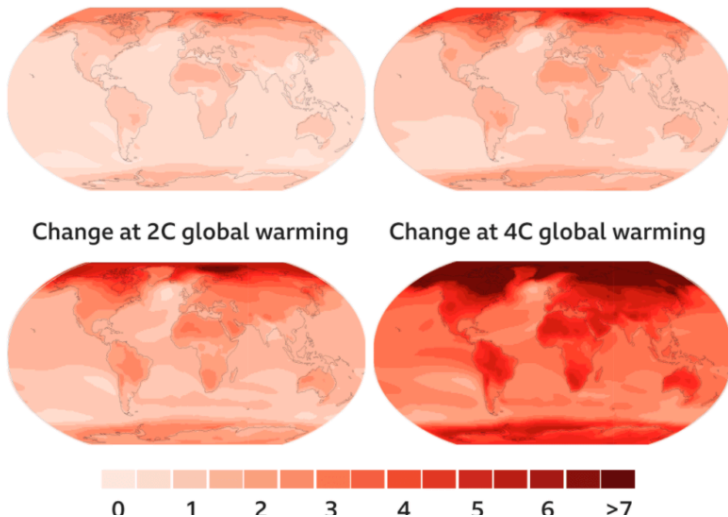


Figure 19. Projected spatial distribution of global temperature when increased by 1.5°C and 4.0°C on average based on IPCC AR6 (image source: IPCC AR6).

7. Mitigation and adaptation to climate change

Mitigation and adaptation are the two major approaches to fight against climate change. **Mitigation** measures are to reduce or prevent the emission of greenhouse gases into the atmosphere; while **adaptation** actions are to reduce the negative impact of climate change and to take the advantage of potential new opportunities.

Mitigation addresses the root causes while adaptation seeks to lower the risks due to climatic changes. Despite various mitigation measures that can reduce the level of the emission of greenhouse gases, substantial climate change appears inevitable. Thus, actions for climate change mitigation and adaptation are both required.

7.1. Mitigation to climate change

The following are possible measures for climate change mitigation:

- Applying new technologies,
- Using renewable energies,
- Electrification of industrial processes and transportation,
- Carbon tax and emissions markets,
- Changing consumer behavior or the way people live.

7.2. Adaptation to climate change

The following are possible activities to adapt to climate change:

- More secure facility locations and infrastructures,
- Landscape restoration (natural landscape) and reforestation,
- Flexible and diverse cultivation to be prepared for natural catastrophes,
- Research and development on possible catastrophes, temperature behavior, etc.,
- Preventive and precautionary measures (evacuation plans, health issues, etc.),
- Assisted migration of species and populations to match their suitable climate conditions.

Summary

Weather changes a lot day by day, but we expect that the climate remains constant for the same time of the year. Climate variability refers to short-term variation while climate change is a change in a long-term trend. Climate change has been observed with shifted mean increased probability of extremes. The main cause of climate change is anthropogenic emissions of greenhouse gases (GHGs). Projected future climates based on a range of GHGs emission scenarios are threatening. Climate change mitigation and adaptation are the two options to fight against climate change, and forest ecosystems play important roles in both.

Topic 1.2. Impacts of Climate Change on Forest Ecosystems

Tongli Wang

Contents

1. Forest ecosystems

1.1. Definition of a forest ecosystem

1.2. Distribution of forest ecosystems

1.3. Biomes

1.4. Classification of forest ecosystem

1.5. Functions of forest ecosystems

2. Impacts of Climate Change on Forest growth and productivity

2.1. Positive impact on forest growth and productivity

2.2. Negative impact on forest growth and productivity

3. Impacts of Climate Change on the vulnerability of forest ecosystems to pests and diseases

3.1. Impacts on the vulnerability of forest ecosystems to pests

3.2. Impacts on the vulnerability of forest ecosystems to diseases

4. Impacts of Climate Change on forest fires

5. Impacts of Climate Change on biodiversity and sustainability of forest ecosystems

5.1. Impacts on biodiversity

5.2. Impacts on sustainability

6. Impacts of Climate Change on Forest Ecosystems' services

[6.1. Effects on water quantity and quality](#)

[6.2. Impact on recreation provided by forest ecosystems](#)

[7. Mitigation of climate change by forest ecosystems](#)

[7.1. Carbon sequestration](#)

[7.2. Renewable materials](#)

[Summary](#)

1. Forest ecosystems

1.1 Definition of a forest ecosystem

A forest ecosystem is a terrestrial unit consisting of plants, animals and micro-organisms (biotic components), all interacting among themselves and with the environmental (abiotic) factors including soil, climate, water, and light. The organisms involved in a forest ecosystem are interdependent on one another for survival and can be broadly classified according to their ecological role as producers, consumers, and decomposers. Healthy forest ecosystems are usually self-sustaining. Larger ecosystems with wide diversity are much more stable and resilient to harmful changes.

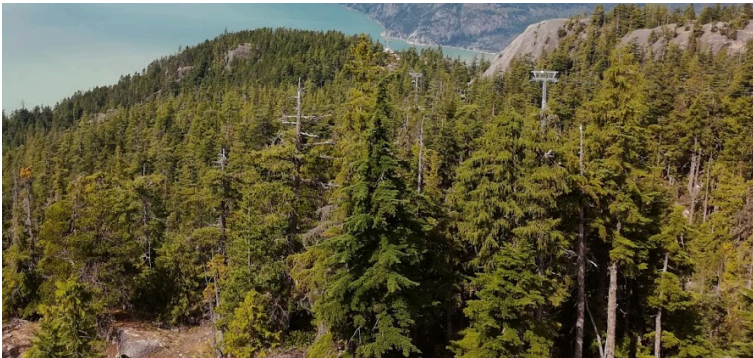


Figure 1. A forest ecosystem British Columbia. Image by Tongli Wang under CC BY-SA.

1.2. Distribution of forest ecosystems

The world's total forest area in 2010 is estimated to be just over 4 billion hectares, occupying 31% of the world land areas, and corresponding to an average of 0.6 ha of forest per capita. The five most forest-rich countries (Russia, Brazil, Canada, the United States of America, and China) account for more than half of the total forest areas (53 percent). For detail, please read the FAO report at <http://www.fao.org/3/i1757e/i1757e02.pdf>.

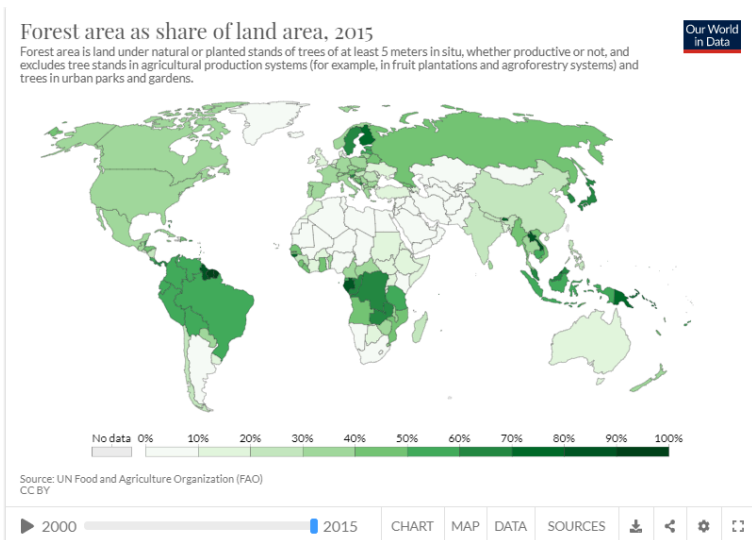


Figure 2. Global forest cover. Image by FAO under CC BY. Source: <https://ourworldindata.org/forests>

1.3. Biomes

A **biome** is a specific geographic area of the planet that can be classified according to the plants and animals that live in it. Temperature, soil, and the amount of light and water help determine what life exists in a biome.

Classification of biomes provides a basis for forest ecosystem classification. However, not all scientists classify biomes in the same way. Some scientists use broad classifications and count as few as six biomes. These are forest, grassland, freshwater, marine, desert, and tundra. Others use more precise classifications and list dozens of different biomes. For example, they consider different kinds of forests to be different biomes, as shown in Figure 3. Tropical rain forests that are warm and wet year-round are one biome. Temperate deciduous forests—those that have cold winters, warm summers, and are dominated by trees that lose their leaves—are a different biome. Taiga forests, which are in cold regions and are dominated by cone-bearing firs and spruces, are yet another biome.

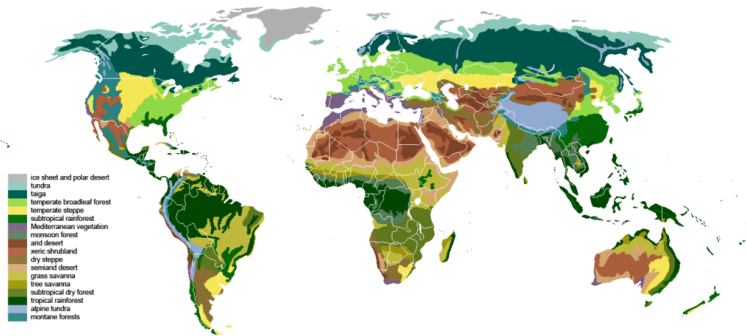


Figure 3. One way of mapping terrestrial biomes around the world. Image by Ville Koistinen under CC BY-SA 3.0. Source: <https://en.wikipedia.org/wiki/Biome>

Biomes move as the climate changes. Ten thousand years ago, parts of North Africa were lush landscapes cut by flowing rivers. Hippopotamuses, giraffes, and crocodiles lived amid abundant trees. Gradually, the climate dried out. Today, this region is part of the Sahara Desert, the world's largest desert.

1.4. Classification of forest ecosystems

Forest ecosystem classification is often in a hierarchal system, from a landscape level at the top and to a site level at the bottom. Forest ecosystems can be classified according to the type of biome in which they exist, particularly for the classification at the landscape level.

At a landscape level, Forest ecosystem classification provides a framework for landscape analysis and planning. It is based on features such as climate, elevation, topography, bedrock formation, and vegetation. For example, the Biogeoclimatic Ecosystem Classification (BEC) divides British Columbia into 16 zones (Figure 4).

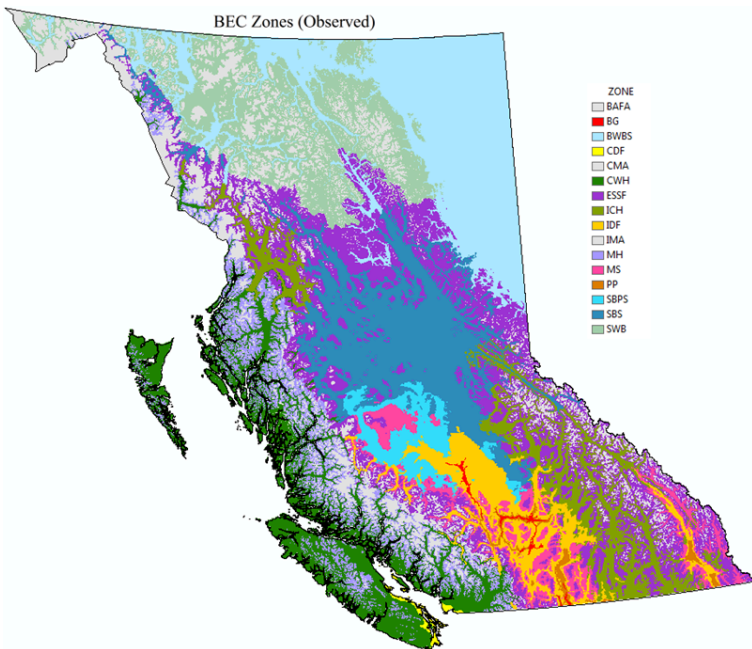


Figure 4. Biogeoclimatic zones in British Columbia, Canada. The full names of the BEC zone can be found on the [official website of the system](#). Image by Tongli Wang under CC BY-SA.

At a stand level, classifying forest ecosystems is based on vegetation, soil, and site attributes, which allows users to recognize similar ecosystem units on the ground and to develop a common strategy for forest resources management.

A forest ecosystem is often named by the primary tree species that form the canopy of the system. For example, the Coastal Douglas-fir forest ecosystem in North America (Figure 3), and Chinese pine forest ecosystem in China.

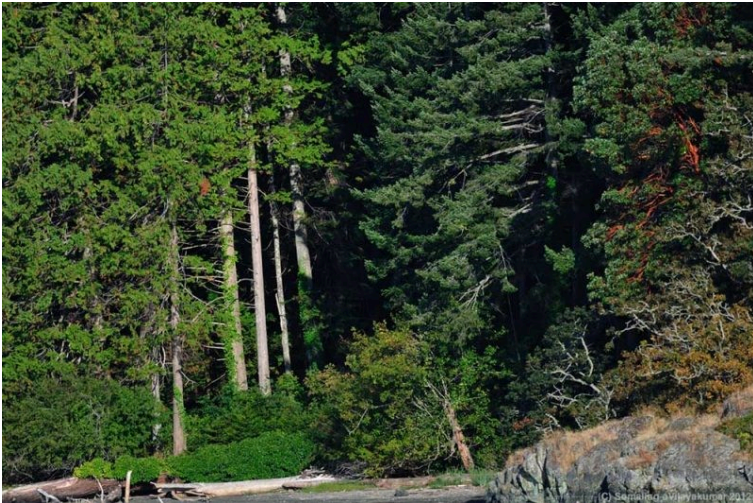


Figure 5. Coastal Douglas-fir Ecosystem in British Columbia. Photo © Vijayakumar Somalinga. Source: <https://islandnature.ca/2013/09/coastal-douglas-fir-forests/>

1.5. Functions of forest ecosystems

In the past, timber production was regarded as the dominant function of forests. In recent years, however, this perception has shifted to include other forest functions and services, such as:

- Water quantity and quality
- Recreation,
- Health and well-being (fresh air),
- Biodiversity,
- Mitigation of climate change

These will be covered in more detail later in this topic.

2. Impacts of Climate Change on Forest growth and productivity

2.1. Positive impact on forest growth and productivity

In the short term, increased temperature and prolonged growth periods (Figure 6) may improve tree growth in cooler regions for some period, where temperature is a limiting factor for tree growth.

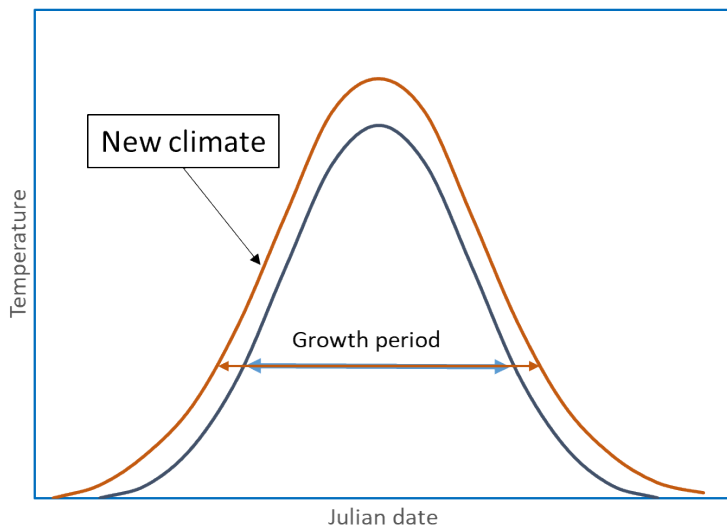


Figure 6. Illustration of a prolonged growth period caused by climate change. Image by Tongli Wang under CC BY-SA.

Elevated CO₂ concentration may promote the photosynthesis rate. Such an effect may increase under warmer temperatures as shown in the graph of Figure 7.

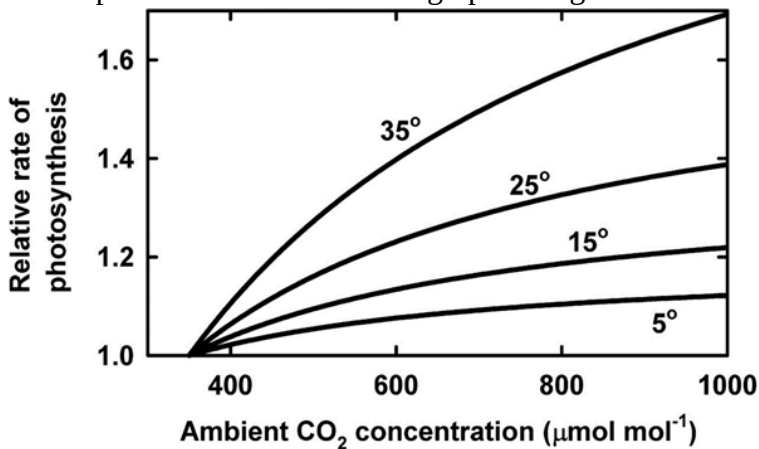


Figure 7. Photosynthetic response to CO₂ concentration, shown at four temperatures, based on

calculations for RuBP regeneration-limited photosynthesis. Image by Kirschbaum (2011). Open access source: <http://www.plantphysiol.org/content/155/1/117>

Elevated CO₂ concentration may also promote water use efficiency. Plants take in carbon dioxide and lose water vapor through stomata. Under elevated carbon dioxide levels, the stomata need to open smaller and the plant loses less water.

2.2. Negative impact on forest growth and productivity

In the long term, climate change will cause a mismatch between the climate that a tree species has adapted to (through long-time evolution and adaptation) and the climate that the species will experience in the future, as illustrated in Figure 8. Such mismatches will likely compromise the growth and productivity of the species. Using ecological models to determine the ranges of suitable climate conditions for each of the major forest tree species and forest ecotypes, and to predict the impact of climate change on their suitable climate conditions will be a major objective of this course.

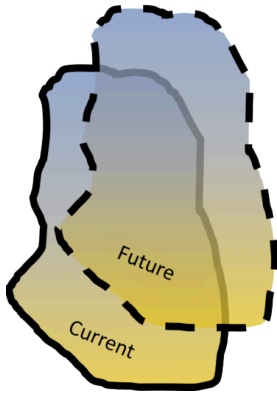


Figure 8. Illustration of the habitats suitable for a given species under current and future climates. Image by Tongli Wang under CC BY-SA.

Climate change will likely increase the risk of drought in some areas due to the increase in both plant transpiration and soil evaporation. Forest mortality associated with drought has been documented recently from all wooded continents (Figure 9) and from diverse forest types and climatic zones. It is clear that climate-induced drought and heat stress have the potential to cause forest dieback across a broad range of forest and woodland types around the world.

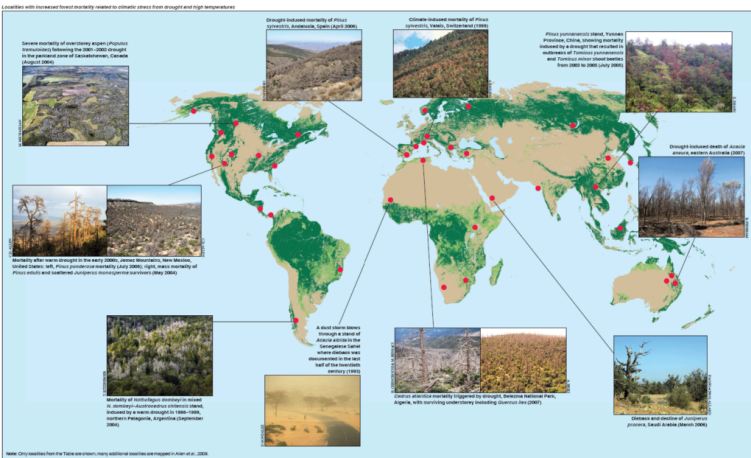


Figure 9. Localities with increased forest mortality related to climate stress from drought and high temperature. Image by FAO (in public domain), source: <http://www.fao.org/3/i0670e10.htm>

3. Impacts of Climate Change on the vulnerability of forest ecosystems to pests and diseases

3.1. Impacts on the vulnerability of forest ecosystems to pests

Climate change may impact the balance between the hosts (trees) and pests, and increase the susceptibility of the trees to pests. For example, warmed winters in British Columbia, Canada, resulted in the majority of the mountain pine beetle not being killed by the cold temperature during winters and led to its outbreak, and devastated 18 million ha forests (Figure 10).



Figure 10. Lodgepole pine destroyed by mountain pine beetle in British Columbia. Image by UBC Micrometeorology (flickr). Source: <https://www.flickr.com/photos/140969380@N07/28371273132/>

3.2. Impacts on the vulnerability of forest ecosystems to diseases

Similarly, climate change may also facilitate the spread of diseases. For example, *Dothistroma* needle blight historically has had only minor impacts on native forest trees, and mortality due to any foliar disease in natural forests is rare. Recently, however, this has dramatically changed in the northern temperate forests of British Columbia, Canada, where *Dothistroma* needle blight has severely affected native lodgepole pine trees (Figure 11) ([Woods et al. 2016](#)).



Figure 11. Severely attacked lodgepole pine stand showing heavy defoliation in the lower crown. Image by A. Woods (in public domain). Source: <https://www2.gov.bc.ca/gov/content/industry/forestry/managing-our-forest-resources/forest-health/forest-pests/foliar-pathogens/dothistroma>

4. Impacts of Climate Change on forest fires

Climate change prolongs natural fire season. A warmer climate makes more hot and dry days. A warmer climate with severer extreme climate events (hot and dry) makes forest fire more likely. More fuels resulted from devastated forests due to climate change-induced breakout of pests and diseases also increase the probability of occurrence for forest fires (Figure 13).



Between 2003 and 2012, approximately 67 million hectares of forest land burned annually.

Figure 12. Forest fire. Image by FAO (public domain). Source: <http://www.fao.org/forestry/firemanagement/en/>

There is evidence from some regions that the trend is towards more fires affecting a larger area and burning with greater intensity. Extreme wildfires have occurred frequently across the globe over the past few years, for example, the 2018 and [2020 California wildfire season](#), [the 2019 Amazon rainforest wildfires](#), and the [2019–20 Australian bushfire season](#).

5. Impacts of Climate Change on biodiversity and sustainability of forest ecosystems

5.1. Impacts on biodiversity

Healthy ecosystems comprise a vast assortment of plant and animal life, from soil microbes to top-level predators like bears and wolves.



Figure 13. Forest biodiversity. Image by CIFOR under CC BY-NC-ND. Source (flickr): <https://www.flickr.com/photos/cifor/35490446384/>

The shift and abnormal change in climate patterns caused by climate change may harm many species and their interactions. Most susceptible species may be affected first, which may sequentially impact some other species, and eventually impact the stability of the entire ecosystem. The impact may start with affecting growth and health, then fitness and survival. Climate change is considered a top threat to biodiversity.



Figure 14. Forest biodiversity at risk. Image by FAO (in public domain). Source: <http://www.fao.org/news/story/en/item/45904/icode/>

5.2. Impacts on sustainability

Healthy ecosystems are generally self-sustaining. A self-sustaining community is composed of interdependent organisms (plants, insects, animals) and their natural environment (Figure 15). It provides the food chain through which energy flows and the biological cycles that recycle essential nutrients and wastes. Climate change may substantially impact the environment, biodiversity, resilience to pests, diseases, and forest fires, thus interrupting the food chain and impacting the sustainability of the system.



Figure 15. A self-sustainable forest ecosystem. Image by Erik Olsson under CC BY. Source: <https://www.flickr.com/photos/158241364@N02/36747678793/>

6. Impacts of Climate Change on Forest Ecosystems' services

6.1. Effects on water quantity and quality

Forests improve stream quality and watershed health primarily by decreasing the amount of stormwater runoff and pollutants that reach our local waters. Tree roots and leaf litter create soil conditions that promote the infiltration of rainwater into the soil. This helps to replenish our groundwater supply and maintain streamflow during dry periods.



Figure 16. Forests and water. Image by Riccardo Maria Mantero under CC BY-NC-ND. Source: <https://www.flickr.com/photos/ljsilver71/32786989401/sizes/l/>

In the past, forests were considered the best land cover to maximize water yield, regulate seasonal flows, and ensure high water quality. Forest hydrology research conducted during the 1980s and 1990s confirmed the important role of upstream forest cover in ensuring the delivery of high-quality water but suggests that, especially in arid or semi-arid ecosystems, forests are not the best land cover to increase downstream water yield due to use of limited water resources by trees (Figure 17).



Figure 17. A forest in a dry area. Image by Water Alternatives under CC BY-NC. Source: https://www.flickr.com/photos/water_alternatives/45824476661/sizes//

Global climate models predict marked changes in seasonal snowfall, rainfall, and evaporation in many parts of the world. In the context of these changes, the influence of forests on water quantity and quality may be negative or positive. Where large-scale forest planting is contemplated for climate change mitigation, it is essential to ensure that it will not accentuate water shortages.

6.2. Impact on recreation provided by forest ecosystems

Forests provide an environment where it is possible to escape from the stresses of modern life, to connect with nature. Forests also provide wild areas where activities

such as mountain biking, orienteering, and other active and challenging sports can be pursued.



*Figure 18. Illustration of weather, climate and climate change.
Images source (in public domain): <https://vtcommunityforestry.org/news/rfp-recreation-planning-assistance-ten-vermont-town-forests>*

The development of forests for recreation is seen as an important aspect of sustainable forestry and there is increasing recognition of the wide-ranging benefits that forests provide to society. Today, recreation is perhaps the most important non-timber service provided by forest ecosystems.



Figure 19. Trails are built into the forest for recreation purposes.
Image by Kevin M Klerks under CC BY. Source:
<https://www.flickr.com/photos/ledicarus/29730757907/sizes/l/>

The negative impacts of climate change mentioned above will likely compromise the recreation values of forest ecosystems as well.

7. Mitigation of climate change by forest ecosystems

7.1. Carbon sequestration

Forest ecosystems can serve as a carbon sink by carbon sequestration through photosynthesis. Forest ecosystems absorb about 1/3 of all anthropogenic CO₂ emissions (Figure 20). A recent study suggests that our ecosystems could support an additional 0.9 billion hectares of

continuous forest. Such a change has the potential to cut the atmospheric carbon pool by about 25% ([Science 2019](#)).

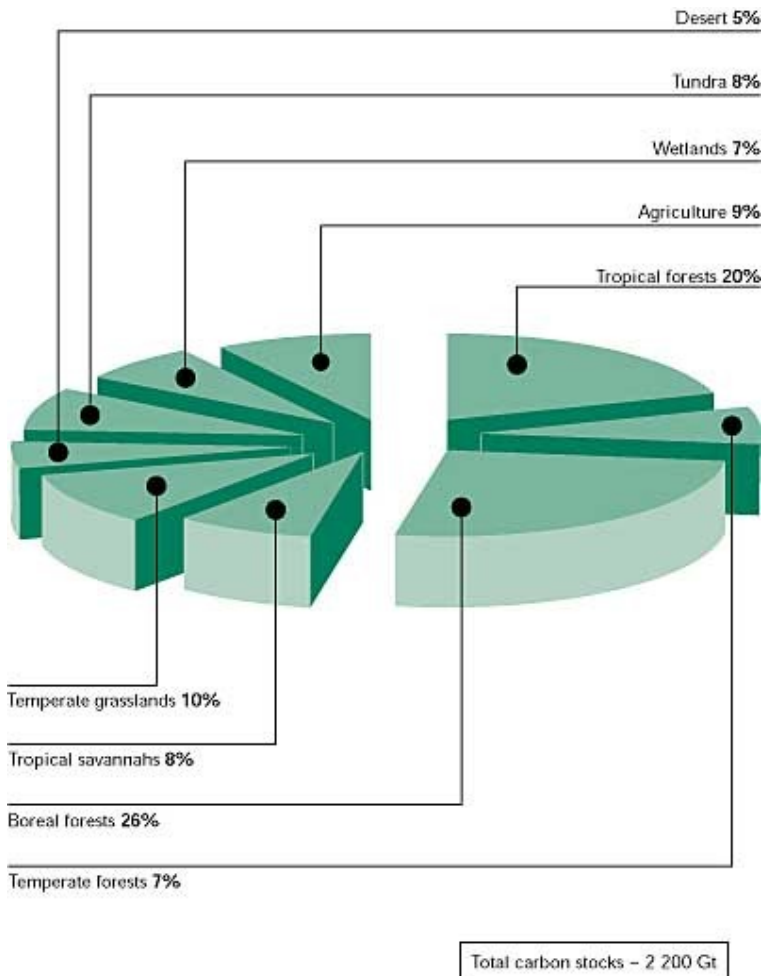


Figure 20. Terrestrial carbon stocks by ecosystem. Image by FAO under public domain. Source: <http://www.fao.org/3/y0900e/y0900e06.htm>

However, for a forest ecosystem to serve as a carbon sink, it needs to be healthy and productive, as shown in Figure

21, thus itself needs to be adaptive to climate change. As mentioned above, climate change will cause a mismatch between the climate that forest tree species have adapted to and the climate that the species will experience in the future. Such mismatch will likely result in maladaptation and compromise the capacity of forest ecosystems to sequester carbon, and even convert the forest ecosystem from a carbon sink into a carbon source, as shown in Figure 22.



Figure 21. The upper: A healthy forest ecosystem. Source: Kootenay Association for Science & Technology, <https://kast.com/beyond-computers-internet-innovation-in-bc-forestry/> (in public domain)



Image credit: Environmental Energetics

Figure 22. A lodgepole pine forest damaged by the mountain pine beetle. Image: Environmental Energetics in the public domain. Source: <https://onetreaplanted.org/blogs/stories/bark-beetles-insects-harm-trees>

7.2. Renewable materials

Forest ecosystems provide green materials (timber) to reduce the use of carbon emission-intensive materials, such as steel and cement. With increasing concerns about climate change, wood materials are increasingly used in construction, and wooden buildings are getting taller and taller (Figure 23).



Figure 23. An 18-story hybrid mass timber residence at the University of British Columbia (UBC). Source: <http://news.ubc.ca>.

Summary

In this lecture, I introduced the definition of forest ecosystems, biomes, and their classifications. Climate change may impose the most important threat to forest ecosystems and their functions. A warming temperature and elevated CO₂ concentration may improve the growth and productivity of the forest ecosystems in the short term. However, in the long run, climatic mismatches may impact the adaptation of plant species and compromise the functions and services of forest ecosystems.

Module II. Introduction to Climatic and Ecological Models

Topic 2.1. Introduction to Climate Data and Climate Models

Tongli Wang



Video #1: The importance of climate data:



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://pressbooks.bccampus.ca/fode010notebook/?p=34#video-34-1>

Video #2: How to get climate data (1/2):



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://pressbooks.bccampus.ca/fode010notebook/?p=34#video-34-2>

Video #2: How to get climate data (2/2):



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://pressbooks.bccampus.ca/fode010notebook/?p=34#video-34-3>

Contents

1. The importance of climate data

1.1. Determining terrestrial biome types

1.2. Controlling the distribution of plant species

1.3. Affecting the local adaptation of populations within a species

1.4. Essential for climate change-related studies

2. How to obtain climate data?

2.1. Weather stations

2.2. Interpolated climate data

2.3. Climate models

2.4. Climate models

3. Climate variables and time intervals

3.1. Essential Climate Variables

[3.2. Observed and calculated climate variables](#)

[3.3. Time scales](#)

[3. Major sources of climate data](#)

[Summary](#)

1. Why climate data are important?

1.1 Determining terrestrial biome types.

Each biome has a specific range of temperatures and levels of precipitation (including both rainfall and snowfall). The patterns of global distributions of climate (Figure 1) and biome classifications (Figure 2) are so similar. Typically, tropical climates have tropical forests while boreal climates support boreal forests. If we know what temperature and precipitation are like in a location, we can often predict what type of biome will be found there. Climate variables have been used to build ecological models to predict biomes and forest ecosystems.

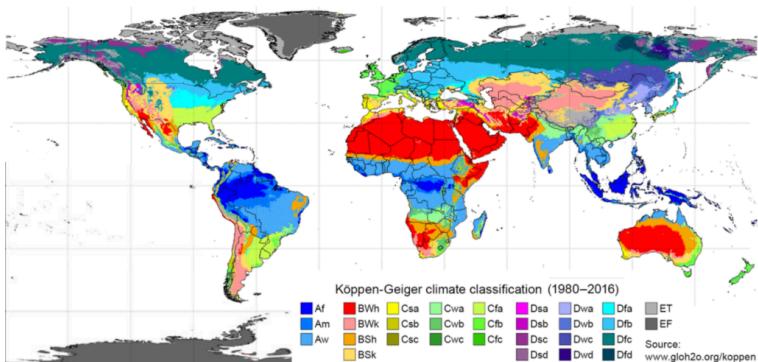


Figure 1. Köppen-Geiger climate classification. Image under CC BY-SA 4.0

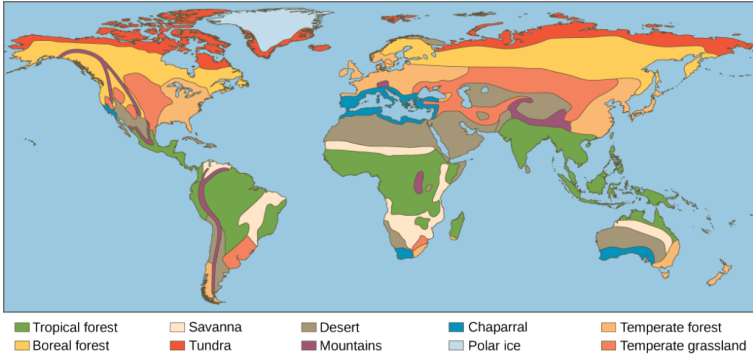


Figure 2. Eight major terrestrial biomes, along with mountains and polar ice. Image credit: Biomes: by OpenStax College, Biology, CC BY 4.0

1.2. Controlling the distribution of plant species

The distributions of plants are primarily controlled by climate ([Tuhkanen, 1980](#); [Woodward and Williams, 1987](#)). Thus, climate variables have widely been used to build species distribution models and climate niches models, which will be covered later in this course.

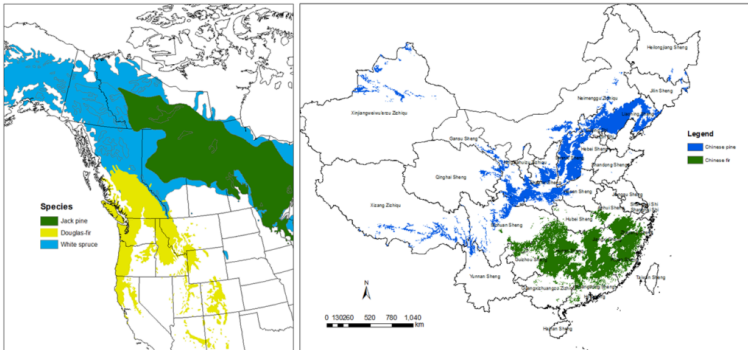


Figure 3. Forest tree species distributions in North America and China. Images by Tongli Wang under CC BY-SA

1.3 Affecting the local adaptation of populations within a species

Climate is also a major factor affecting the adaptation and productivity of forest tree populations. Population transfer functions are built to determine the effect of transfer distance of a population in terms of climate variables (Figure 4, left), and population response functions (Figure 4, right) are developed to model the response of populations to climate.

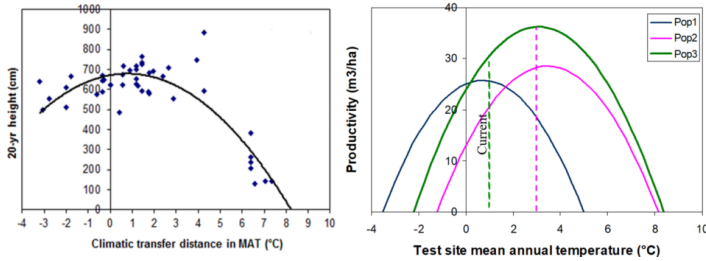


Figure 4. A climatic transfer function (left) and response function (right) for lodgepole pine populations based on mean annual temperature (MAT). Images by Tongli Wang under CC BY-SA.

1.4. Essential for climate change-related studies

Climate data are essential in most of the climate change-related studies and applications. Historical climate data are used to build and validate various types of models as mentioned above. Projected future climate data are used to assess the impact of climate change on ecosystems, species, and populations. Paleoclimate data can be used to backcast for studying the paleo history of plant migration and evolution (Figure 5).

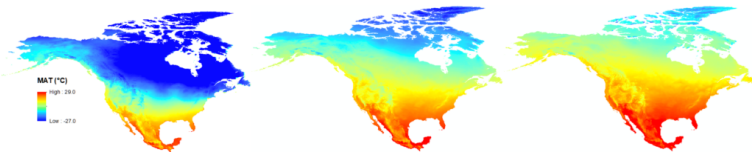


Figure 5. Maps of temperatures for the periods of last glacial maximum (21k years ago, left), current (middle) and the end of this century (right). Image by Tongli Wang under CC BY-SA.

2. How to obtain climate data?

2.1. Weather stations

Observations from weather stations (Figure 6) provide the most accurate climate data. However, weather stations are expensive to establish and maintain. Thus, the number of weather stations is limited, especially in developing countries. Even in developed countries, weather stations are mostly located in areas with easy access near cities or in valleys. The representation of climate data from weather stations for climate conditions of forest stands or plots is far from being adequate. Therefore, climate models are necessary to generate climate data for full coverages.

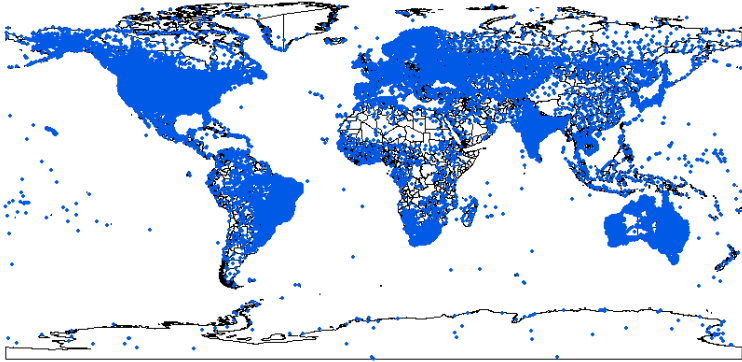


Figure 6. Distribution of weather stations over the world. Image by Tongli Wang under CC BY-SA.

2.2. Interpolated climate data

Spatially interpolated climate data generated using various interpolation approaches can provide climate data for everywhere within its coverage. The interpolation is based

on observations from weather stations. The quality of the climate data depends on both the density of the weather stations and the performance of the climate model used to make the interpolations. In most cases, interpolated climate data are delivered at certain spatial resolutions, which also affect the accuracy of the climate data for specific locations, especially for locations in mountains, where climate varies a lot within small areas due to the changes in elevation. There are many climate datasets that provide interpolated climate data. The most widely used ones include WorldClim, PRISM, CRU, and ClimateNA, which are to be described later in this topic.

2.3. Climate models

Climate models are systems of differential equations based on the basic laws of physics, fluid motion, and chemistry. They are used for a variety of purposes from the study of the dynamics of the climate system to projections of future climate.

Climate models vary in complexity:

- A simple radiant heat transfer model treats the earth as a single point and averages outgoing energy.
- This can be expanded vertically (radiative-convective models) and/or horizontally.
- Finally, (coupled) atmosphere–ocean–sea ice global climate models solve the full equations for mass and energy transfer and radiant exchange, such as General Circulation Models (GCMs).

GCMs discretize the equations for fluid motion and energy transfer and integrate these over time. Atmospheric GCMs (AGCMs) model the atmosphere and impose sea surface temperatures as boundary conditions. Coupled atmosphere-ocean GCMs (AOGCMs, e.g. HadCM3, EdGCM, GFDL CM2.X, ARPEGE-Climat) combine the two models.

The first general circulation climate model that combined both oceanic and atmospheric processes was developed in the late 1960s at the NOAA Geophysical Fluid Dynamics Laboratory. AOGCMs represent the pinnacle of complexity in climate models and internalize as many processes as possible. However, they are still under development and uncertainties remain. They may be coupled to models of other processes, such as the carbon cycle, so as to better model feedback effects. Such integrated multi-system models are sometimes referred to as either “earth system models” or “global climate models.”



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://pressbooks.bccampus.ca/fode010notebook/?p=34#oembed-1>

Video 1. This video explores how global climate models provide weather forecasts and climate projections, including details on model resolution and reliability. This video has been produced by CSIRO in collaboration with the Australian Bureau of Meteorology with funding from

Inspiring Australia. It is posted on Youtube at <https://youtu.be/Ir7J0eEuWgk>

2.4. Coupled Model Intercomparison Project (CMIP)

CMIP, led by World Climate Research Program, is to conduct intercomparison among models from >30 groups and to set standards for all the groups. CMIP is a framework for climate model experiments, allowing scientists to analyze, validate and improve GCMs in a systematic way. The 2013 IPCC fifth assessment report (AR5) included climate models from CMIP5, while the upcoming 2021 IPCC sixth assessment report (AR6) will feature new CMIP6 models.

CMIP5 uses Representative Concentration Pathways (RCPs) for four greenhouse gas concentration (not emissions) trajectories: RCP2.6, RCP 4.5, RCP 6.0, and RCP 8.5. RCP values are peaking values in radiative forcing before 2100 and decline. For example, the peaking value for RCP2.6 is about 2.6 W/m^2 .

CMIP6 updated CMIP5 scenarios and the new scenarios are called SSP1-2.6, SSP2-4.5, SSP4-6.0, and SSP5-8.5. Each of which results in similar 2100 radiative forcing levels as their predecessor in CMIP5. A number of new scenarios are added, including: SSP1-1.9, SSP4-3.4, SSP5-3.4 and SSP3-7.0. The levels of CO₂ emissions in CMIP6 are higher than that in CMIP5 in RCP2.6 and RCP8.5, but lower in RCP6.0.

3. Climate variables and time intervals

3.1. Essential Climate Variables

Essential Climate Variables (ECVs) are the climate variables that critically contribute to the characterization of the earth's climate. The ECVs of the earth's surface include temperature, precipitation, wind speed and direction, radiation, pressure, and water vapour. However, only temperature, precipitation, and their combinations, sometimes radiation, are often used in developing ecological models.

3.2. Observed and calculated climate variables

Hourly temperature and precipitation are the most commonly observed variables. Daily, monthly, seasonal and annual minimum, maximum and average values are often calculated.

Biologically relevant climate variables, such as degree-days (in various forms), dryness, moisture deficit, etc. are also often calculated for ecological modelling.

3.3. Time scales

Climate variables can be obtained at different time intervals or frequencies, such as hourly, daily, monthly, and annually. Decadal and normal period (30 years) averages are often used on many occasions. As mentioned in Module 1, climate refers to long-term climatic conditions, typically 30-year averages.

Daily climate variables are commonly provided by

weather stations. Daily climate variables are used to calculate monthly, seasonal and annual climate variables. They are also used to calculate additional biologically-important climate variables, such as various forms of degree-days, frost-free periods, and extreme temperatures. However, the size of daily climate data is often too large to be delivered and to handle model applications. Thus, daily climate data are not available in most interpolated climate datasets.

Monthly, seasonal and annual climate variables are most widely used in ecological models, especially in climate niche and species distribution models. The file size for climate variables at these time intervals is relatively small and feasible for the use of model predictions.

Annual climate variables have special significance. Mean annual temperature (MAT) and mean annual precipitation (MAP) are the most widely used climate variables. Annual chilling degree-days, growing degree-days, and annual climatic moisture deficit (CMD) are the top important biologically relevant climate variables.

4. Major sources of climate data

4.1. WorldClim

WorldClim is a set of global climate layers (gridded climate data) with a spatial resolution of about 1 km^2 in Version 1 and 2, and lower spatial resolutions in Version 2. The [ANUSPLIN software](#) is used to interpolate climate variables observed at weather stations for both versions. ANUSPLIN is a program for interpolating noisy multi-variate data using thin plate smoothing splines. Latitude,

longitude, and elevation were used as independent variables.

WorldClim Version 1

WorldClim version 1 has average monthly climate data for minimum, mean, and maximum temperature and for precipitation for 1960-1990. The data layers were generated through interpolation of average monthly climate data from weather stations. Variables included are monthly total precipitation, and monthly mean, minimum and maximum temperature, and 19 derived bioclimatic variables (<https://www.worldclim.org/bioclim>).

Available climate data for free download (<https://www.worldclim.org/version1>) include:

- Current conditions (interpolations of observed data, representative of 1960-1990)
- Future conditions: downscaled global climate model (GCM) data from CMIP5 (IPPC Fifth Assessment)
- Past conditions (downscaled global climate model output)

WorldClim Version2

WorldClim version 2 (<https://www.worldclim.org/version2>) has average monthly climate data for minimum, mean, and maximum temperature and for precipitation for 1970-2000 only for now. Climate variables can be downloaded for different spatial resolutions, from 30 seconds ($\sim 1 \text{ km}^2$) to 10 minutes ($\sim 340 \text{ km}^2$). More climate variables are included in this version, including solar

radiation, wind speed, and water vapour. Future projections from some CMIP 6 GCMs are also available.

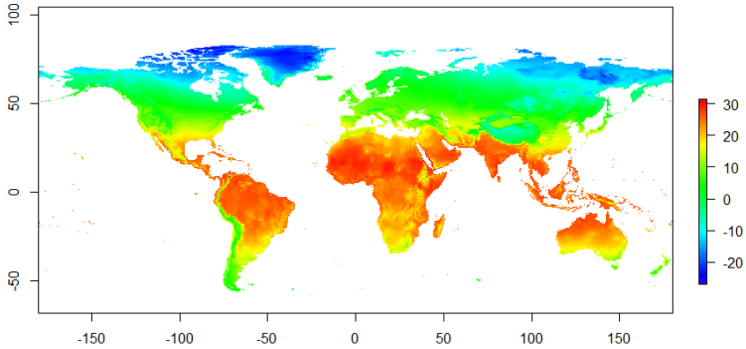


Figure 7. Global mean annual temperature (°C). Images by Tongli Wang under CC BY-SA

4.2 CRU climate data

The gridded [Climatic Research Unit \(CRU\) Time-series \(TS\)](#) data version 4.05 data are month-by-month variations in climate over the period 1901-2020, provided on high-resolution (0.5×0.5 degree) grids, produced by CRU at the University of East Anglia.

The CRU TS4.05 variables are cloud cover, diurnal temperature range, frost day frequency, potential evapotranspiration (PET), precipitation, daily mean temperature, monthly average daily maximum and minimum temperature, and vapour pressure for the period January 1901 – December 2020.

The monthly dataset was generated using Angular Distance Weighting (ADW) interpolation approach, which weights the data points by distance relative to a correlation decay distance and gives more weight to isolated data points. The time-series dataset can be downloaded at https://crudata.uea.ac.uk/cru/data/hrg/cru_ts_4.05/

The $0.5^\circ \times 0.5^\circ$ spatial resolution is too low for ecological modeling at a local, especially in mountainous regions.

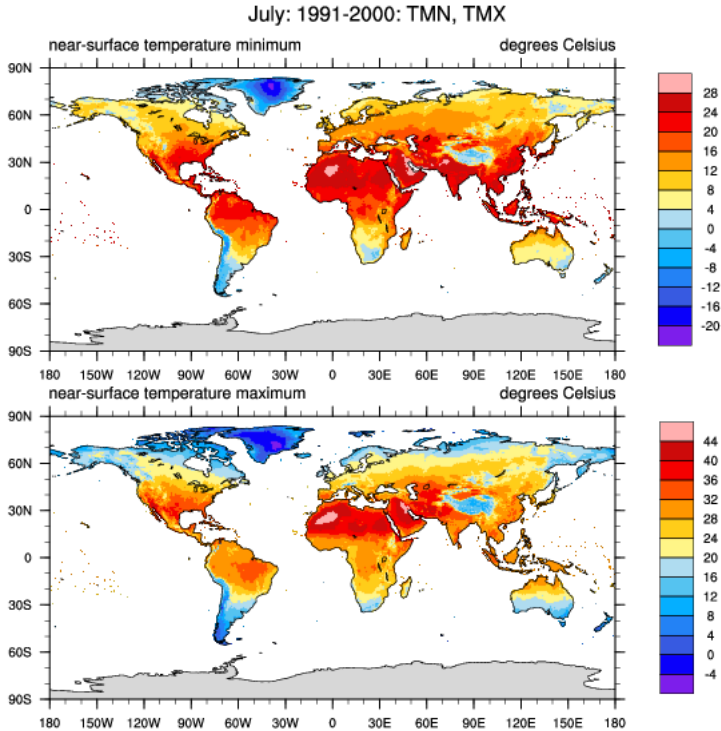


Figure 8. The coverage of the CRU time-series climate data and the patterns of temperatures. <https://climatedataguide.ucar.edu/climate-data/cru-ts-gridded-precipitation-and-other-meteorological-variables-1901>

4.3 PRISM climate data

The PRISM Climate Group gathers climate observations from a wide range of monitoring networks, applies sophisticated quality control measures, and develops

spatial climate datasets to reveal short- and long-term climate patterns. The resulting datasets incorporate a variety of modeling techniques and are available at multiple spatial/temporal resolutions, covering the period from 1895 to the present. It offers these datasets to the public, either free of charge or for a fee (depending on dataset size/complexity and funding available for the activity).

The PRISM data were developed using the PRISM model that incorporates weather station data, a digital elevation model, and expert knowledge of climate patterns such as rain shadows, coastal effects, orographic lift, and temperature inversions over topographically delineated “facets” (Daly et al. 2002). PRISM data have clear advantages over other products in reflecting these effects, especially for precipitation.

The PRISM climate datasets are available for the United States except for AK. The datasets include PRISM normals cover the period 1981-2010, monthly for Historical Past (1895-1980) and for Recent Years (Jan 1981 – May 2019) at <http://www.prism.oregonstate.edu/>.

4.4 ClimateNA

[ClimateNA](#) is a standalone MS Windows software application (Wang et al. 2016) that extracts and downscales gridded (4 x 4 km) monthly climate data for the reference normal period (1961-1990) from PRISM (Daly et al. 2008) and WorldClim (Hijmans et al. 2005) to scale-free point locations. It also calculates many (>200) monthly, seasonal and annual climate variables. The downscaling is achieved through a combination of bilinear interpolation and dynamic local elevational adjustment. ClimateNA also uses the scale-free data as a baseline to downscale

historical and future climate variables for individual years and periods between 1901 and 2100. A time-series function is available to generate climate variables for multiple locations and multiple years.

The program can read and output comma-delimited spreadsheet (CSV) files. The new version (v6.00) can also directly read digital elevation model (DEM) raster (ASC) files and output climate variables in raster format for mapping. The spatial resolution of the raster files is up to the user's preference. The coverage of the program is shown in Figure 9.

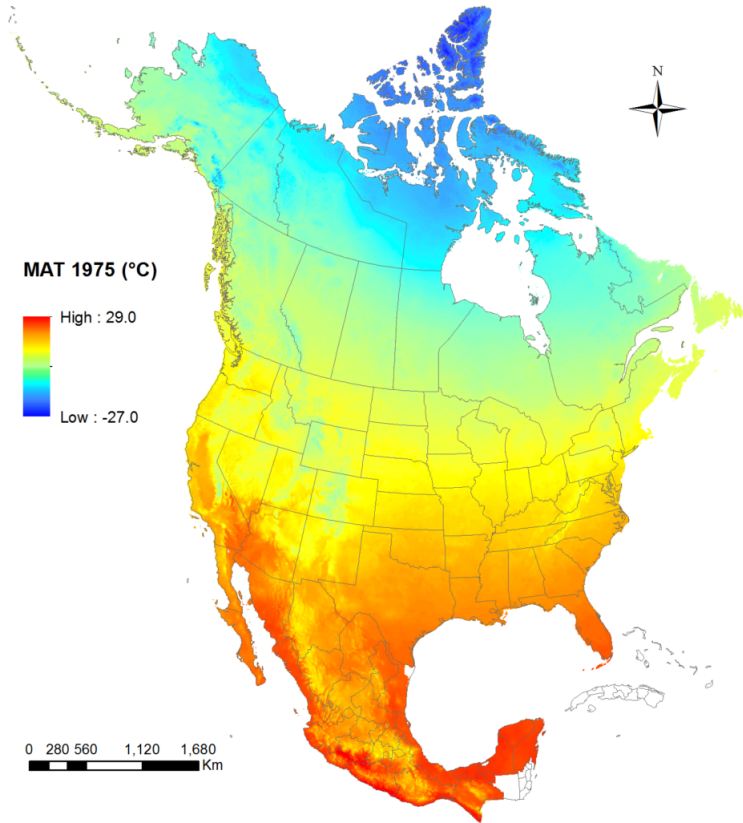


Figure 9. The coverage of ClimateNA. The extent of the coverage: longitude = $-179.167 \sim -52.625^\circ$; latitude = $14.453 \sim 83.203^\circ$. Image by Tongli Wang under CC BY-SA.

Summary

Climate is a key abiotic factor that determines terrestrial biome types. Climate also controls the distribution of plants and affects the local adaptation of populations within a species. Thus, climate data are essential for climate change-related studies. There are many sources to obtain

climate data, including observations from weather stations, interpolated surfaces, and climate models' predictions. However, the quality of the climate data and availability of climate variables vary considerably. It is important to know the sources and to understand their differences for students to conduct climatic and ecological modeling.



By Tongli Wang, Climate Modelling and Forest Applications is licensed under [Creative Commons <CC BY-NC-ND 4.0> License](https://creativecommons.org/licenses/by-nc-nd/4.0/)

Topic 2.2. Introduction to Ecological Models

Tongli Wang



Video#1: An overview and the definition of ecological models:



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://pressbooks.bccampus.ca/fode010notebook/?p=36#video-36-1>

Video#2: Types of ecological models:



One or more interactive elements has been excluded from this version of the text. You can view them online here:

[https://pressbooks.bccampus.ca/
fode010notebook/?p=36#video-36-2](https://pressbooks.bccampus.ca/fode010notebook/?p=36#video-36-2)

Video#3: Some important process-based ecological models:



One or more interactive elements has been excluded from this version of the text. You can view them online here:

[https://pressbooks.bccampus.ca/
fode010notebook/?p=36#video-36-3](https://pressbooks.bccampus.ca/fode010notebook/?p=36#video-36-3)

Contents

1 Overview

[1.1 definition](#)

[1.2 Model building](#)

[1.3 Model applications](#)

2 Types of models

[2.1 Analytic models](#)

[2.2 Simulation models](#)

[2.3 Empirical and process-based models in climate change applications](#)

3 Building process-based models

[3.1 Model design](#)

[3.2 Model validation](#)

[4 Widely used process-based models](#)

[4.1 3-PG](#)

[4.2 FORECAST](#)

[4.3 Tree and Stand Simulator \(TASS\)](#)

[4.4 LANDIS-II](#)

1 Overview

The impacts of climate change on forest ecosystems are likely to compromise tree growth, increase vulnerability to disturbances, and eventually shift in species distributions. Thus, it is likely to require changes in forest planning and natural resource management to mitigate and adapt to climate change. Ecosystem models have widely been used to understand the relationships between ecosystem and environmental conditions, and to provide a “best estimate” about how forests might work in the future and thus guide decision-making in forest resources management.

1.1 Definition

An ecosystem model, also called an ecological model, is usually a mathematical representation of an ecological system. It is a simplified form of a highly complex ecosystem in the real world. The ecological system can range in scale from an individual population to an ecological community, or even an entire biome.

1.2 Model building

Based on data gathered from the field, ecological relationships—such as the relation of sunlight and water availability to photosynthetic rate, or frost resistance to cold temperature and the length of growth period—are derived, and these are combined to form ecosystem models.

Due to the complexity of ecosystems, it is impossible to capture every possible element of a system in a model. Thus, modelers need to take decisions about which of its features to include and which to disregard. These decisions are usually guided by the aims that the modeler is attempting to achieve. It is desirable to make an ecological model to be general (the model can be applied to a wide range of real-life systems), realistic (the model's conclusions have a close match to a real-life system) and precise (the model's predictions for a specific set of circumstances have little or no uncertainty). Trading-off one desideratum in order to achieve greater performance in the other two is often involved in model-building processes.

1.3 Model applications

Ecological models enable researchers to simulate large-scale experiments that would be too costly or unethical to perform on a real ecosystem. They also enable the simulation of ecological processes over very long periods of time (i.e., simulating a process that takes centuries in reality, can be done in a matter of minutes in a computer model). More importantly, ecological models can be used to predict the state of ecosystems under future climates.

Ecosystem models have applications in a wide variety of disciplines, such as natural resource management, ecotoxicology, and environmental health, agriculture, and wildlife conservation. Applications of ecological models in forest resources management to mitigate the negative impact of climate change and help forest trees to adapt to a rapidly changing environment have become particularly important.

At present, our understanding of the relationships between ecosystems and the changing environment is relatively weak. The vulnerability of ecological systems and the services provided by them, to the impact arising from climate change remain largely unknown. Ecological models have been widely used to assess the impact of climate change in terms of productivity, carbon storage, and vulnerability to disturbances, and to provide information for developing forest adaptive strategies, such as assisted migration at the species level and the population level, which is also called assisted gene-flow.

2. Types of models

There are two major types of ecological models:

- 1) Analytic models and
- 2) Simulation models.

2.1 Analytic models

Analytic models are typically relatively simple (often linear) systems, that can be accurately described by a set of mathematical equations whose behavior is well-known.

Analytic models are also called correlative models or empirical models.

For example, a model of tree height growth of populations originated from different climates that assumes the transfer distance of populations affects the tree height growth in a quadratic relationship as:

$$Y = b_0 + b_1 * MAT_td + b_2 * MAT_td^2$$

Where Y is the tree height; MAT_td is the transfer distance in terms of mean annual temperature; MAT_td² is the quadratic form of the transfer distance; and b₀, b₁, and b₂ are parameters to be determined during the model fitting process.

A model can be acceptable if the amount of variance explained by the model is high (or the relationship is strong) based on the statistical significance ($P < 0.05$), such as the one illustrated in Figure 1.

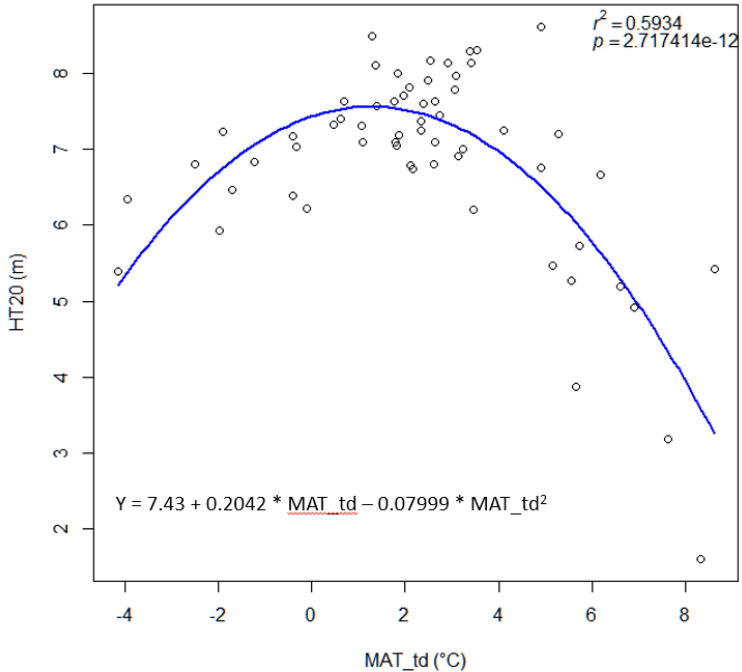


Figure 1. Illustration of an analytic model based on the relationship between climatic transfer distance (MAT_{td}) and the height growth (HT₂₀) of tree populations. Image by Tongli Wang under CC BY 4.0.

Although the solution to the above simple question is fairly transparent, analytical models describing more complex systems can often become quite complicated. However, for those comfortable with mathematics, an analytical solution does provide a concise preview of a model's behavior that is not as readily available with a simulation model. We will focus on this type of models in this course, and more details on analytic or empirical models will come in later lectures.

2.2 Simulation models

Simulation models use numerical techniques to solve problems for which analytic models are impractical or impossible. They are also called numerical models or process-based models. A simulation model approach focuses on simulating detailed physical or biological processes that explicitly describe the behavior of a system.

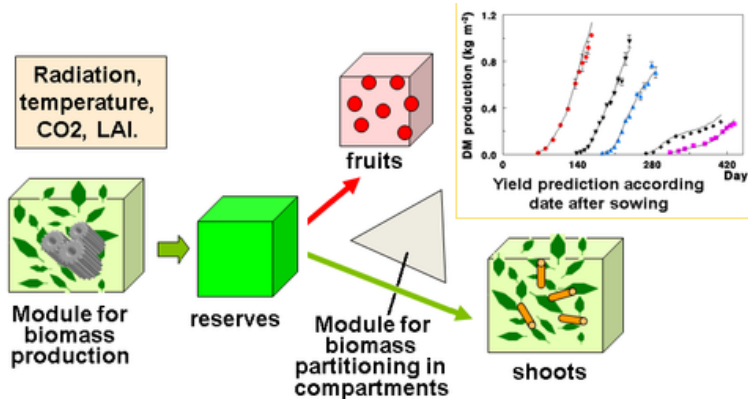


Figure 2. Illustration of a typical process-based model application. The biomass production model computes the biomass produced by the leaf compartment from the LAI and environmental conditions. The partitioning model then splits the biomass between the fruits and the other organ compartments. The model then loops for a new development cycle. Image by Drawing and graph E. Heuvelink, WAGENINGEN UNIVERSITY, P. de Reffye, CIRAD under CC BY NC SA.

As the name stated, a process-based model can simulate the process or dynamics over time. For example, many process-based models can simulate the tree growth from seedlings to harvest, or simulate the stand performance over rotations.

2.3 Empirical and process-based models in climate change applications

Both empirical and process-based models are applied in modeling the impact of climate change on species distributions, productivity, and vulnerability to stress.

The empirical approach relies on correlative relationships in line with mechanistic understanding, but without fully describing system behaviors and interactions. For example, the model represents the relationship between the transfer distance and tree height of populations. We can see that a strong relationship is represented by the model, but the model does not show the physiological process driving this relationship.

The process-based models, on the other hand, can be more comprehensive and incorporate mechanisms explicitly. For example, many process-based models include the process of how radiation, water, and temperature affect the rate of photosynthesis, and how the biomass generated from the photosynthesis allocates to different parts of the plants, as shown in Figure 2. However, the accuracy of the model predictions is often lower than an empirical model. This is because there are many factors affecting the performance of trees, populations, or an entire ecosystem, and many of them cannot be incorporated into the process-based model. With the accumulation of knowledge and advances in modeling algorithms and computing capacity, prediction accuracies of process-based models are expected to be improved with time.

Empirical models based on the occurrence data of a species, particularly with machine-learning algorithms, have demonstrated clear superiority over process-based models in predictions of species distributions. These

models are often called species distribution models (SDMs). The occurrence of the species resulted from the process of long-term local adaptation to environmental conditions, interaction with other species, and many other factors. These processes are difficult to model by a process-based model. In contrast, the empirical models do not need to consider these processes, but just relate the occurrence locations with their environmental variables. The relationships are mostly very strong. However, the extrapolation capacity of correlative models may be limited as their depending relationship may change beyond the coverage of observations. More detailed comparisons are listed in Table 1.

Table 1 | Relative differences in the characteristics of process-based and empirical modeling approaches.

	Process-based	Empirical
Relationship type	Causal	Correlative
Relative comprehensiveness	More comprehensive	Less comprehensive
Incorporation of mechanism	Explicit	Implicit
Primary source of error	Unknown parameters and processes	Extrapolation
Model uncertainty	Higher	Lower
Data requirements	Higher	Lower
Spatial scale for calibration	Smaller	Smaller to larger
Spatial scaling of prediction	Smaller to Larger	Best at scale of calibration

Source: <https://www.frontiersin.org/articles/10.3389/fpls.2013.00438/full>, under CC BY.

3. Building process-based models

3.1 Model design

Ecological systems are composed of an enormous number

of biotic and abiotic factors that interact with each other in ways that are often unpredictable, or so complex as to be impossible to incorporate into a computable model.

Because of this complexity, ecosystem models typically simplify the systems to a limited number of components that are well understood and relevant to the problem that the model is intended to solve. Thus, the process of model design needs to begin with a specification of the problem to be solved or the objectives of the model.

The process of simplification typically reduces an ecosystem to a small number of state variables and mathematical functions that describe the nature of the relationships between them. The number of ecosystem components that are incorporated into the model is limited by aggregating similar processes and entities into functional groups that are treated as a unit.

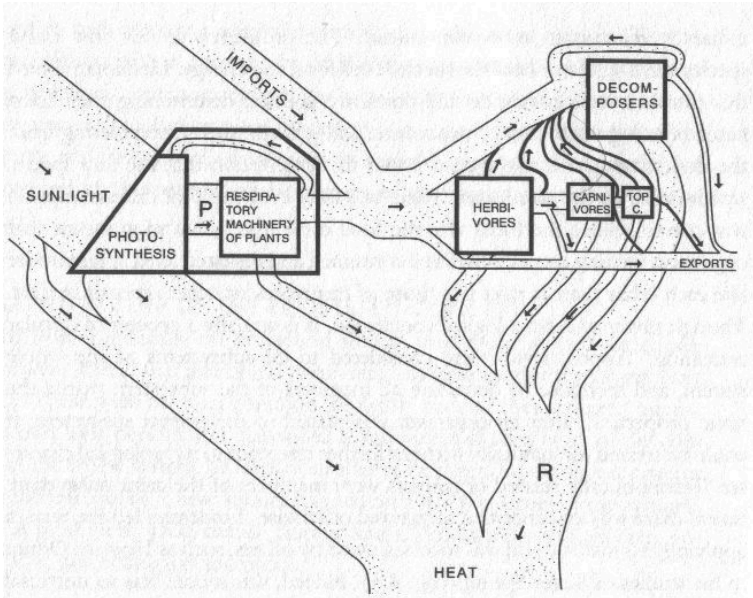


Figure 3. Diagram of the Silver Springs model. Note the aggregation into functional groups such as “herbivores” or “decomposers”. Image by ODUM from an open source at https://commons.wikimedia.org/wiki/File:Silver_Spring_Model.jpg.

Another important factor in the ecosystem model structure is the representation of the space used. Historically, models have often ignored the confounding issue of space. However, for many ecological problems, spatial dynamics are an important part of the problem, with different spatial environments leading to very different outcomes.

Spatially explicit models (also called “spatially distributed” or “landscape” models) attempt to incorporate a heterogeneous spatial environment into the model. A spatial model is one that has one or more state variables that are a function of space or can be related to other spatial variables.

3.2 Model validation

After construction, models are validated to ensure that the results are acceptably accurate or realistic.

One method is to test the model with multiple sets of data that are independent of the actual system being studied. This is important since certain inputs can cause a faulty model to output correct results.

Another method of validation is to compare the model's output with data collected from field observations.

Model predictions almost always involve errors. Researchers frequently specify beforehand how much of a disparity they are willing to accept between parameters output by a model and those computed from field data.

4. Widely used process-based models

In process-based models, the environment variables and vegetation are linked up upon the comprehensive physiological processes and functions between them. They often have a high requirement for data input. The process-based approach focuses on simulating detailed physical or biological processes that explicitly describe system behavior. They can be comprehensive and mechanism explicit. All process-based models include some empirical information. Uncertainty in process-based model outputs could be higher than for the empirical approach due to greater model parameters and data inputs to represent the many processes in the system. The application of physiological rules makes process-based models possible to mimic the vegetation growth and display the future species distribution, as well as the forest dynamics.

We will introduce several process-based forest ecological models here.

4.1 3-PG

The 3-PG (Physiological Principles Predicting Growth) forest growth model has been most widely used. The 3-PG model is a simplified physiological processes-based stand growth model developed by [Landsberg and Waring in 1997](#).

A 3-PG model calculates the radiant energy absorbed by forest canopies and converts it into biomass production. The efficiency of radiation conversion is modified by the effects of nutrition, soil drought (the model includes continuous calculation of water balance), atmospheric vapor pressure deficits, and stand age.

3-PG is a deliberate attempt to bridge the gap between conventional empirical growth and yield models, and complex process-based carbon-balance models. It can be applied to plantations, or to even-aged, relatively homogeneous forests. It is a generic stand model, in the sense that its structure is not site or species-specific. However, parameterization is necessary for improving model effectiveness when it is applied for different tree species. The 3-PG model is usually more powerful and effective after the parameter values have been adjusted for aimed tree species.

3-PG consists of five simple sub-models:

- the assimilation of carbohydrates,
- the distribution of biomass between the foliage, roots and stems,
- the determination of stem number, soil water

balance, and

- conversion of biomass into variables of interest to forest managers.

The 3-PG model has been used to simulate forest dynamics and productivity and estimate future migration direction and distribution under changing climate. A 3-PG model can also be applied together with other tools for more comprehensive and convincing performance.

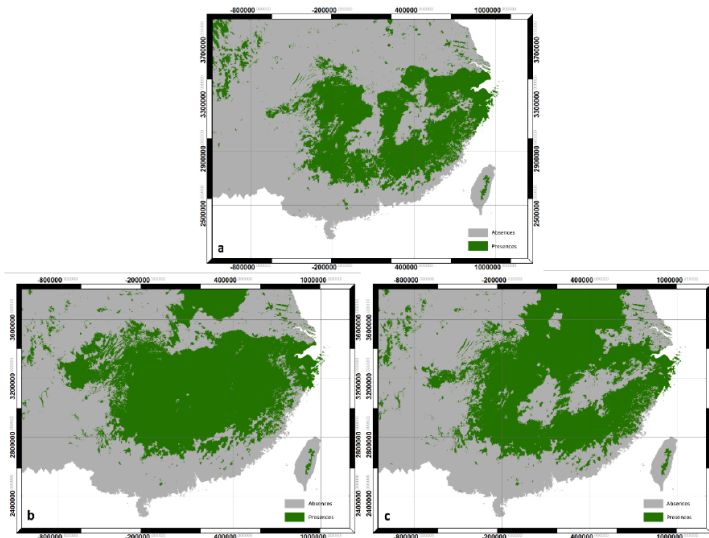


Figure 4. Chinese fir distribution under the current (a) and two future climate projections (b and c). Image by Lu et al. (2015) from *Forests*, an open-source at <https://doi.org/10.3390/f6020360>.

The 3-PG model can be downloaded for free from:

<https://3pg.forestry.ubc.ca/software/>

4.2 FORECAST

[FORECAST](#) is a management-oriented, stand-level, forest-growth, and ecosystem-dynamics model. The model was designed to accommodate a wide variety of silvicultural and harvesting systems and natural disturbance events (e.g., fire, wind, insect epidemics) in order to compare and contrast their effect on forest productivity, stand dynamics, and a series of biophysical indicators of non-timber values. The model was developed by [Kimmins and his colleagues](#) at the University of British Columbia.

FORECAST can project stand growth and ecosystem dynamics. The projection is based upon a representation of the rates of key ecological processes regulating the availability of, and competition for, light and nutrient resources (a representation of moisture effects on soil processes, plant physiology and growth, and the consequences of moisture competition is being added).

The rates of these processes are calculated from a combination of historical bioassay data (such as biomass accumulation in plant components and changes in stand density over time) and measures of certain ecosystem variables (including decomposition rates, photosynthetic saturation curves, and plant tissue nutrient concentrations) by relating ‘biologically active’ biomass components (foliage and small roots) to calculated values of nutrient uptake, the capture of light energy, and net primary production.

Using this ‘internal calibration’ or hybrid approach, the model generates a suite of growth properties for each tree and understory plant species that is to be represented in a subsequent simulation. These growth properties are used to model growth as a function of resource availability and competition. They include (but are not limited to):

(1) photosynthetic efficiency per unit foliage biomass and its nitrogen content based on relationships between foliage nitrogen, simulated self-shading, and net primary productivity after accounting for litterfall and mortality;

(2) nutrient uptake requirements based on rates of biomass accumulation and literature- or field-based measures of nutrient concentrations in different biomass components on sites of different nutritional quality (i.e. fertility);

(3) light-related measures of the tree and branch mortality derived from stand density and live canopy height input data in combination with simulated vertical light profiles. Light levels at which mortality of branches and individual trees occur are estimated for each species.

The FORECAST model has been widely applied as a long-term management evaluation tool in a variety of forest ecosystems. The FORECAST model simulates the effect of light and nutrient availability on forest productivity but it has no explicit representation of moisture or temperature on ecosystem processes. Thus, the model is not directly driven by climate data. [FORECAST Climate](#) has been developed to incorporate these two factors. The foundation for FORECAST Climate was established through the creation of a dynamic linkage with a forest hydrology model including direct feedback to the core processes driving forest ecosystem production.

4.3 Tree & Stand Simulator (TASS)

The Tree and Stand Simulator (TASS) is a biologically based, spatially explicit, individual tree model. By

simulating the potential effects of silvicultural decisions and prescriptions, TASS provides yield projections for timber supply reviews and helps resource managers identify investment opportunities. The model is developed and maintained by the Ministry of Forests, Lands, Natural Resource Ministry of Forests, Lands, Natural Resource Operations & Rural Development.

TASS currently exists in 3 main forms:

1. **TASS III** is the all-new public-release Windows™ version, which begins to extend TASS into more complex stand structures with multiple-species and -age cohorts. The initial release is limited to lodgepole pine (*Pinus contorta* var. *latifolia*) and white spruce (*Picea glauca*).
2. **TASS II** (commonly referred to as TASS) is the well-established, in-house version described below. Although, the concepts largely apply to TASS III, as well.
3. [TIPSY \(Table Interpolation for Stand Yields\)](#) provides direct operational access to yield tables generated by TASS II. For scenarios outside the limited range of TIPSY, a custom TASS run service is available by contacting the [Growth and Yield Application Specialist](#).

Photos in Figure 5 show trees remaining after variable retention harvesting by Weyerhaeuser Company Limited (the real world), and the TASS simulations. They are amazingly similar.



Figure 5. Illustration of TASS simulations of forest stands. Image by TASS group in a public domain at https://www.for.gov.bc.ca/hfd/pubs/docs/p/p074/P074_g_y.pdf.

The TASS is a height-driven, biologically oriented, spatially explicit individual tree model (Mitchell 1975). It was designed to produce potential growth and yield tables for even-aged managed stands. It simulates the growth of individual trees in a three-dimensional space with competitions. Tree height, bole diameter, branch length, and crown form and foliage are updated annually. These measures can be affected by inter-tree competition, thinning, pruning, animal damage, site characteristics, defoliation, mortality, and fertilization, as well as genetic variation (or genetic treatments, as shown in Figure 6) in tree growth.

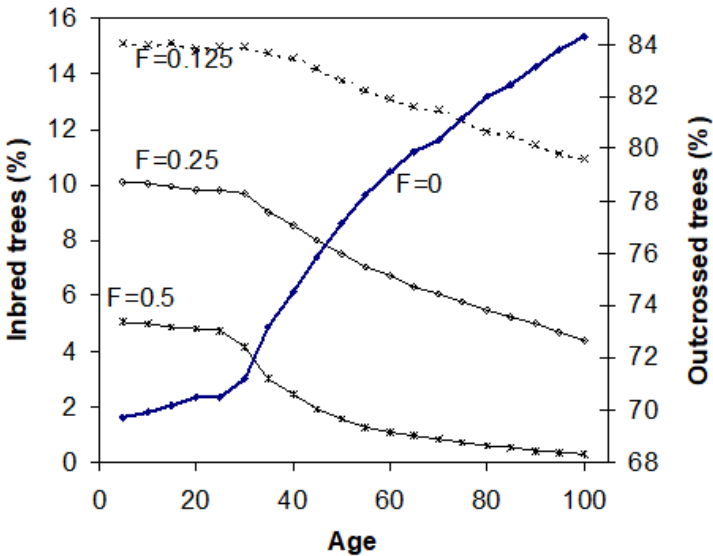


Figure 6. Changes in proportions in inbred trees by TASS simulations with the site index of 35m and planting density of 1890 trees/ha. Image by Tongli Wang, under under CC BY 4.0.

TASS can predict silvicultural treatment response by modelling individual tree crown dynamics and their relationship to bole growth and wood quality. This makes TASS particularly well suited for predicting response to treatments such as

- Espacement,
- fertilization,
- pruning,
- pre-commercial and commercial thinning.

However, effects of climate change have not been incorporated into TASS models. Scientists are working on

functions and parameters to link population response to climate into the modeling algorithms.

The software can be download for free at:

<https://www2.gov.bc.ca/gov/content/industry/forestry/managing-our-forest-resources/forest-inventory/field-forms-and-software/software-download>

4.4 LANDIS-II

The LANDIS-II forest landscape model simulates future forests (both trees and shrubs) at decadal to multi-century time scales and spatial scales spanning hundreds to millions of hectares. The model simulates change as a function of growth and succession and, optionally, as they are influenced by a range of disturbances (e.g., fire, wind, insects), forest management, land-use change. Climate and climate change affect processes throughout the model. LANDIS-II is highly customizable with dozens of libraries ('extensions') to choose from.

LANDIS-II is optimized for the simulation of spatial processes and the interactions between spatial processes and patterns. Landscapes within LANDIS-II are represented as a grid of interacting cells with a user-defined spatial resolution (cell size) and extent. Practicable cell sizes can range from a few meters up to a kilometer. All LANDIS models assume that individual cells have homogeneous light environments. Cells are aggregated into ecoregions with homogeneous climate and soils and a user-defined extent, thereby creating a hierarchy of spatial interactions.

The LANDIS-II community is very active with 100s of users and developers worldwide. Everyone is welcome to join the LANDIS-II community and to contribute. All

model components are free and open-source. There are active bulletin boards for Users and Developers. And there are meetings and training held every year in various locations across the US and Canada.

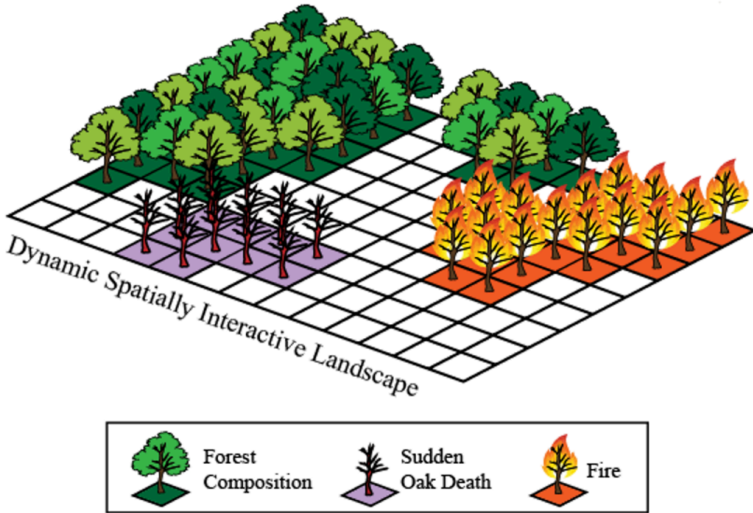


Figure 7. Illustration of Landis-II model simulations of tree growth spatially explicit. Images by Landis_II group at a public domain: <http://www.landis-ii.org/>

The following video shows an example of Landis II simulations:

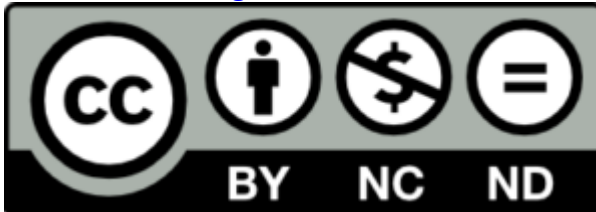


One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://pressbooks.bccampus.ca/fode010notebook/?p=36#oembed-1>

The software of the model can be install from:

<http://www.landis-ii.org/install>



By Tongli Wang, Climate Modelling and Forest Applications is licensed under [Creative Commons <CC BY-NC-ND 4.0> License](https://creativecommons.org/licenses/by-nc-nd/4.0/)

Module III. Scale-free Climate Models and Their Applications

Topic 3.1. Introduction to ClimateNA and ClimateAP

ClimateNA and ClimateAP

Tongli Wang



Video #1: Introduction to ClimateNA and ClimateAP – Overview:



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://pressbooks.bccampus.ca/fode010notebook/?p=38#video-38-1>

Video #2: Introduction to ClimateNA and ClimateAP – How do they work?



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://pressbooks.bccampus.ca/fode010notebook/?p=38#video-38-2>

Video #3: Introduction to ClimateNA and ClimateAP –
Main features:



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://pressbooks.bccampus.ca/fode010notebook/?p=38#video-38-3>

Contents

[1. Overview](#)

[2. Data sources](#)

[2.1. Baseline data](#)

[2.2. Historical data](#)

[2.3. Future climate data](#)

[2.4. Paleoclimate data](#)

[3. Dynamic local downscaling](#)

[3.1. Baseline data](#)

[3.2. Historical and Future climate data](#)

[4. Development of additional climate variables](#)

[4.1. Calculated climate variables](#)[4.2. Derived climate variables](#)[5. Main features](#)[5.1. Scale-free climate data](#)[5.2. All-in-one package](#)[5.3. Multiple-location and multiple-GCM processing](#)[5.4. Time-series functions](#)[5.5. Map-in and map-out capacity](#)[6. Web versions](#)[7. Spatial data](#)

1. Overview

[ClimateNA](#) and [ClimateAP](#) are standalone MS Windows software packages that generate scale-free climate data for historical (1901 – last year), future (2011-2100) and paleo (back to 21,000 years back) periods and years. They also calculate and derive many (>200) monthly, seasonal and annual climate variables to facilitate ecological modeling. Time-series functions are available to generate climate variables for multiple locations and multiple years for historical or future years in a single run.

The scale-free feature allows users to generate climate data for specific locations (instead of grid averages from other climate models) for better accuracy and to generate climate surfaces at any spatial resolution, which is particularly useful for applications at regional and local scales. The scale-free feature is achieved through a dynamic local downscaling to the gridded climate data from [PRISM](#) and [WorldClim](#).

The application packages can be downloaded for free. No installation is required. The program can read and output comma-delimited spreadsheet (CSV) files. The new version of ClimateNA (v6.00) can also directly read digital elevation model (DEM) raster (ASC) files and

output climate variables in raster format for spatial analysis and mapping. The spatial resolution of the raster files is up to the user's preference. ClimateNA covers the entire North America (Figures 1).

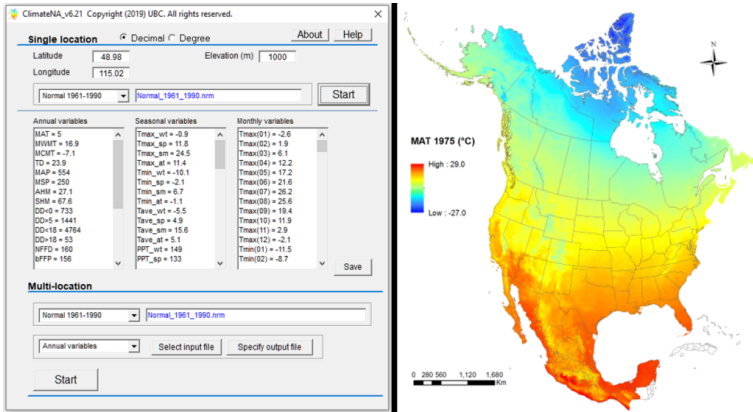


Figure 1. The interfaces and coverages of ClimateNA. Images by Tongli Wang under CC BY 4.0.

ClimateAP covers a major part of the Asia Pacific region (Figure 2).

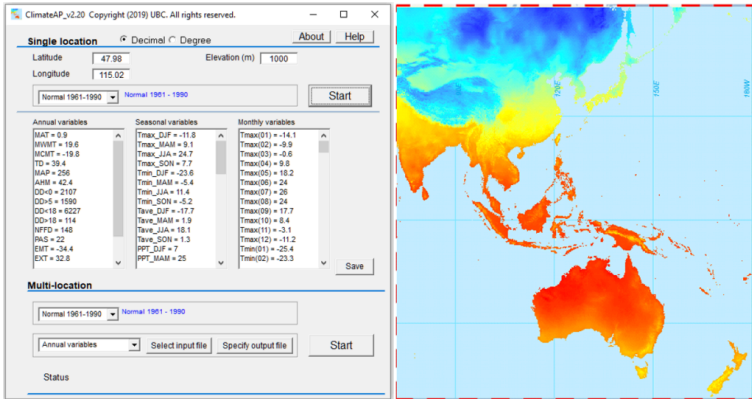


Figure 2. The interfaces and coverages of ClimateAP. Images by Tongli Wang under CC BY 4.0.

ClimateNA and ClimateAP have been widely in climate change-related studies and applications in almost all fields including forestry, forest genecology, plant ecology, hydrology, meteorology, and agriculture. Currently, there are over 1,700 subscribers and a total of over 1,800 citations.

2. Data sources

2.1. Baseline data

The monthly baseline data for 1961-1990 normals were compiled from the following sources and unified at 4 x 4 km spatial resolution:

- British Columbia: PRISM at 800 x 800 m from Pacific Climate Impact Consortium;
- Prairies provinces: PRISM at 4 x 4 km from the PRISM Climate Group

(<http://www.prism.oregonstate.edu/>);

- United States: PRISM at 800 x 800 m from the PRISM Climate Group (Daly et al. 2008);
- The rest: WorldClim (Hijmans et al. 2005);
- The monthly solar radiation data (1948-2010) were provided by Dr. Robbie Hember at the University of British Columbia (to be replaced soon).

2.2. Historical data

Historical monthly data are developed by ourselves for ClimateNA through interpolation of weather station observations from The Global Historical Climatology Network daily ([GHCNd](#)). For ClimateAP, the historical monthly data are from the [Climate Research Unit](#) (CRU ts4.04). Both datasets are provided at the spatial resolution is $0.5 \times 0.5^\circ$ and updated annually.

2.3. Future climate data

The climate data for future periods, including the 2020s (2011-2040), 2050s (2041-70), and 2080s (2071-2100), were from General Circulation Models (GCMs) of the Coupled Model Intercomparison Project (CMIP6) included in the IPCC Sixth Assessment Report (IPCC 2021). Thirteen GCMs were selected for four greenhouse gas emission scenarios, called Shared Socioeconomic Pathways (SSPs), including SSP2-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5. When multiple ensembles are available for each GCM, an average was taken over the available (up to five) ensembles. Ensembles among the

13 GCMs are also available. Time-series of monthly projections are provided for the years between 2011-2100 for each of the selected GCMs and their ensembles.

2.4. Paleoclimate data

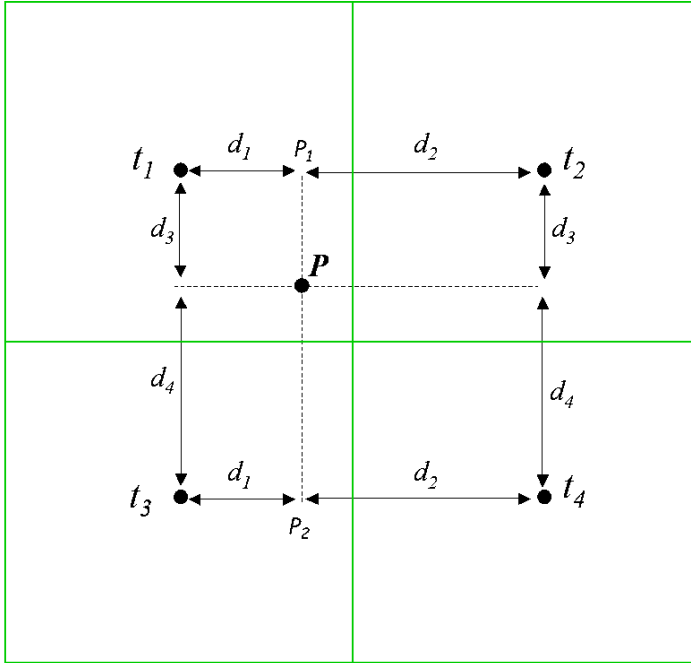
ClimateNA (version 6) (not yet in ClimateAP and ClimateNA version 7) contains monthly paleoclimate data were from four GCMs of the CMIP5 (paleoclimate data are not available yet in CMIP6). The paleo periods covered: 1) Last Millennium or past 1000 years, starting on January 01, 0850; 2) MidHolocene (6,000 years ago); and 3) Last Glacial Maximum (21,000 kyrs ago). Monthly averages were taken over the first 50 years of each period, which is the minimum period available in among the four GCMs.

3. Dynamic local downscaling

3.1. Baseline data

Instead of using the midpoint values of each grid cell to represent all locations in the grid, ClimateNA and ClimateAP use a combination of bilinear interpolation and local elevation adjustment approaches to downscale the baseline monthly grid data (4×4 km) to scale-free point locations. In other words, the programs predict climate data for every specific location.

The program first uses bilinear interpolation to estimate values between midpoints of the four neighbor grids to generate a seamless surface for each monthly climate variables, avoiding step-artifacts at grid boundaries.



$$f(P) = t_1 d_2 d_4 + t_2 d_1 d_4 + t_3 d_2 d_3 + t_4 d_1 d_3$$

Figure 3. Illustration of the bilinear interpolation algorithm. t_1 , t_2 , t_3 , and t_4 are mid-point values of the four neighbouring grids. d_1 , d_2 , d_3 and d_4 are proportions of the bilinear distances to the point of interest (P). Image by Tongli Wang under CC BY 4.0.

The algorithms then retrieve monthly climate data and elevation values for a location from the corresponding grid cell plus eight surrounding cells. The climate and elevation values of the nine cells are used to calculate differences in a climate variable and in elevation between all 36 possible pairs. A simple linear regression of the differences in the climate variable on the difference in elevation is then established, and the slope of the regression is used as the

empirical lapse rate for each climate variable at each specific location. The lapse rates are estimated for all the 48 primary monthly variables in ClimateNA and 36 in ClimateAP.

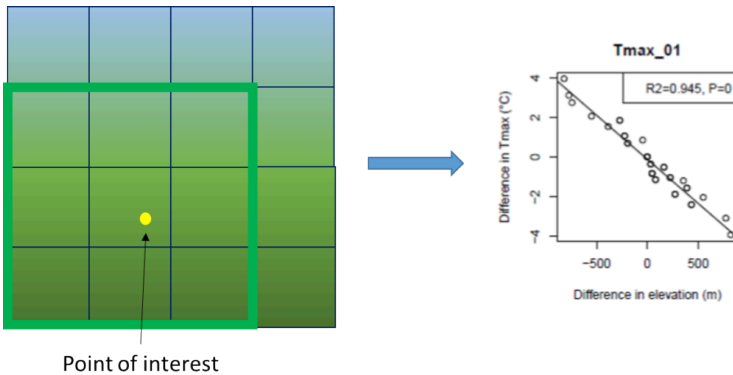


Figure 4. Illustration of the dynamic local regression to estimate a local lapse rate for a climate variable (T_{max_01}). Image by Tongli Wang under CC BY 4.0.

To avoid over-adjustments due to a weak linear relationship, each lapse rate was weighted by the R-square value of the local linear regression. The estimated lapse rate is applied to adjust the bilinearly interpolated values of the location to achieve the downscaling. As the local regressions are dynamically developed along with locations of inquiry, this downscaling method is called a “dynamic local downscaling” approach.

3.2. Historical and Future climate data

Historical, paleo, and future climate data are first converted to anomalies relative to the reference normal period (1961-1990), bilinearly interpolated, then added onto the

downscaled baseline monthly climate normal data (scale-free) to arrive at the final climate surface at a specified resolution or for point data. This approach is called a delta downscaling approach, as illustrated in Figure 5.

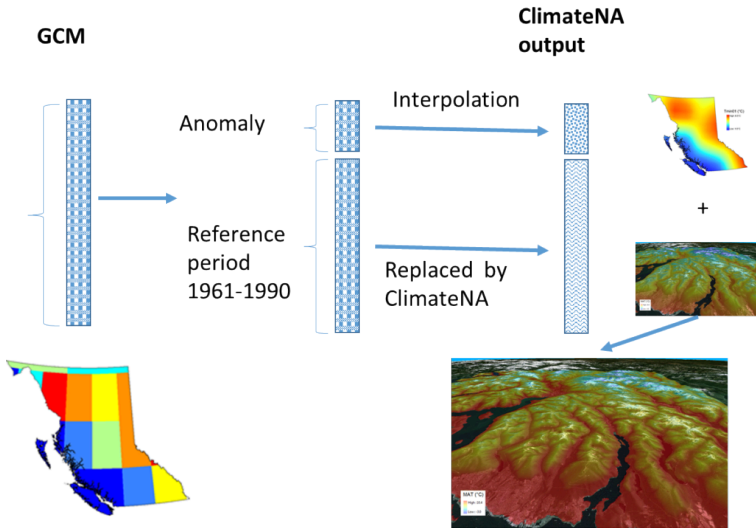


Figure 5. Illustration of a delta approach of downscaling future climate data from GCMs. Image by Tongli Wang under CC BY 4.0.

With this approach, the original baseline portion (i.e., absolute values for the 1961–1990 normal period) of the historical and future climate data were replaced by the scale-free climate data generated by ClimateNA. Because elevation-adjusted baseline data generated by ClimateNA typically has much higher spatial resolution and accuracy than the historical and future climate data, this process improved the accuracies for both historical and future climate data. However, the approach does not change or improve the anomaly portions of the historical and future climate data except the effect of bilinear interpolation.

4. Development of additional climate variables

The baseline dataset contains 36 primary monthly climate variables, including monthly minimum (Tmin) and maximum (Tmax) temperatures, and monthly total precipitation (Prep). ClimateNA and ClimateAP generate over 200 biologically relevant climate variables to facilitate applications in ecology and forest resource management. Of these climate variables, some are directly calculated from the primary monthly climate variables, and others are derived (or estimated) based on observations from weather stations.

4.1. Calculated climate variables

Monthly mean temperatures (Tave) are calculated from monthly Tmin and Tmax. Annual and seasonal mean temperatures and total precipitation are calculated from Tave and Prep, such as mean annual temperature (MAT) and mean annual precipitation (MAP). Other often-used annual climate variables include mean warmest month temperature (MWMT), mean coldest month temperature (MCMT), continentality (TD), which is the difference between MWMT and MCMT, and mean summer precipitation (MSP), ect.. A full list of calculated climate variables can be found at <http://climatena.ca/>

4.2. Derived climate variables

Many of these additional variables need to be calculated using daily climate data, which are not available in

ClimateNA. We estimated these variables based on empirical or mechanistic relationships between these variables calculated using daily observations and monthly climate variables from weather stations from entire North America for ClimateNA, and from the Asia Pacific for ClimateAP. We called these variables “derived climate variables”. These variables were developed at the monthly scale, then summed up to seasonal and annual scales. The steps included: 1) calculating derived climate variables for each month (e.g., degree days) from daily weather station data; 2) building relationships (or functions) between the derived climate variables and observed (or calculated) monthly climate variables; 3) applying the functions in ClimateNA and ClimateAP to estimate derived climate variables using monthly climate variables.

The widely used derived climate variables include growing degree-days ($DD > 5$), cooling degree-days ($DD < 0$), frost-free days (NFFD), frost-free period (FFP), extreme minimum (EMT) and maximum (EXT) temperatures, and climatic moisture deficit (CMD). A full list of derived climate variables can be found at <http://climatena.ca/>

The derived climate variables are at high accuracies based on the relationships between derived climate variables and monthly temperatures, as shown in Figure 6, for example. Please check [Wang et al. \(2016\)](#) for details.

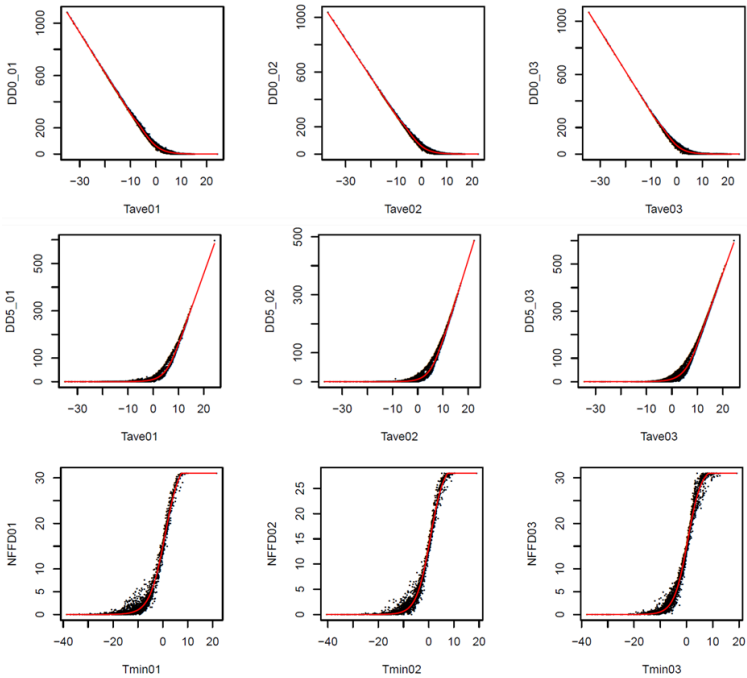


Figure 6. Relationships between estimated climate variables, including cooling degree-days (DD0), growing degree-days (DD5), and number of frost-free days (NFFD), and monthly temperatures. Image by Tongli Wang under CC BY 4.0.

5. Main features

5.1. Scale-free climate data

ClimateNA and ClimateAP generate scale-free climate data that facilitates users to obtain climate data for specific locations, instead of getting grid averages from other data sources. Thus, the scale-free climate data can better reflect the climatic conditions of the locations of interest, as illustrated in Figure 7.

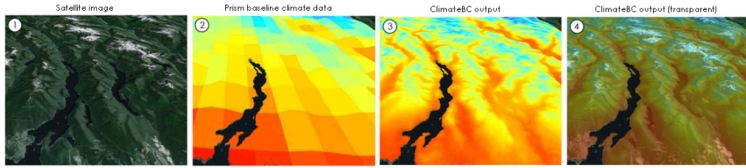


Figure 7. Comparisons between scale-free and gridded (PRISM) climate data in term of reflecting the effect of topography on climate variation. Image by Tongli Wang under CC BY 4.0.

The scale-free feature considerably improves the accuracy of climate variables. For example, ClimateBC (a subset of ClimateNA) predictions of monthly maximum temperatures have much stronger relationships with observations from weather stations (Figure 8).

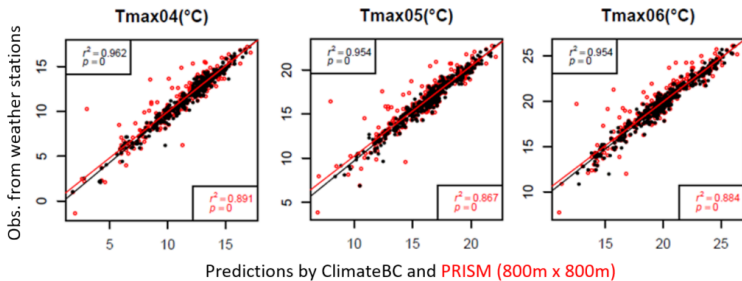


Figure 8. Comparisons of scale-free (ClimateBC) and gridded (PRISM) climate data against observations from weather stations. Image by Tongli Wang under CC BY 4.0.

Scale-free climate data improve the prediction accuracy of ecological models. Climate variables of ecological sample plots predicted by ClimateNA or ClimateAP can better represent climatic conditions of the specific locations of the sample plots, thus can help to build more accurate ecological models. Figure 9 demonstrates an evident improvement in the prediction accuracy of an ecological

niche model modeling the biogeoclimatic zones in British Columbia.

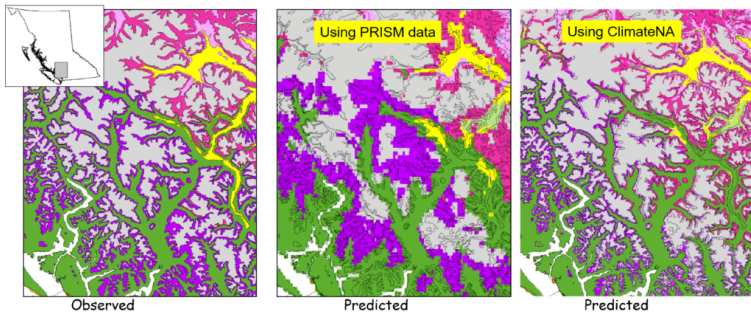


Figure 9. The effect of scale-free climate data on prediction accuracy of an ecological niche model predicting the biogeoclimatic zones in British Columbia. Image by Tongli Wang under CC BY 4.0.

ClimateNA and ClimateAP can generate spatial climate data surfaces at any spatial resolution. The output spatial resolution of the climate variables is determined by user's input dataset. This is particularly important for regional or local applications of ecological models, as a finer spatial resolution is often required to reflect the situation meaningful for practical applications (Figure 10).

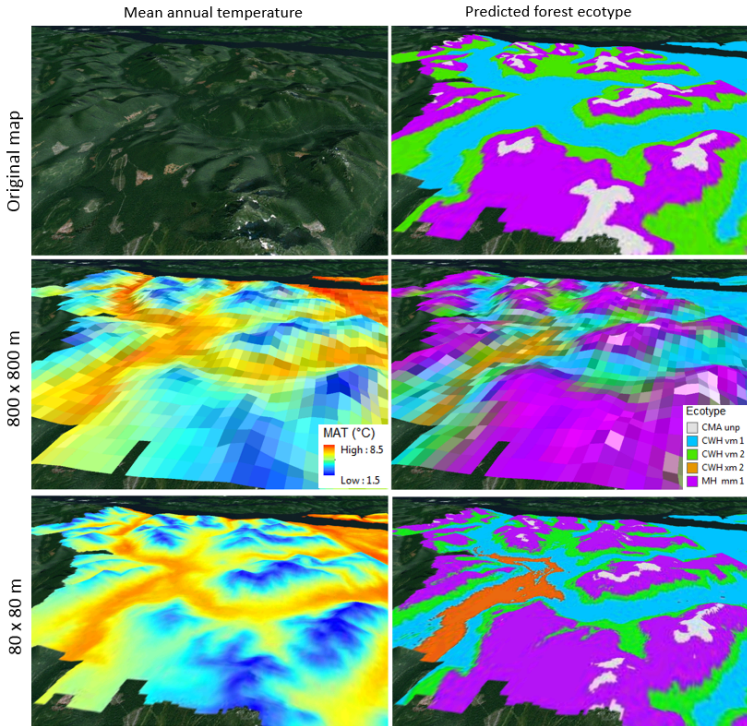


Figure 10. Illustration of climate data at a finer spatial resolution necessary for the prediction of forest ecotypes at a local scale. Image from Wang et al. (2016) in PLOS One under open access.

5.2. All-in-one package

ClimateNA and ClimateAP integrate and downscale climate data for historical and Future years and periods. ClimateNA also integrates and downscales climate data for paleo periods (Figure 11). This prevents users from the need to obtain climate data from different sources with different climate variables with different data formats.

5.5. Map-in and map-out capacity

Since version 6.0, ClimateNA can read DEM raster file and generate climate variables in raster format for mapping. In previous versions, users need to convert a digital elevation model (DEM) to a spreadsheet file (i.e., a CSV file) as an input file, and also need to convert the output file to raster files. The new versions can directly read a DEM file in raster format and generate climate variables in raster format (as illustrated in Figure 13), which can be directly imported to a mapping software including R.

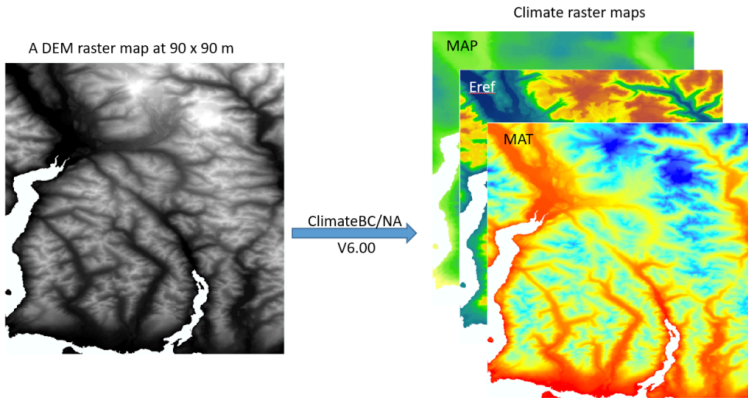


Figure 13. Illustration of ClimateNA directly reading a DEM raster file and generating climate variables in raster file format. Image by Tongli Wang under CC BY 4.0.

6. Web versions

Web versions of [ClimateNA](#) and [ClimateAP](#) are available to facilitate easy access to climate data for users dealing with a small number of locations. Users just need to click at the locations on Google Maps to obtain climate variables

without the need to type in latitude, longitude, and elevation of the locations as they are automatically retrieved through the Google Map Services. The obtained climate data can be saved to a local computer.

The web versions also allow users to visualize the spatial patterns of major climate variables. Users can use Google Maps' features to navigate and to zoom in or out around the locations of interest. The maps of the major climate variables can be downloaded in a picture or raster format.

In addition, the web versions serve as a platform to host the spatial data of forest ecosystems and species distributions under current and future climates. This feature has been widely used as a basis for foresters to develop forest adaptive strategies and as demonstrations of the impact of climate change on ecosystems and forest tree species in high school and university classes.

The web version of ClimateNA has two subsets, [ClimateBC](#) and [ClimateWNA](#), to focus on regional interests (Figure 14). ClimateBC focuses on British Columbia. In addition to climate data, it hosts forest ecosystems and species distributions. ClimateWNA hosts several major forest tree species in western North America.

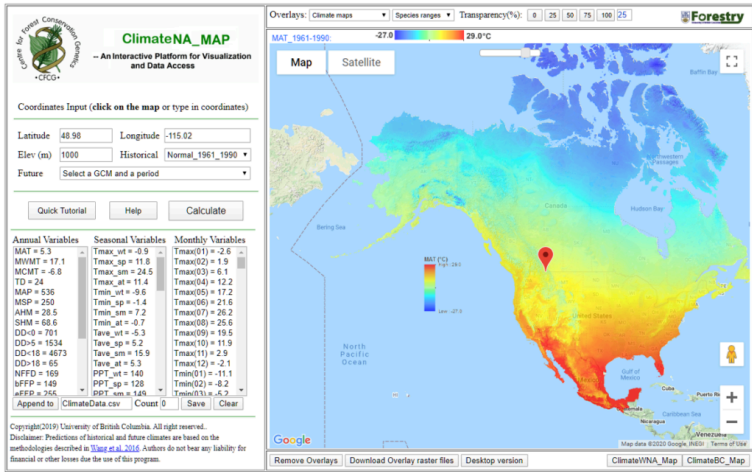


Figure 14. The web version interface of ClimateNA with links to ClimateBC and ClimateWNA. Image by Tongli Wang under CC BY 4.0.

The web version of ClimateAP has five subsets, ClimateChina, ClimateLaos, ClimateMalaysia, ClimateMyanmar, and ClimateTaiwan, for different economies (Figure 13). Each hosts climate surfaces and ecosystem projections of its own region.

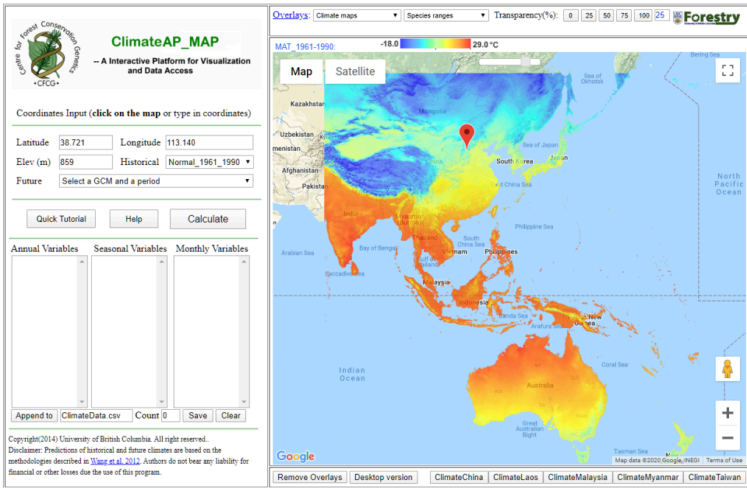


Figure 15. The web version interface of ClimateAP with links to ClimateChina, ClimateLaos, ClimateMalaysia, ClimateMyanmar, and ClimateTaiwan. Image by Tongli Wang under CC BY 4.0.

7. Spatial data

Pre-generated spatial climate data in raster format for western North America (800m x 800m) and the entire North America (4000m x 4000m) are available at <https://climatena.ca/spatialData>

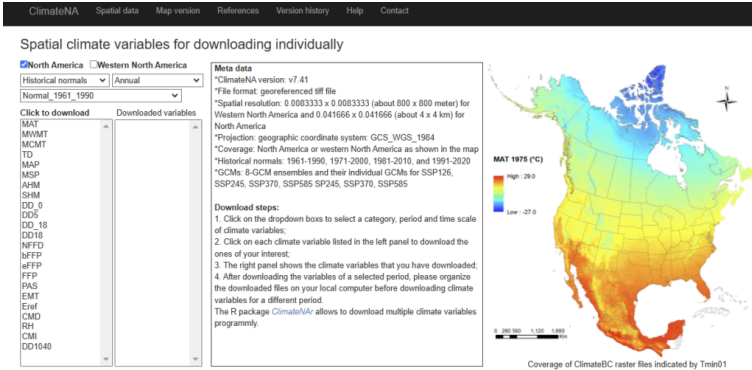


Figure 14. Screenshot of ClimateNA spatial data portal.

Pre-generated spatial climate data in raster format for entire North America. The datasets were produced by [AdpatWest](#), which includes climate data for reference and future periods with many GCMs.

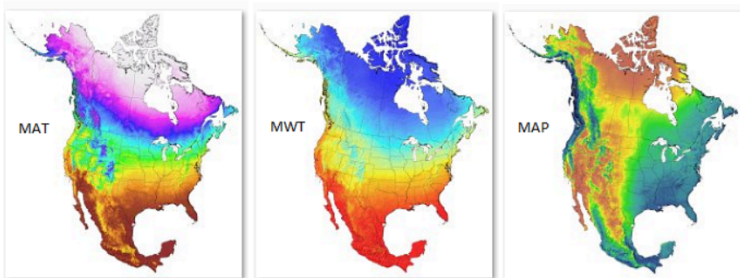


Figure 15. Spatial climate variables generated by AdaptWest using ClimateNA for entire North America. Image by AdaptWest (in a public domain).

8. ClimateNAr package

8.1 Main features

ClimateNAr (v2.0.0 or later) includes the R version of

ClimateNA, *ClimateNAr*. It is a significant step in developing the functionality of ClimateNA. It has the following features:

- Full integration into a programming workflow in R.
- Independent of the desktop version – working on MacOS and Linux.
- Using an external input file or an internal dataframe.
- Using DEM raster files in any format: *.tif, *.asc, or GIS grid.
- Output raster files are in *.tif format, which is much smaller and efficient.
- Output climate variables customizable: Choose any combination of climate variables.
- Smaller package and faster processing speed: 165 MB in total, 1/10 of the desktop package, and Processing speed is 10 times faster.

8.2



By Tongli Wang, Climate Modelling and Forest Applications is licensed under [Creative Commons <CC BY-NC-ND 4.0> License](https://creativecommons.org/licenses/by-nc-nd/4.0/)

Topic 3.2. Using ClimateNA/AP to Generate Point and Spatial Climate Data

Tongli Wang



Note: Tutorial videos are embedded within the related content below. Thus, no additional videos are provided here.

Contents

- [1. Download and installation](#)
 - [1.1. Download](#)
 - [1.2. Installation](#)
- [2. Using the interactive mode](#)
 - [2.1. Coordinates input](#)
 - [2.2. Period selection](#)
 - [2.3. Climate data generating and saving](#)
- [3. Using multi-location mode](#)
 - [3.1. Input file preparation](#)
 - [3.2. Multi-location processing](#)

- [4. Multiple GCMs and RCPs processing](#)
- [5. Generating spatial climate data](#)
 - [5.1. Installing R packages](#)
 - [5.2. Preparing an input file from a DEM file](#)
 - [5.3. Using a DEM as an input file](#)
- [6. Using time-series functions](#)
 - [6.1. Historical Time Series](#)
 - [6.2. Future time series](#)
 - [6.3. Time-series raster files](#)
- [7. Using the web version](#)
 - [7.1. Easy access to climate data](#)
 - [7.2. Visualizing spatial climate variables](#)
 - [7.3. Visualizing ecological niche maps of tree species and forest ecotypes](#)
- [8. Using pre-generated spatial data](#)
- [9. Using the R version of ClimateNA](#)
 - [9.1 Download and installation](#)
 - [9.2 Input parameters](#)
 - [9.3 Return](#)
 - [9.4 Examples](#)

1. Download and installation

1.1. Download

ClimateNA and ClimateAP are free for download from their websites at

<http://climatena.ca/> and <http://web.climateap.net/>

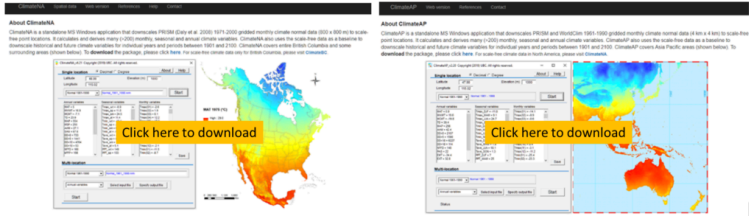


Figure 1. Screenshots of ClimateNA and ClimateAP websites.
Image by Tongli Wang under CC BY 4.0.

By clicking on the interface of each model, you will be led to the downloading page. To download, users are required to register to a mailing list first, which is only for informing new releases and important announcements. After the registration, a downloading link will appear on the page and also be sent to your email. The packages are compressed zip files with sizes of about 1.20 GB for ClimateNA (version 7) and 300 MB for ClimateAP.

1.2. Installation

No installation is required. Users just need to simply unzip all the files and subfolders into a folder on your hard disk and double-click the file ClimateNA _v7.**.exe". The program does not work on network drives for some unknown reasons.

The exe file can be started even without uncompress the package. However, it does not work properly, and an error message will appear if the "Start" button is clicked.

The program runs on all Windows.

A Youtube video is available to demonstrate the steps of downloading and installation.



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://pressbooks.bccampus.ca/fode010notebook/?p=39#oembed-1>

Video 1. A tutorial for downloading and installation of ClimateNA (or ClimateAP). Video by Tongli Wang under CC BY 4.0.

[Back to top](#)

2. Using the interactive mode

2.1. Coordinates input

The upper section of the interface is for processing single locations. Users need to manually type in the latitude, longitude, and elevation of a location of interest.

ClimateNA_v7.10 Copyright (2021) UBC. All rights reserved. X

Single location Decimal Degree About Help

Latitude Elevation (m)

Longitude

Start

Annual variables	Seasonal variables	Monthly variables
MAT = 5	Tmax_wt = -0.9	Tmax(01) = -2.6
MVMT = 16.9	Tmax_sp = 11.8	Tmax(02) = 1.9
MCMT = -7.1	Tmax_sm = 24.5	Tmax(03) = 6.1
TD = 23.9	Tmax_at = 11.4	Tmax(04) = 12.2
MAP = 554	Tmin_wt = -10.1	Tmax(05) = 17.2
MSP = 250	Tmin_sp = -2.1	Tmax(06) = 21.6
AHM = 27.1	Tmin_sm = 6.7	Tmax(07) = 26.2
SHM = 67.6	Tmin_at = -1.1	Tmax(08) = 25.6
DD<0 = 733	Tave_wt = -5.5	Tmax(09) = 19.4
DD>5 = 1441	Tave_sp = 4.9	Tmax(10) = 11.9
DD<18 = 4764	Tave_sm = 15.6	Tmax(11) = 2.9
DD>18 = 53	Tave_at = 5.1	Tmax(12) = -2.1
NFFD = 160	PPT_wt = 149	Tmin(01) = -11.5
bFFP = 156	PPT_sp = 133	Tmin(02) = -8.7

Save

Multi-location

Figure 2. Interface of ClimateNA (about the same for ClimateAP). Image by Tongli Wang under CC BY 4.0.

The latitude and longitude can be in decimal degrees with the default option. They can also be in degree, minute and second if the “Degree” option is selected. Elevation is always in meters. If elevation is missing (or not available), the program still generates climate data based on the bilinearly interpolated elevation values of the baseline dataset, thus, elevation adjustment is not applied in that case.

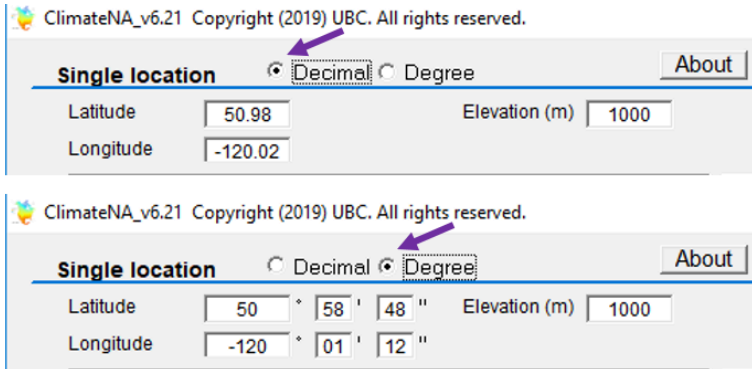


Figure 3. Two options for lat and lon input using ClimateNA (the same for ClimateAP). Image by Tongli Wang under CC BY 4.0.

2.2. Period selection

The default period in both ClimateNA and ClimateAP is the normal period 1961-1990, which is also the period defined as a standard reference period for long-term climate change assessments by the World Meteorological Organization.

More normal period – selection for more normal periods is available by selecting “More Normal Data”. There are nine normal periods available over the period of 1901-2019.

Decadal data – a decadal period can be selected in the same as for selecting a normal period. There are 12 decadal periods available including an incomplete decadal period for the current decade (2011-2018).

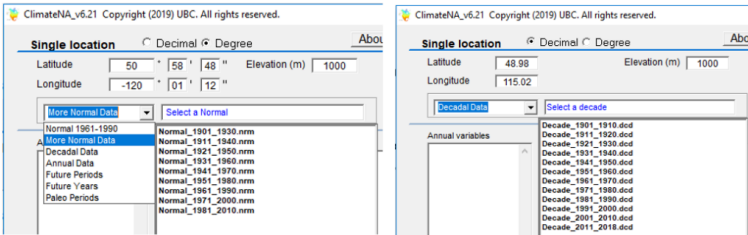


Figure 4. Selection of a normal period (left) or a decadal period (right). Image by Tongli Wang under CC BY 4.0.

Annual data – to generate climate data for a specific year, “Annual Data” needs to be selected in the period dropdown menu. A list of all years will appear for the period of 1901-2019, and an individual year can then be selected.

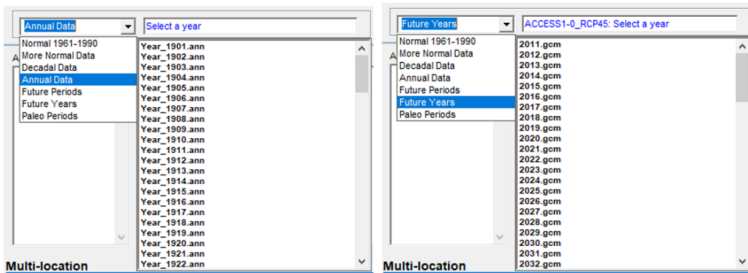


Figure 5. Selection of an individual year for the historical period of 1901-2018 (left) and the future period of 2011-2100 (right). Image by Tongli Wang under CC BY 4.0.

Future periods – A General Circulation Model (GCM), a Representative Concentration Pathway (RCP), and a future period can be selected under “Future Periods”. There are 15 GCMs for the greenhouse gas emission scenarios RCP4.5 and RCP8.5, and 13 GCMs for RCP2.6. There is also ensemble for each RCP and period over all the GCMs. These ensembles are widely used in climatic and ecological modelling applications.

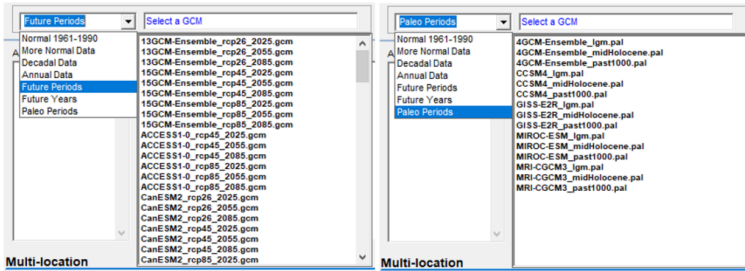


Figure 6. Selection of a GCM and a future period (left) or a GCM and paleo period (right). Image by Tongli Wang under CC BY 4.0.

2.3. Climate data generating and saving

After a period is selected as mentioned above, it is ready to generate climate data by clicking on the “Start” button. Data generated for annual, seasonal and monthly climate variables are shown in Figure 7. The data values and the coordinates and elevation can be saved as a CSV file by clicking the Save button.

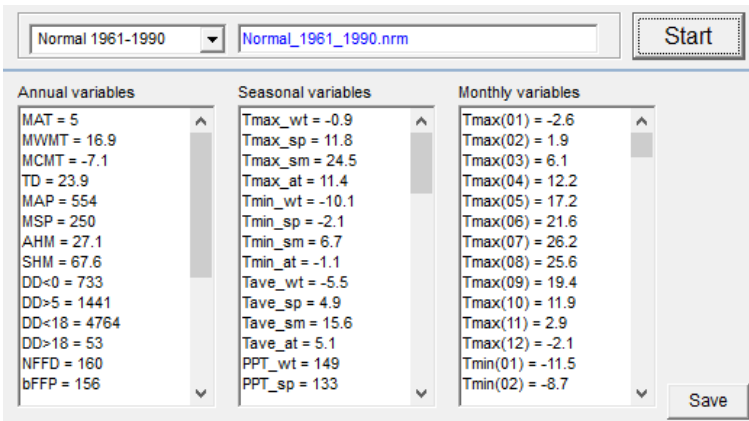


Figure 7. Data generated for annual, seasonal and monthly climate variables. Image by Tongli Wang under CC BY 4.0.

[Back to top](#)

3. Using multi-location mode

The lower section of the interface of the programs is for processing multiple locations. It requires users to prepare an input file that comprises identifications, latitude, longitude and elevation of the locations of interest. The output file keeps all the columns of the input file, but attach climate variables to each location (or each row). Since version 6.0, ClimateNA can also directly read a digital elevation model (DEM) and generate output files in raster format for mapping. The multi-location mode allows time-series functions to generate climate data for multiple years and processing multiple GCMs, RCPs and periods in a single run.

3.1. Input file preparation

An input file can be prepared either in Microsoft Excel, a text editor or R. The file must comprise five columns with the first two columns for IDs (e.g., “ID” and “Site”) and three columns for latitude (lat), longitude (long), elevation (el) as shown in the example below. The columns must be in the same order as shown. If elevation is not available, you have to put in “.” Instead and no elevation adjustment will be applied to that location. No additional columns are allowed. The input file can have almost unlimited number of locations (or rows).

If the file is prepared in Excel, it must be save as a comma-delimited spreadsheet file (a CSV file). If the file is prepared in a text editor, the values must be separated by comma (“,”). If there is a missing value, you need to enter a “.” between two commas. If prepared in R, it need to be save using “write.csv” command.

The image shows two windows. The left window is Microsoft Excel with a spreadsheet containing the following data:

	A	B	C	D	E
1	ID1	ID2	lat	long	el
2	site1	region1	46.98	-113.02	1000
3	site1	region2	51.25	-115.55	550
4					

The right window is Notepad showing the text content of the file:

```
ID1,ID2,lat,long,el
site1,region1,46.98,-113.02,1000
site1,region2,51.25,-115.55,550
```

Figure 8. The format of Data generated for annual, seasonal and monthly climate variables. Image by Tongli Wang under CC BY 4.0.

For making climate maps, an input file can also be converted from a digital elevation model (DEM). This will be explained later in this lecture.

3.2. Multi-location processing

After selecting a period and climate variables, you can select your input file, for example, “Input_test.csv”. Then, click the “Open” button.

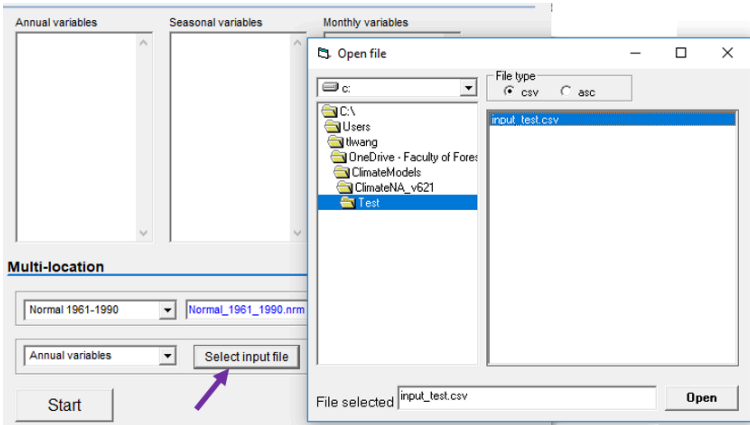


Figure 9. Steps for selecting an input file. Image by Tongli Wang under CC BY 4.0.

The next step is to specify an output file by clicking on the “Specify output file” button. A default output file name is provided, which contains all the information about the input file, period and climate variable category. You may not need to change it in most cases. Then, it is ready to hit the “Start” button to run the process.

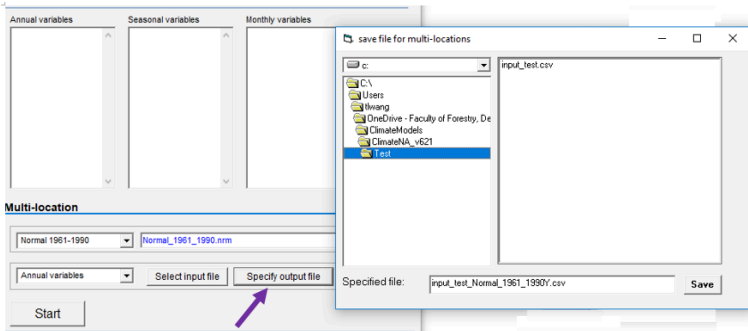


Figure 10. Steps for specifying an output file. Image by Tongli Wang under CC BY 4.0.

A Youtube tutorial shows all the steps and more.



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://pressbooks.bccampus.ca/fode010notebook/?p=39#oembed-2>

[Back to top](#)

4. Multiple GCMs and scenarios processing

In multi-location mode, climate data can be generated for multiple GCMs, SSPs and periods in a single run. Here are the steps:

- 1) Click on “Future Periods” in period selection dropdown box, and a list of GCMs, SSPs and periods will appear on the left.
- 2) Click on the items in the list to make selections, and selected items will appear in the list on the right. If an item is selected by mistake, it can be removed from the list by clicking on it.
- 3) After the selection is completed, click on the “Finish Selection” button.
- 4) Select climate a variable category, open an input file and specify an output file name.

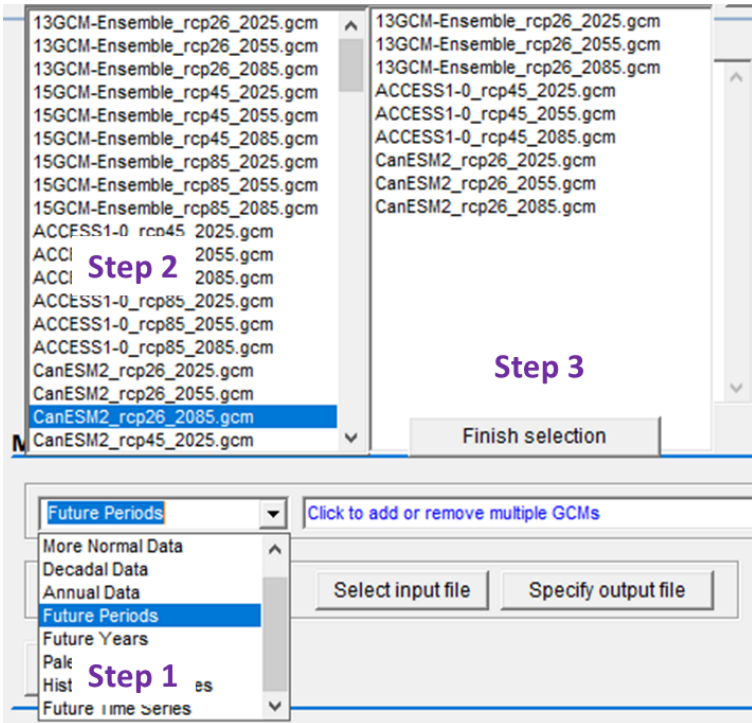


Figure 11. Steps for selecting GCMs, RCPs and periods. Image by Tongli Wang under CC BY 4.0.

All the climate data are generated in one output file. A column is added on the left to show the information of the GCM, scenarios and period.

	A	B	C	D	E	F	G	H	I
1	GCM	id1	id2	Latitude	Longitude	Elevation	MAT	MWMT	MCMT
2	13GCM-Ensemble_rcp26_2025.gcm	1	site1	49.5	-126.5	250	9.3	16	3.6
3	13GCM-Ensemble_rcp26_2025.gcm	2	site2	50.84259	-127.3	1000	9.8	16.1	4.4
4	13GCM-Ensemble_rcp26_2025.gcm	3	site3	55	-126.815	150	5.7	17.6	-9.2
5	13GCM-Ensemble_rcp26_2055.gcm	1	site1	49.5	-126.5	250	9.9	16.6	4.2
6	13GCM-Ensemble_rcp26_2055.gcm	2	site2	50.84259	-127.3	1000	10.4	16.7	5.1
7	13GCM-Ensemble_rcp26_2055.gcm	3	site3	55	-126.815	150	6.4	18.2	-8.3
8	13GCM-Ensemble_rcp26_2085.gcm	1	site1	49.5	-126.5	250	10	16.7	4.5
9	13GCM-Ensemble_rcp26_2085.gcm	2	site2	50.84259	-127.3	1000	10.5	16.7	5.4
10	13GCM-Ensemble_rcp26_2085.gcm	3	site3	55	-126.815	150	6.5	18.1	-8
11	ACCESS1-0_rcp45_2025.gcm	1	site1	49.5	-126.5	250	9.3	16	3.6
12	ACCESS1-0_rcp45_2025.gcm	2	site2	50.84259	-127.3	1000	9.7	16	4.5
13	ACCESS1-0_rcp45_2025.gcm	3	site3	55	-126.815	150	5.7	17.9	-8.1
14	ACCESS1-0_rcp45_2055.gcm	1	site1	49.5	-126.5	250	10.3	16.9	4.6
15	ACCESS1-0_rcp45_2055.gcm	2	site2	50.84259	-127.3	1000	10.8	16.8	5.6
16	ACCESS1-0_rcp45_2055.gcm	3	site3	55	-126.815	150	6.7	18.3	-8.2
17	ACCESS1-0_rcp45_2085.gcm	1	site1	49.5	-126.5	250	11	17.7	5.2
18	ACCESS1-0_rcp45_2085.gcm	2	site2	50.84259	-127.3	1000	11.4	17.8	6.2
19	ACCESS1-0_rcp45_2085.gcm	3	site3	55	-126.815	150	7.6	20.1	-7

Figure 12. Output of multiple GCMs, RCPs and periods. Image by Tongli Wang under CC BY 4.0.

All the steps are also shown in a Youtube:



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://pressbooks.bccampus.ca/fode010notebook/?p=39#oembed-3>

[Back to top](#)

5. Generating spatial climate data

ClimateNA and ClimateAP can generate spatial climate data at any spatial resolution. A DEM file is required either for preparation of an input file for ClimateAP or directly used as an input file for ClimateNA. The spatial resolution

of the DEM file determines the spatial resolution of the output raster files.

5.1. Installing R packages

We need R to prepare input files and to process output data. Here we will show the steps to install R on Windows, and to install and load R packages. As we will provide the code for each processing work, no programming is required for this lecture. Here are the steps:

- 1) Download R: <https://mirror.rcg.sfu.ca/mirror/CRAN/bin/windows/base/R-4.3.3-win.exe>

- 2) Click to run the exe file “R-4.3.2-win.exe. Follow the instruction on the screen to complete the installation.

- 3) Click the newly installed R-4.3.2 to start R.

- 4) Click “File”, then “New script” to open the R code editor

- 5) Type “install.packages(‘terra’)” and run it by right click and select “Run line or selection”, and follow the instruction on the screen to complete the package installation.

Now you are ready to run the code provided later on.

5.2. Preparing an input file from a DEM file

DEM files can be obtained from various sources. For example, free Digital Elevation Models can be downloaded at https://www.opendem.info/link_dem.html. However, to get a DEM for a specific area is not readily available. The following Youtube tutorial video shows you how.



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://pressbooks.bccampus.ca/fode010notebook/?p=39#oembed-4>

For ClimateAP, an additional step is required. to convert a DEM file into an input file in R. The following code is used to convert the DEM file “bc80k.asc” in ascii raster format (or “bc80k.tif” in tiff format) to an input file in CSV format (“bc80k.csv”) using *terra* package:

```
library(terra);
setwd('your working folder')
r <- rast('c80k.asc');
xyv <- as.data.frame(r,xy=T); head(xyv)
xyv <-
data.frame(ID1=row(xyv)[,1],ID2=row(xyv)[,1],xy
v[,c(2,1,3)]); head(xyv)
names(xyv)[3:5] <- c('lat', 'lon', 'elev');head(xyv)
write.csv(xyv,'c80k.csv',row.names=F)
```

A DEM file can be converted into an input file either in ArcGIS Map, Python or other programs as long as the output file meets the requirements as mentioned in 3.1.

ClimateAP (or ClimateNA) can use this input file to generate climate data in CSV format as shown in 3.2. To convert the CSV file (e.g., “bc80k_1975Y.csv”) into spatial layers, the following code can be used in R:

```
library(terra);
setwd(' your working folder')
csv <- read.csv('bc80k_1975Y.csv');
rr <- rast(csv[,c(4,3,6:9)], type='xyz'); rr
```

plot(rr); #plot all variables

Plot(rr,1:2); # plot the first two variables

5.3. Using a DEM as an input file

Since version 6.0, ClimateNA can directly read a DEM raster file as an input file and generate climate variables in raster format for mapping. This makes it much easier to generate spatial climate variables. Here are the steps:

1. Have a DEM raster file (*.asc) for the area of interest at the resolution you want;
2. Click the “Select input file” button and a file open box will show up;
3. Navigate to the folder containing the raster file;
4. Select the File type: “asc” as shown in the following screenshot and only asc raster files will show up:

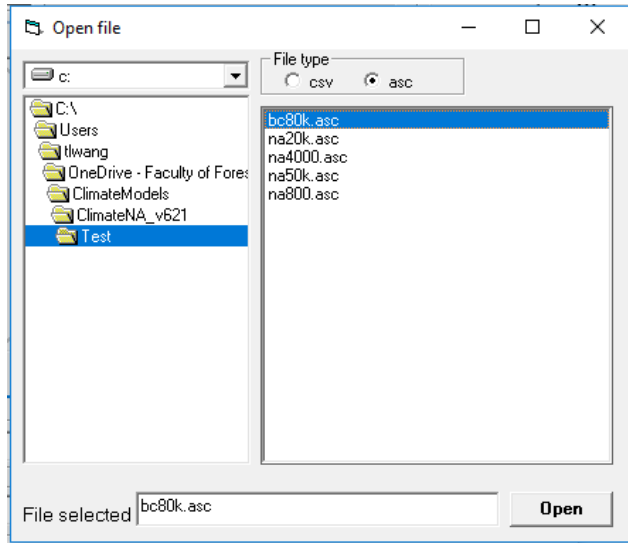


Figure 13. A screenshot of the file-open box to illustrate the selection of a raster input file. Image by Tongli Wang under CC BY 4.0.

5. Select an asc input file and click “Open” button;
6. Click the “Specify output file” button, and a default folder will show up

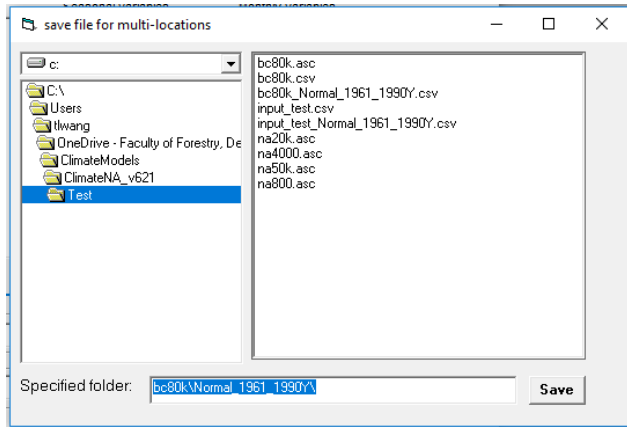


Figure 14. A screenshot of file-save box for specifying an output folder. Image by Tongli Wang under CC BY 4.0.

7. Click the “Save” button;
8. Click the “Start” button, and climate variables in ASCII raster format will be generated;
9. Projection: geographic coordinate system, GCS_WGS_1984.
10. Drag and drop an output file into ArcGIS to show the map, or display the map in R using the following code:

```
library(raster);
setwd(' your working folder')
mat <- raster('mat.asc');
plot(mat)
```

A Youtube video shows the steps:





One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://pressbooks.bccampus.ca/fode010notebook/?p=39#oembed-5>

[Back to top](#)

6. Using time-series functions

ClimateNA and ClimateAP can generate time-series climate data for historical (1901-2018) or future years (2011-2100) for multiple locations in a single run. The time-series functions work only for the multi-location process.

6.1. Historical Time Series

Here are the steps to generate historical time-series climate data:

- 1) Select “Historical Time Series” in the period selection dropdown box;
- 2) Select a variable category (monthly, seasonal, annual or all variables);
- 3) Input the starting and ending years in the pop-up boxes;
- 4) Specify input and output files, and click the

“Calculate TS” button.

The climate data for all the years involved will be saved in a single CSV file. The information of years is added as the first column to the output file. The size of the file can be huge.

6.2. Future time series

The steps are slightly different for generating time-series data for a future period as it involves selection of GCM, RCP and GCM model runs. Here are the steps:

- 1) Select “Future Time Series” in the period selection dropdown box;

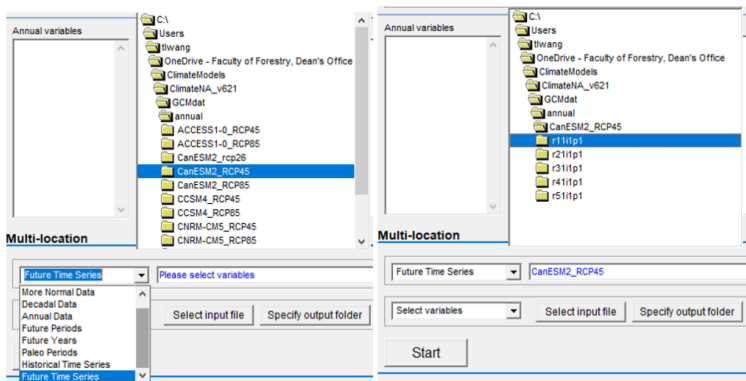


Figure 15. Illustration of selecting a GCM, RCP and model run for a future time-series. Image by Tongli Wang under CC BY 4.0.

- 2) Select a variable category (monthly, seasonal, annual or all variables);
- 3) Input the starting and ending years in the pop-up boxes;

- 4) Specify input and output files, and click the “Calculate TS” button.

6.3. Time-series raster files

The two time-series functions also work for generating spatial raster files. The only differences in the steps include:

- 1) Select an ASCII raster input file as in 5.2;
- 2) Specify an output folder instead of an output file. Although the default folder works in most cases, this step cannot be skipped.

The following Youtube video shows all the steps mentioned above.



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://pressbooks.bccampus.ca/fode010notebook/?p=39#oembed-6>

[Back to top](#)

7. Using the web version

The web versions of ClimateNA and ClimateAP can be accessed at: <https://climatena.ca/mapVersion>

<http://climateap.net/default.aspx>

7.1. Easy access to climate data

Different from the desktop versions, typing in coordinates is optional as they can be automatically obtained by clicking on the Google Map. There are two dropdown boxes for selecting either a historical period or years, or a future period. There are only three GCMs available in contrast to 15 GCMs in the desktop versions. Here are the steps:

- 1) Click at the location of interest on the Google Map and the lat, lon and elve will be automatically obtained through Google Service. Alternatively, lat, lon and elve can be typed in.
- 2) Select either a historical or a future period;
- 3) Click the “Calculate” button to generate climate data (this step can be skipped);
- 4) Click the “Append” button to save the climate data to the temperate memory;

Get climate data for other locations or periods by repeating the steps 1) – 4);

- 5) Click the “Save” button to save the data to your local computer.

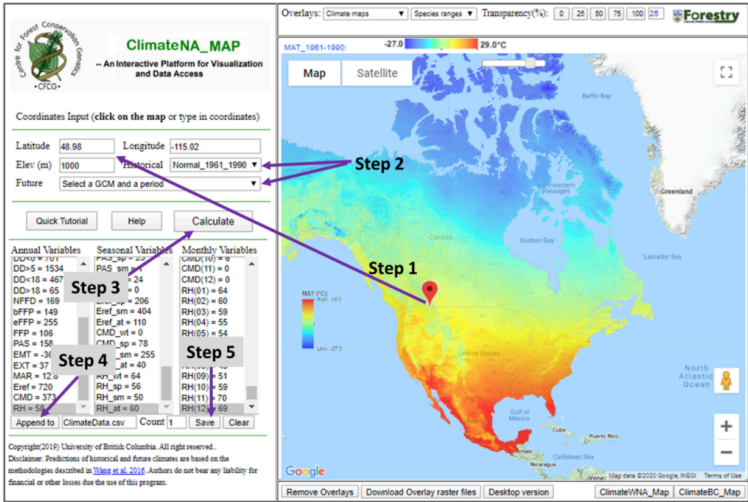


Figure 16. Steps of using the web version of ClimateNA. Image by Tongli Wang under CC BY 4.0.

7.2. Visualizing spatial climate variables

Several major climate variables are overlaid on the Google Maps and they can be visualized using the Google Map features. When a climate variable is shown, it can be downloaded as a raster file to your local computer by clicking on “Download overlay raster files”.

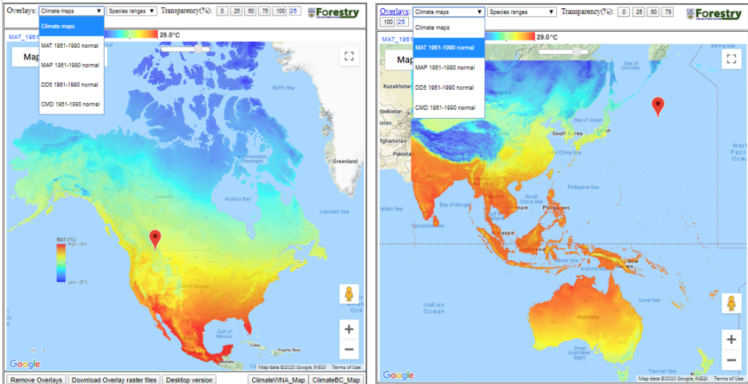


Figure 17. Visualization of a climate map in ClimateNA (left) and ClimateAP (right). Image by Tongli Wang under CC BY 4.0.

7.3. Visualizing ecological niche maps of tree species and forest ecotypes

Ecological niche maps of some major forest tree species are also overlaid onto the Google Maps in ClimateNA and ClimateAP, and their subprograms, such as ClimateBC_Map and ClimateChina.

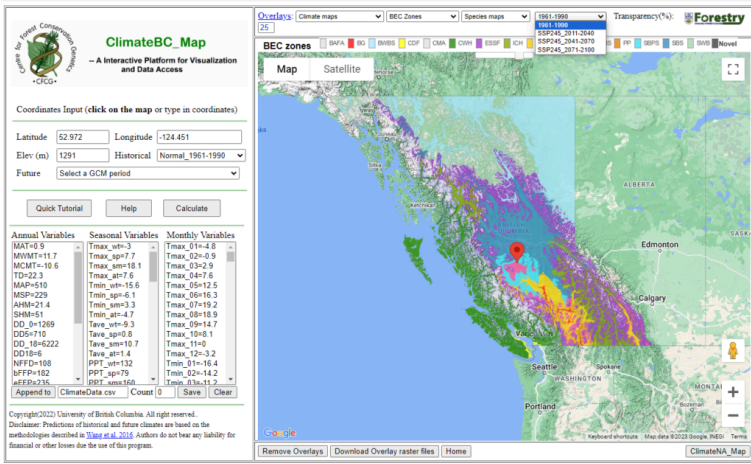


Figure 18a. Biogeoclimatic ecosystem classifications

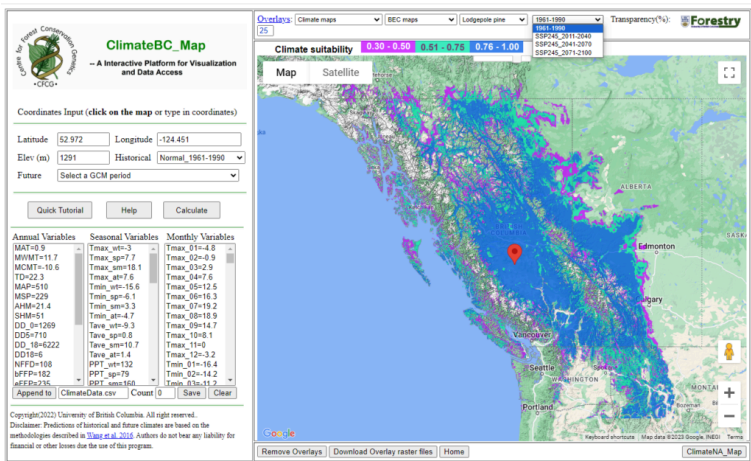


Figure 18b. Climate niche of lodgepole pine in BC

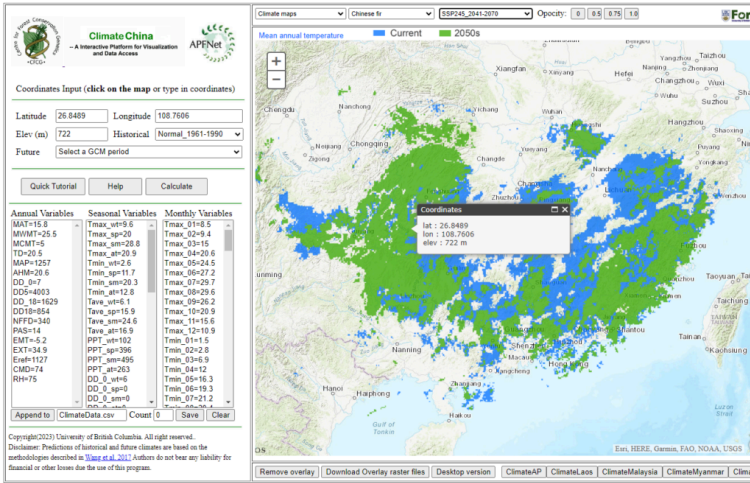


Figure 18c. Climate niche of Chinese fir

[Back to top](#)

8. Using pre-generated spatial data

A large volume of pre-generated spatial climate data using ClimateNA is available for download. However, such spatial data are not available yet for ClimateAP.

Spatial climate data can be downloaded individually or in groups for ClimateBC (Figure 19). A similar feature will be available for ClimateNA soon.

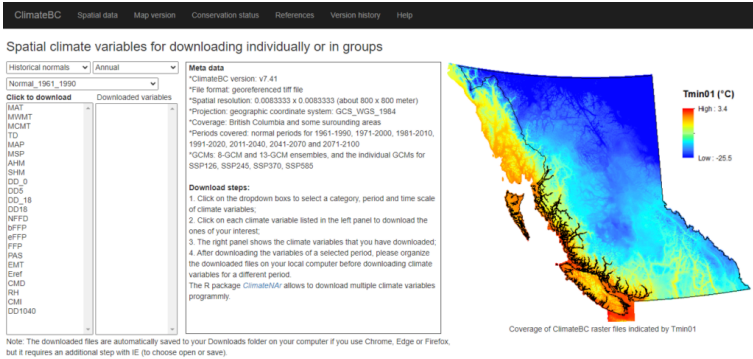


Figure 19a. ClimateBC pre-generated raster files can be downloaded

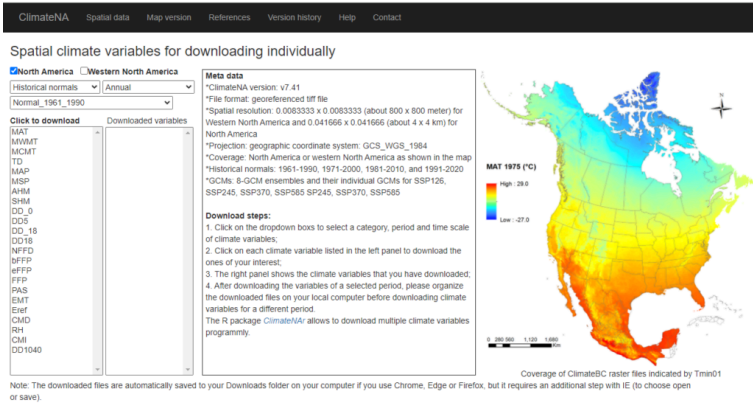
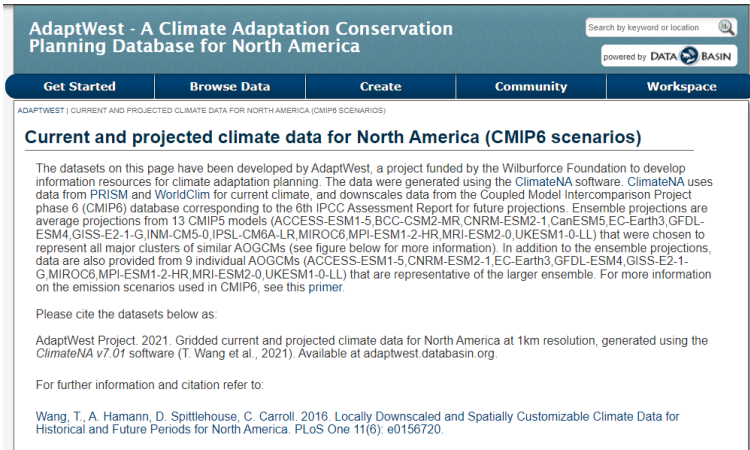


Figure 19b. ClimateNA pre-generated climate layers can be downloaded

A large volume of spatial climate for the entire North America has been generated using ClimateNA by WestAdapt at the spatial resolution of 1 km (Figure 20).



AdaptWest - A Climate Adaptation Conservation Planning Database for North America

Search by keyword or location

powered by DATA BASIN

Get Started | Browse Data | Create | Community | Workspace

ADAPTWEST | CURRENT AND PROJECTED CLIMATE DATA FOR NORTH AMERICA (CMIP6 SCENARIOS)

Current and projected climate data for North America (CMIP6 scenarios)

The datasets on this page have been developed by AdaptWest, a project funded by the Wilburforce Foundation to develop information resources for climate adaptation planning. The data were generated using the ClimateNA software. ClimateNA uses data from PRISM and WorldClim for current climate, and downscales data from the Coupled Model Intercomparison Project phase 6 (CMIP6) database corresponding to the 6th IPCC Assessment Report for future projections. Ensemble projections are average projections from 13 CMIP6 models (ACCESS-ESM1-5, BCC-CSM2-MR, CNRM-ESM2-1, CanESM5, EC-Earth3, GFDL-ESM4, GISS-E2-1-G, INM-CM5-0, IPSL-CM6A-LR, MIROC6, MPI-ESM1-2-HR, MRI-ESM2-0, UKESM1-0-LL) that were chosen to represent all major clusters of similar AOGCMs (see figure below for more information). In addition to the ensemble projections, data are also provided from 9 individual AOGCMs (ACCESS-ESM1-5, CNRM-ESM2-1, EC-Earth3, GFDL-ESM4, GISS-E2-1-G, MIROC6, MPI-ESM1-2-HR, MRI-ESM2-0, UKESM1-0-LL) that are representative of the larger ensemble. For more information on the emission scenarios used in CMIP6, see this primer.

Please cite the datasets below as:

AdaptWest Project. 2021. Gridded current and projected climate data for North America at 1km resolution, generated using the *ClimateNA v7.01* software (T. Wang et al., 2021). Available at adaptwest.databasin.org.

For further information and citation refer to:

Wang, T., A. Hamann, D. Spittlehouse, C. Carroll. 2016. Locally Downscaled and Spatially Customizable Climate Data for Historical and Future Periods for North America. *PLoS One* 11(6): e0156720.

Figure 20. Pre-generated spatial climate data using ClimateNA. Image in public domain.

9. Using the R version of ClimateNA

9.1 Download and installation

The package can be downloaded at: <https://register.climatena.ca/>

Steps to install in RStudio: Tools => Install Packages => Install from: Package Archive File (.zip)

Steps to install in R Console: Packages => Install Package from local file ...

or Directly unzip the package to your library.

You also need to install other R packages (e.g., *terra* and *data.table*) if they have not yet been installed on your computer, using the following code:

```
install.packages('terra')
install.packages('data.table')
```

9.2 Input parameters

ClimateNAr(inputFile, periodList, varList, outDir)

- inputFile – The full name of the input file
- periodList – a list of periods
- varList – a list of climate variables to generate
- outDir – the folder to save the output files

9.3 Return

- All the output files are saved into the outDir designated folder
- The return dataset contains the climate variable of the last period only

9.4 Examples

```
library(ClimateNAr)
csvFile = 'C:/temp/test_na.csv' # download from:
https://climatena.ca/downloads/test\_na.csv
varList=c('MAT','MAP','DD5','CMI','RH')
periodList=
c('Normal_1961_1990.nrm','Year_1902.ann','8GCMs_ensemble_ssp245_2041-2070.gcm')
outDir= 'C:/temp/'
test <-
climateNAr(inputFile=csvFile,periodList,varList,outDir);
head(test)
library(ClimateNAr)
tif <- 'C:/temp/na20k.tif' # download from:
https://climatena.ca/downloads/bc4000.tif
```

```
varList=c('MAT','MAP','DD5','CMI','RH')
periodList=
c('Normal_1961_1990.nrm','8GCMs_ensemble_ssp245_
2041-2070.gcm')
outDir= 'C:/temp/'
test <-
climateNAr(inputFile=tif,periodList,varList,outDir); test
```



By Tongli Wang, Climate Modelling and Forest Applications is licensed under [Creative Commons <CC BY-NC-ND 4.0> License](https://creativecommons.org/licenses/by-nc-nd/4.0/)

Module IV. Introduction to Niche-based Models

Topic 4.1. Introduction to Niche-based ecological Models

Tongli Wang



Viedo #1: The concept of ecological niche:




One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://pressbooks.bccampus.ca/fode010notebook/?p=40#video-40-1>

Viedo #2: Impacts of climate change and species occurrence based niche modeling:



One or more interactive elements has been

 excluded from this version of the text. You can view them online here: <https://pressbooks.bccampus.ca/fode010notebook/?p=40#video-40-2>

Viedo #3: Provenance trials based niche modeling:



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://pressbooks.bccampus.ca/fode010notebook/?p=40#video-40-3>

Contents

- [1. Adaptation of plant species to climate](#)
- [2. Ecological niche](#)
 - [2.1. Fundamental niche](#)
 - [2.2. Realized niche](#)
 - [2.3. Niche conservatism](#)
- [3. Impacts of climate change on ecological niche](#)
 - [3.1. Niche shift](#)
 - [3.2. Mismatch at the population level](#)
- [4. Climatic niche modeling](#)
 - [4.1. Occurrence data-based modeling](#)
 - [4.2. Provenance trials-based modeling](#)

1. Adaptation of plant species to climate

Of many environmental elements, the climate is the primary factor regulating geographic distributions of plant species. The climate was relatively stable over the last 10,000 years (Figure 1). This period started at around -8500 A. D. after the last Ice Age. This provided an opportunity for plant species to evolve and migrate to geographic locations where climatic conditions are suitable for them to survive, grow and reproduce.

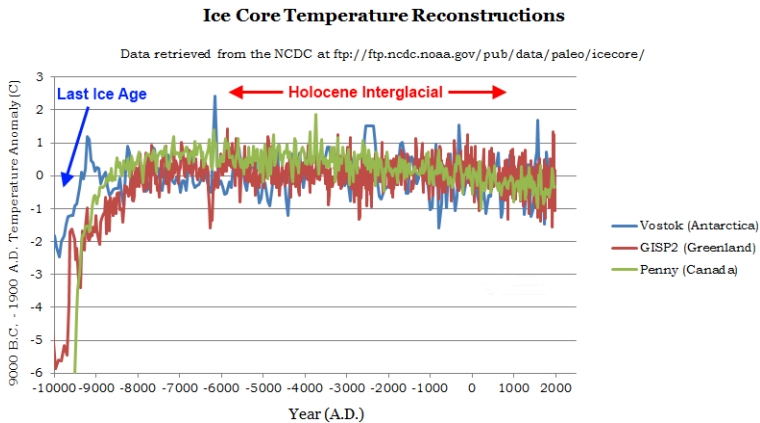


Figure 1. Temperature graph reconstructed from ice cores. Image from an open-source: <https://qph.fs.quoracdn.net/main-qimg-ece88c323825c131daa378b3a6697600.webp>

Through long-term evolution and selection, plant species are adapted to a certain range of climatic conditions over the globe. Species that adapted to similar climatic conditions are grouped into biomes (Figure 2). It clearly shows that different climates support different biomes. For example, the tropical climate supports tropical rainforests, while boreal forests distribute in cold climates in the north.

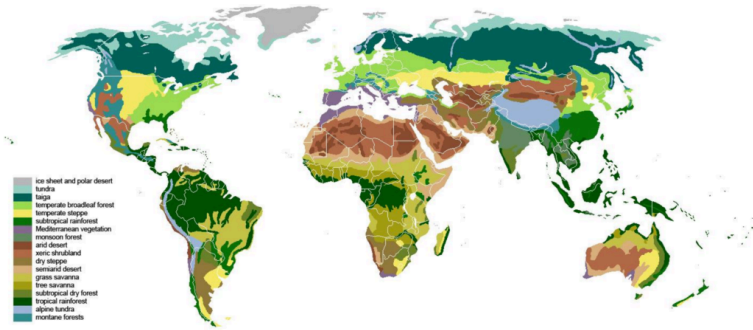


Figure 2. Biomes around the world. Image by Ville Koistinen under CC BY-SA 3.0. Source: <https://en.wikipedia.org/wiki/Biome>

At the species level, each species is adapted to a certain range of climatic conditions (Figure 3). For example, in North America, Douglas-fir is adapted to western North America up to the middle of British Columbia, while spruce and jack pine are distributed further north under a cooler climate. In China, Chinese fir is adapted to a warm and wet climate in the south, while Chinese pine is distributed in the north and northwest adapted to a cool and dry climate.

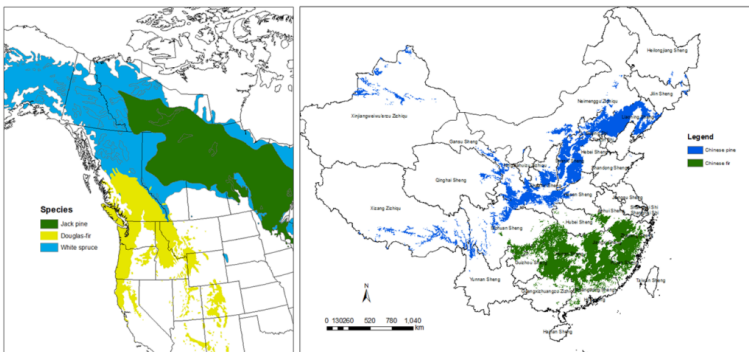


Figure 3. Distributions of some major forest tree species. Image by Tongli Wang under CC BY 4.0.

2. Ecological niche

For a species to maintain its population, its individuals must survive and reproduce. Certain combinations of environmental conditions are necessary for individuals of each species to tolerate the physical environment, obtain energy and nutrients, and avoid predators. The total requirements of a species for all resources and physical conditions determine where it can live and how abundant it can be at any one place within its range. These requirements are termed abstractly the ecological niche.

An ecological niche is originally defined as the position of a species within an ecosystem, describing both the range of conditions necessary for the persistence of the species and its ecological role in the ecosystem. An ecologist, G.E. Hutchinson (1957), suggested that the niche could be modeled as an imaginary space with many dimensions, in which each dimension or axis represents the range of some environmental condition or resource that is required by the species, as illustrated in Figure 4. Hutchinson's definition of ecological niche formed the basis of niche-based ecological models.

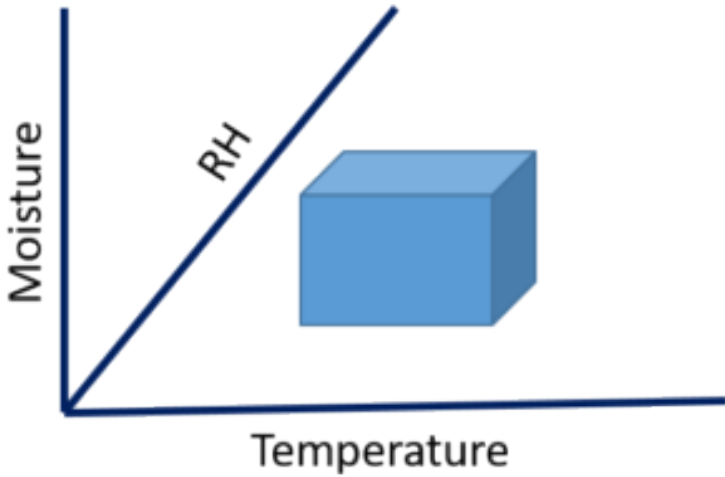


Figure 4. Three climate variables determine the ecological niche of a plant species. Image by Tongli Wang under CC BY 4.0.

Climate is the most important environmental factor determining species distribution. When only climatic components are considered in studying the ecological niche of a species, the ecological niche is also called a climatic niche. The term climatic niche is widely used in ecological modeling related to climate change.

2.1. Fundamental niche

A fundamental niche is the full range of (physical) environmental conditions that a species can occupy and use, without any other limiting factors present which could constrain the population. In other words, the fundamental niche is the potential niche that the species could occupy if there were no limiting factors. Within its fundamental niche, a species can survive, grow, regenerate and evolve.

2.2. Realized niche

A realized niche is the range of conditions actually occupied by a given species. In the real world, a species can only occupy a part of its fundamental niche because of limiting factors, such as competitors, parasites, disease, and physical barriers. Among all the limiting factors, interspecies competition is considered the most important one. Thus, sometimes fundamental niche and realized niche are termed precompetitive and postcompetitive niches, respectively. The difference between a fundamental niche and a realized niche is shown in Figure 5.

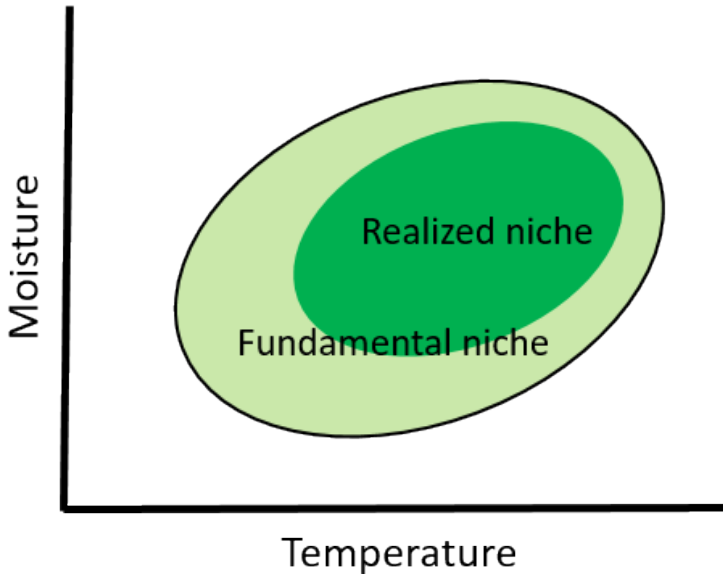


Figure 5. Illustration of fundamental niche and realized niche in two dimensions. Image by Tongli Wang under CC BY 4.0.

The realized niche of a species is reflected by its geographical distribution over the landscape. For example, lodgepole pine is distributed in western North America

as shown in Figure 6. Through determining the climatic conditions of the geographic areas occupied by the species, the realized niche for this species can be defined as shown in Figure 7.

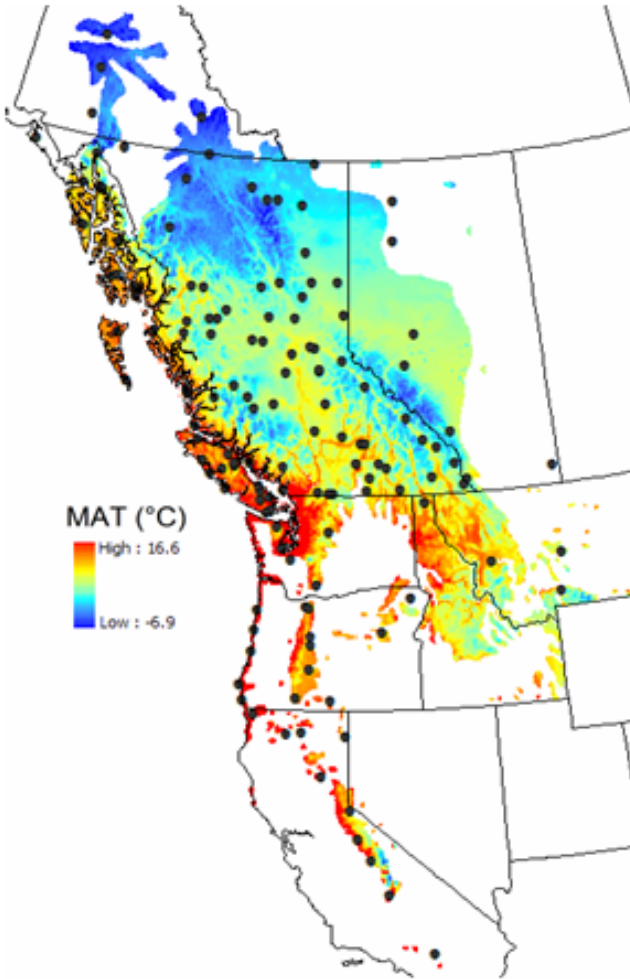
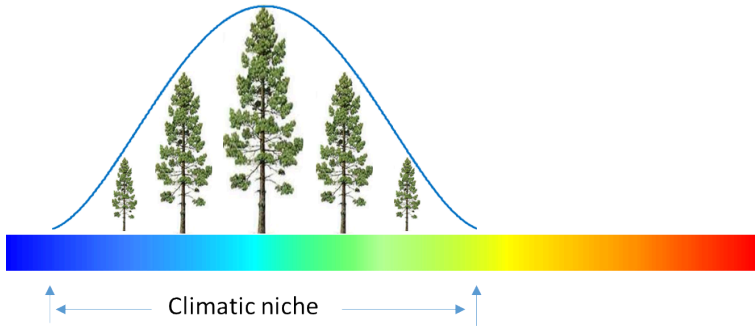


Figure 6. Geographic distribution of lodgepole pine in western North America. Image by Tongli Wang under CC BY 4.0.



*Figure 7. Illustration of the climatic niche of lodgepole pine.
Image by Tongli Wang under CC BY 4.0.*

2.3. Niche conservatism

Many aspects of the fundamental niche can be conserved over long evolutionary timescales. The tendency of species to retain aspects of their fundamental niche over time is called niche conservatism.

Although debate exists in the literature regarding niche conservatism, the evidence is overwhelming in favor of this theory. Many congeneric plant species in Europe and North America had similar geographic ranges on each continent, a pattern interpreted as evidence for niche conservatism. Niche conservatism is observed in six environmental variables in a sample of more than 1000 species of higher plants in Europe ([Wiens, J. J. and C. H. Graham \(2005\)](#)).

Niche conservatism is important because of its relevance to predicting species' responses to global climate change. A tendency toward conservatism would mean that species have only a limited capacity to adapt to new environmental conditions. On the other hand, conservatism also implies that the spread of invasive species is limited by the environmental conditions of their native ranges,

laying robust foundations to predict the extent of potential invasions. In support of this view, [Lavergne et al. \(2013\)](#) showed that declines of European bird species over the last decades were related to their past rates of climatic and habitat niche evolution, suggesting that the degree of niche conservatism may be pivotal for forecasting invasion risk and biodiversity changes.

3. Impacts of climate change on ecological niche

3.1. Niche shift

Climate change is shifting the climatic niche away from the currently occupying geographic locations. Due to the long lifecycle and slow rate of migration of forest trees, such a shift is causing a mismatch between the climate that the species is adapted to and the climate that the species will experience in the future (Figure 8). This is probably the most important challenge for forest species to adapt to a changing climate.

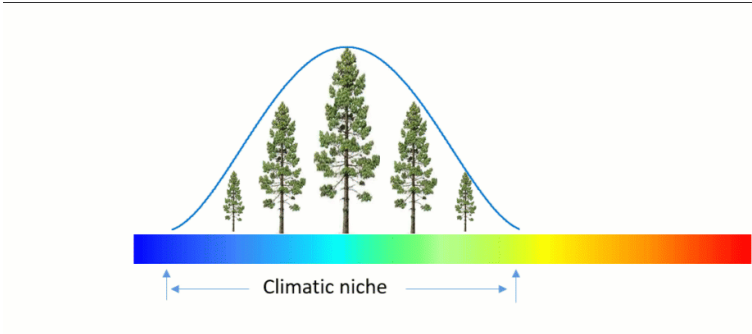


Figure 8. Illustration of the impact of climate change on the shift of climatic niche of a tree species. Image by Tongli Wang under CC BY 4.0.

As a result, trees left out of their climate niche will be suffering from sub-optimal climatic conditions and become vulnerable to insects, diseases, and other hazards (Figure 9). This will likely be the case for existing forests in the lagging areas of the species' climatic niche. Aspen dieback might be an example of the climate niche mismatch (Figure 10).

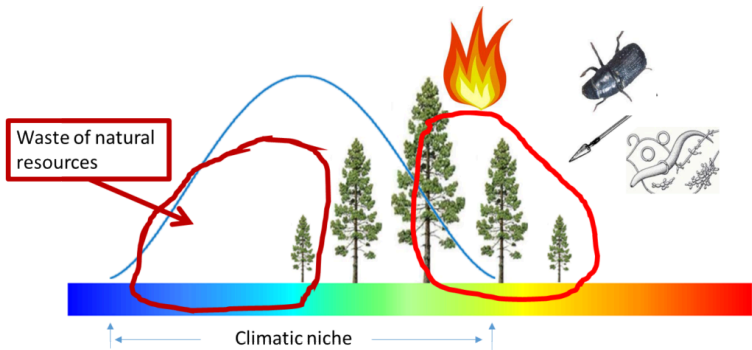


Figure 9. Consequences of the mismatch between species distribution and its climatic niche. Image by Tongli Wang under CC BY 4.0.



Figure 10. Dieback of aspen stands. Photo by Phil Kemp, US Forest Service.

On the other hand, some areas may become suitable for the species to occupy, but it may need a long period for the species to be able to migrate there. Thus, the potential of these areas is not to be used for a long time if we completely rely on the natural migration of the species.

3.2. Mismatch at the population level

As shown in Figure 11, climatic mismatches occur in two dimensions. One is the mismatch in a climatic niche at the species level (horizontally), and another is the mismatch at the population level (vertically). Due to the geographical shift in the climatic niche, some populations will be left out of the climatic niche of the species as discussed previously. Even for populations that remain within the climate niche, they are no longer under their optimal climate conditions. Survival may not be a problem, but productivity may be compromised. The mismatches in both dimensions need to be considered for developing forest adaptive strategies.

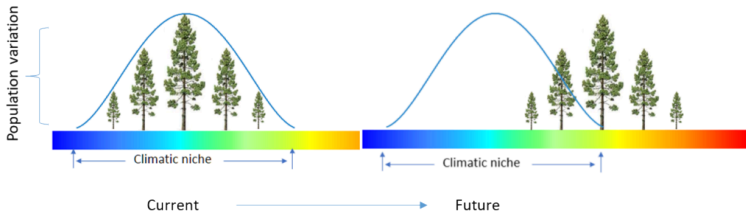


Figure 11. Illustration of the mismatches in two dimensions, horizontally at the species level and vertically at the population level. Image by Tongli Wang under CC BY 4.0.

4. Climatic niche modeling

A climatic niche model is to model or define the climatic niche of a given species and to predict its geographic distribution for different climatic conditions, including future climate projections.

4.1. Species occurrence-based modeling

A climatic niche model can be built based on the relationship between the geographical occurrence of a species and its corresponding climatic variables (Figure 12a). Such a model reflects the **realized niche** of the species (Figure 12b) because the species occurrence is the result of not only the environmental conditions but also migration, local adaptation, and interactions with other species and organisms under recent/historical climatic conditions. The models of this form are typically referred to as climatic niche models, bioclimatic envelope models, or ecological niche models. These terms are often used interchangeably. Occurrence data-based climatic niche models are widely used because the occurrence data is probably the easiest to obtain for a plant species, and models' prediction accuracies are reasonably high. Species occurrence-based models are analytical models, as discussed previously.

Climatic niche models are widely called correlative

species distribution models (SDMs). However, as a climatic niche model defines and predicts the suitable climatic habitat rather than the actual distribution of a species, which involves a series of evolutionary and ecological processes. Therefore, a climatic niche model is, strictly speaking, not a species distribution model (SDM).

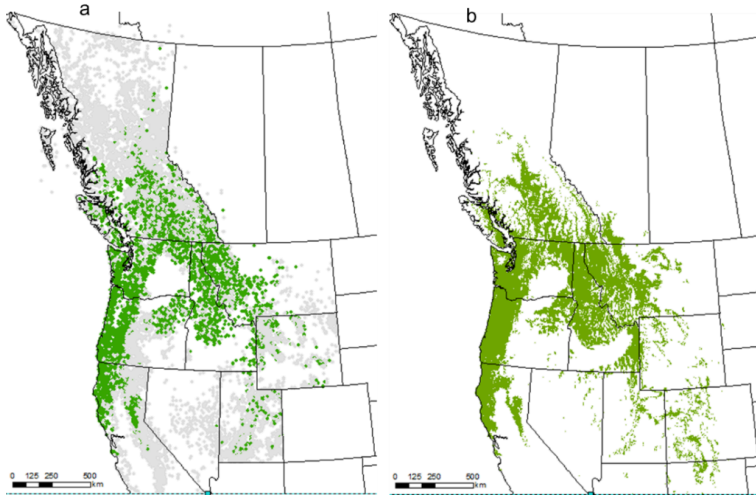


Figure 12. Spatial distribution of the occurrence data (left, green color for presence and grey for absence) and climatic niche (right) of Douglas-fir in North America. Image by Tongli Wang under CC BY 4.0.

A predicted current spatial distribution of the climatic niche may not match well with the current distribution of the species in some areas. Typically, the extent of the climatic niche is greater than the observed range of the distribution. It may not suggest that the model is wrong; instead, it indicates that those areas are potentially suitable for this species to be there. This phenomenon is probably attributable to physical barriers to migration or human disturbances.

4.2. Provenance trials based modeling

Provenance trails in forest trees are traditionally used to determine optimal seed sources for reforestation. Populations collected from a wide range of environments (provenances) are tested in field common-garden experiments under one or multiple planting sites spanning a range of climatic conditions. These materials are now widely used to explore among-population variation along climatic gradients (i.e., genecology functions) and responses of populations to climatic conditions (i.e., response functions). The latter can be used to model the climatic niche at both the population and the species levels.

A response function can be developed for an individual population. Response functions are used to predict the response of adaptive traits, including growth and survival, to climate. The range of the climatic conditions suitable for the population can be used to define the climatic niche of the population. When individual response functions are integrated into a universal response function (Figure 13, left), they can be used to model the climatic niche of the species (Figure 13, right).

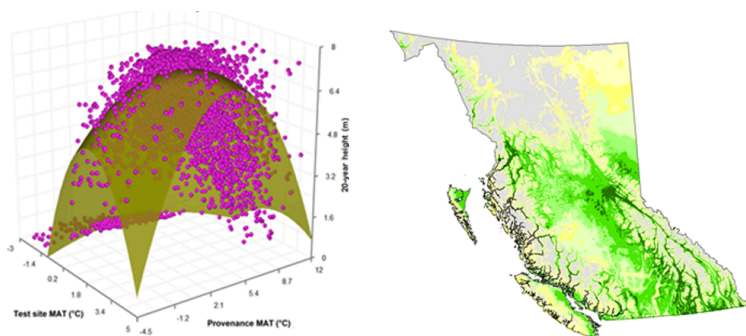


Figure 13. A universal response surface (left) and climatic niche (right) of lodgepole pine in British Columbia. Image by Tongli Wang under CC BY 4.0.

Since the trees growing in the common-garden

experiments are free from interspecific competitions and the limit of migration, the climatic niche predicted by a response function can be considered as the **fundamental niche** of the population or the species. This will be discussed in detail later.



By Tongli Wang, Climate Modelling and Forest Applications is licensed under [Creative Commons <CC BY-NC-ND 4.0> License](https://creativecommons.org/licenses/by-nc-nd/4.0/)

Topic 4.2. Introduction to Genecology functions

Tongli Wang



Video #1 and #2: Concept of genecology



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://pressbooks.bccampus.ca/fode010notebook/?p=42#video-42-1>



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://pressbooks.bccampus.ca/fode010notebook/?p=42#video-42-2>

Video #3: Transfer functions:



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://pressbooks.bccampus.ca/fode010notebook/?p=42#video-42-3>

Video #4: Genecology functions and predictors:



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://pressbooks.bccampus.ca/fode010notebook/?p=42#video-42-4>

Contents

[1. Overview](#)

[2. Definition of genecology](#)

- [2.1. Definition](#)
- [2.2. Some relevant terms](#)
- [2.3. Objectives of genecology](#)
- [2.4. Provenance test](#)
- [3. Transfer functions](#)**
 - [3.1. Transfer functions](#)
 - [3.2. Features and limitations](#)
- [4. General transfer functions](#)**
 - [4.1. General transfer functions](#)
 - [4.2. Relative performance-based general transfer functions](#)
 - [4.3. Limitations of general transfer functions](#)
 - [4.4. Universal transfer functions](#)
- [5. Genecology functions](#)**
 - [5.1. Genecology functions](#)
 - [5.2. The difference in genecology functions among environments](#)
- [6. Geographic vs. climatic variables](#)**
 - [6.1. Representation of environmental variables](#)
 - [6.2. Accessibility](#)
 - [6.3 local vs. universal](#)
 - [6.4. Applicability for climate change](#)
 - [6.5. Non-climatic factors](#)

1. Overview

As mentioned in the previous topic and shown in Figure 1, climatic mismatches occur in two dimensions. One is the mismatch in the climatic niche at the species level (horizontally), and another is the mismatch at the population level (vertically). The niche-based models at the

species level can be used to predict the shift in the climatic niche at the species level, while genecology functions are to predict among-population variations within a species.

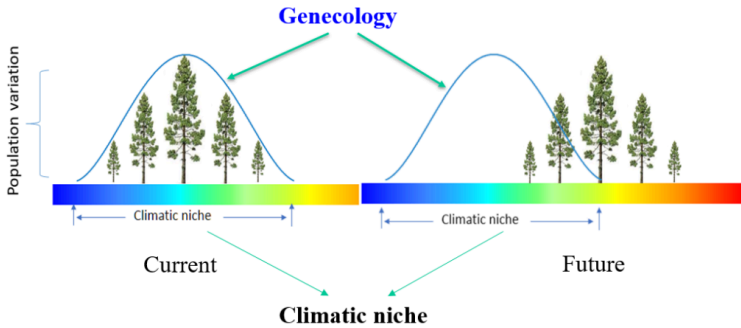


Figure 1. Illustration of mismatches caused by climate change in terms of climatic niche and among-population variation, and the role of gynecology in coping with the mismatches. Image by Tongli Wang under CC BY 4.0.

2. Definition of genecology

2.1. Definition

Genecology is the study of intraspecific genetic variation in relation to environmental conditions. In other words, it is to study the relationships between genetic variation and environmental gradients within a species.

Genecology is often referred to as ecological genetics, which is actually more widely used. However, strictly speaking, ecological genetics is the study of how ecologically relevant traits evolve in natural populations, which covers inter-species interactions. Nevertheless, these two terms are used interchangeably.

2.2. Some relevant terms

Provenance – the place of seed origin (or seed source).

Population – a group of individuals of the same species that live in a particular geographical area, and have the capability of interbreeding. Individuals from the same provenance are also called a population.

Provenance is often used to refer to a population although it is not accurate. Thus, “provenance” and “population” are often used interchangeably in genecology.

Genetic variation – the differences in DNA among individuals if defined at the individual level, or the differences in allele frequency among populations if defined at the population level.

Genetic cline – a geographic or climatic pattern of change in a trait.

Genetic and environment interaction (GxE) – different genotypes respond to environmental variation in different ways.

2.3. Objectives of genecology

Operational objectives:

- To determine how far seed can be moved
- To delineate geographic breeding zones
- To select optimal provenances

Theoretical objectives:

- Testing evolutionary and ecological theories about factors determining the evolution of species range

- Predicting response of populations to climate change
- Dependency on climatic drivers
- Determining the underlying genetic basis of adaptive traits: Ecological genomics

2.4. Provenance test – materials for genecology

Provenance test (or provenance trial) – a common garden experiment that compares trees or plants collected from different provenances. Typically provenance tests compare widely separated seed sources. The test may contain multiple test sites. Provenance tests are the main source of experimental materials for conducting genecology, and they are available for many commercially important forest tree species in developed countries, especially in North America and Europe. They were mostly established during the 1970s.

Figure 2 shows a comprehensive provenance test for lodgepole pine. Populations from 140 provenances range-wide were planted in 60 common garden trials (planting sites) in British Columbia (BC), Canada. At each site, an incomplete block design was used to include 60 populations due to a large number of populations and limited test sites with relatively uniform conditions. Thus, not every population was tested at all the 60 sites except seven reference populations. Tree height was measured at multiple ages. This provenance test will be used as an example to demonstrate various approaches in genecology.

Provenance trials for lodgepole pine

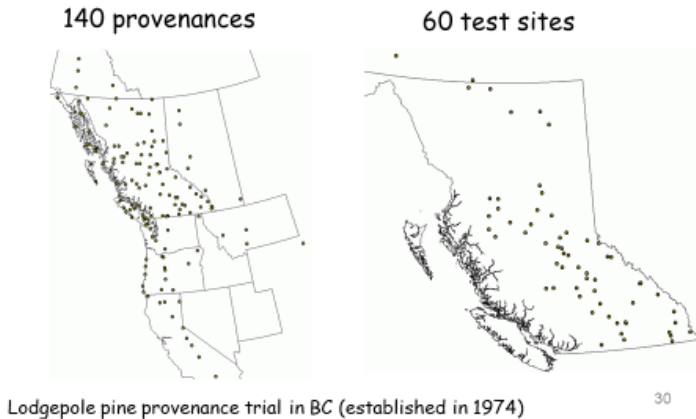


Figure 2. A provenance test comprising 140 provenances of lodgepole pine in western North America and tested in British Columbia. Image by Tongli Wang under CC BY 4.0.

3. Transfer functions

Transfer functions are most commonly used in genecology to reflect the effect of transfer distance of populations in terms of geographical or climatic variables. A provenance test is required for developing a transfer function.

As illustrated in Figure 3, populations are transferred from their origins to a test site in the mid of BC in the provenance test. Those populations grow under new environmental conditions, and their performances may be affected. As the populations are grown under the same climatic conditions (at the same test site), the difference in the performance of populations is supposed to be attributable to genetic variation among the populations.



Figure 3. Illustration of population transfers in the provenance test of lodgepole pine. Image by Tongli Wang under CC BY 4.0.

3.1. Transfer functions

The genetic variation observed in a provenance trial at each test site may be associated with the population transfer distance. It is hypothesized that the longer the transfer distance, the larger the transfer effect. Thus, the transfer distance is used as a predictor in a transfer function using a quadratic function (or a second-degree polynomial function) as:

$$Y = a + bx + cx^2 \quad [1]$$

Where, Y is the performance of a population; x is the transfer distance; and a , b , and c are parameters to be estimated during the model fitting.

A quadratic function is most often used in building a transfer function or other functions in genecology for its simplicity and the bell-shaped curve fitting well with a biological interpretation. A simple linear function is also

often used, but it needs to be extremely cautious for extrapolation beyond the range of the data coverage.

The transfer distance can be in a physical distance for geographical variables (e.g., latitude, longitude, or elevation), or in units for climatic variables (e.g., temperature and precipitation). Climatic variables are increasingly used as they have many advantages over geographical variables, particularly for predicting the impact of climate change on forests.

Climatic variables are needed for all the provenances to build a transfer function using the transfer distance in climatic variables, as shown in Figure 4. A map of mean annual temperature (MAT) covering the lodgepole pine distribution and a transfer function for a test site in the mid of BC.

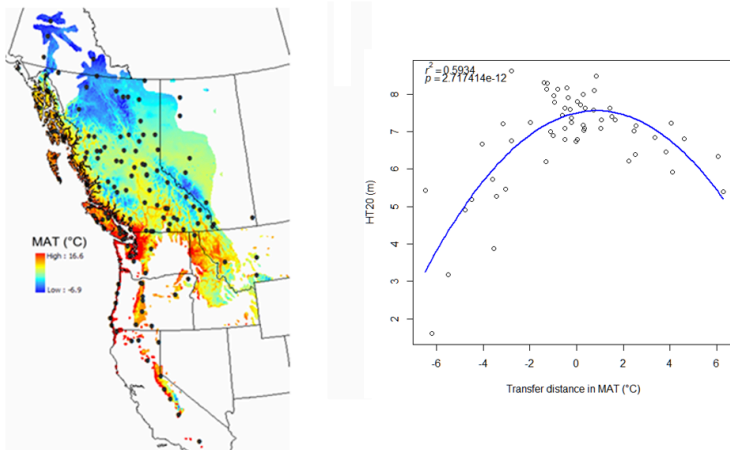


Figure 4. A map of mean annual temperature (MAT) covering the lodgepole pine distribution (left) and a transfer function for a test site in the mid of BC (right). Image by Tongli Wang under CC BY 4.0.

The transfer distance in MAT (MAT_CTD) is calculated as:

$$MAT_CTD = MAT_s - MAT_p \quad [2]$$

Where, MAT_s is the MAT of the test site, MAT_p is the MAT of the provenance. Then, the transfer function is written as:

$$Y = a + bMAT_CTD + cMAT_CTD^2 \quad [3]$$

Populations with negative transfer distances on the left of the plot were transferred northward (from warmer climates to cooler climates). On the other hand, populations with positive transfer distances on the right of the plot were transferred southward (from cooler climates to warmer climates). Populations with zero and around zero transfer distance are considered local populations. The local populations usually have the best performance; however, the peak of the curve is not always exactly at the zero transfer distance point, as shown in Figure 4.

3.2. Features and limitations

A transfer function represents the transfer effect of populations. It reflects the **genetic effect** of climate on populations. It is intuitive and easy to build. Only one test site is required to build a transfer function.

However, there are several limitations for transfer functions. A transfer function built for one site may be different from the one built for another site, even with the same group of populations. Thus, a transfer function developed for one site, strictly speaking, only represents the transfer effect of populations under the particular environment of that test site. The number of test sites is limited for most provenance tests, and people tend to

oversell the results drawn from these sites. Therefore, we need to be aware of such a risk of using a transfer function built for one environment to represent the situations for other environments.

4. General transfer functions

4.1. General transfer functions

Due to the limitation of transfer functions discussed above, general transfer functions were proposed if multiple test sites are available. The only difference between a transfer function and a general transfer function is that data from all test sites are pooled together in developing a general transfer function. The formulas [1] – [3] all apply to a general transfer function.

The curve of a general transfer function shown in Figure 5 was developed based on the observations from the same provenance test mentioned above. The relationship between the performance and transfer distance of populations in a transfer function is usually weak because the environmental effects of specific tests and population and test site interactions are ignored (to be discussed below).

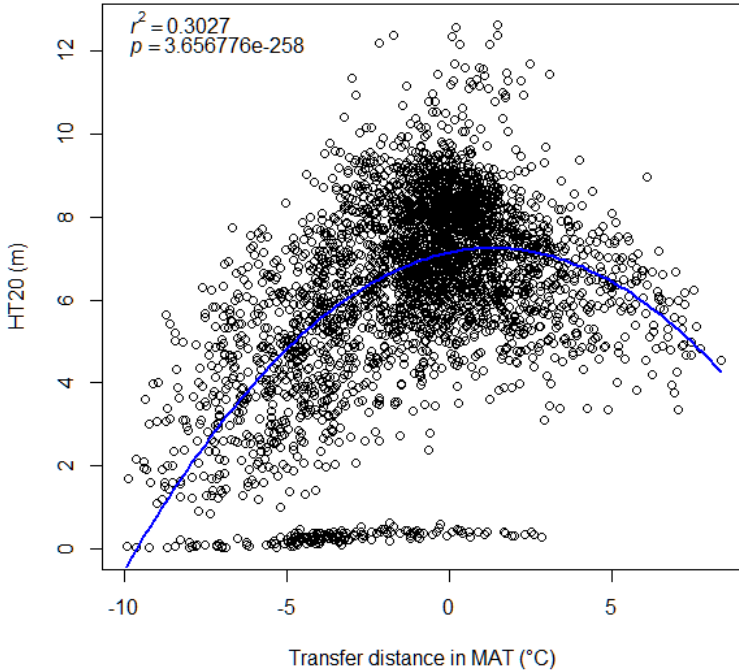


Figure 5. A general transfer function that combines observations across all tests. Image by Tongli Wang under CC BY 4.0.

4.2. Relative performance-based general transfer functions

The most important reason for a general transfer function being weak is the difference in the performance of populations among sites. For example, the outliers at the bottom of Figure 5 are populations growing at several cold sites in the north. Their total height at age 20 is less than a meter while the average tree height can reach over 5 meters at other test sites. Such differences are not reflected in the general transfer function. An approach was developed to overcome this problem by using relative performance.

With this approach, the performances of populations are converted to proportions relative to the performance of the local population. It can substantially improve the relationship of the general transfer function, as shown in Figure 6. It is worth reminding that the relative performance needs to be converted back to actual performance in applications of this model.

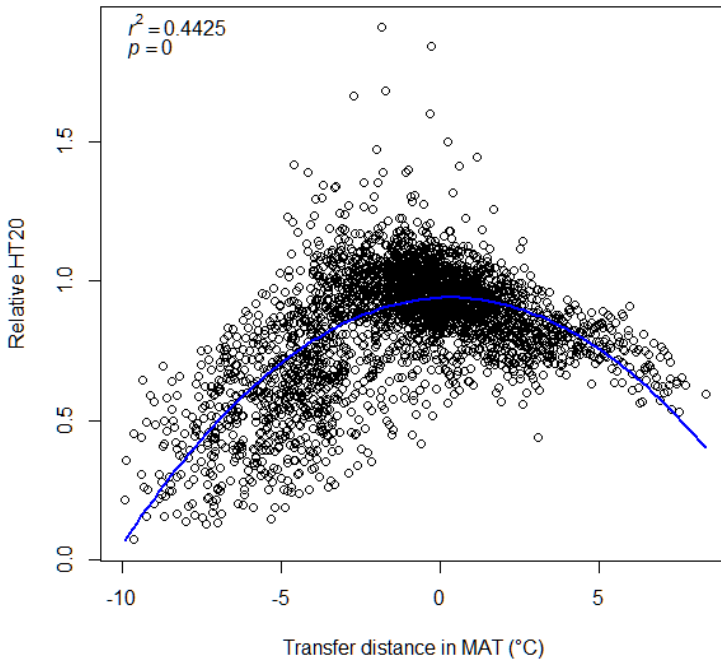


Figure 6. A general transfer function using relative performance. Image by Tongli Wang under CC BY 4.0.

4.3. Limitations of general transfer functions

In addition to the weak relationships of general transfer functions as mentioned above, there are other limitations for general transfer functions. As discussed earlier, the

transfer effects of populations are different among test sites. By pooling the data across all test sites to develop a general transfer function, the difference in the transfer effect of populations among test sites is averaged out, as illustrated in Figure 7. This compromises the prediction power of general transfer functions. In addition, the same amount of transfer distance may have different transfer effects. For example, for a -2°C transfer distance in temperature, the transfer from 2°C to 0°C may be different from the transfer from 10°C to 8°C . This point is related to population and test site interactions and will be further discussed in the following sections. Nevertheless, general transfer functions are the simplest to implement; thus it is still commonly used.

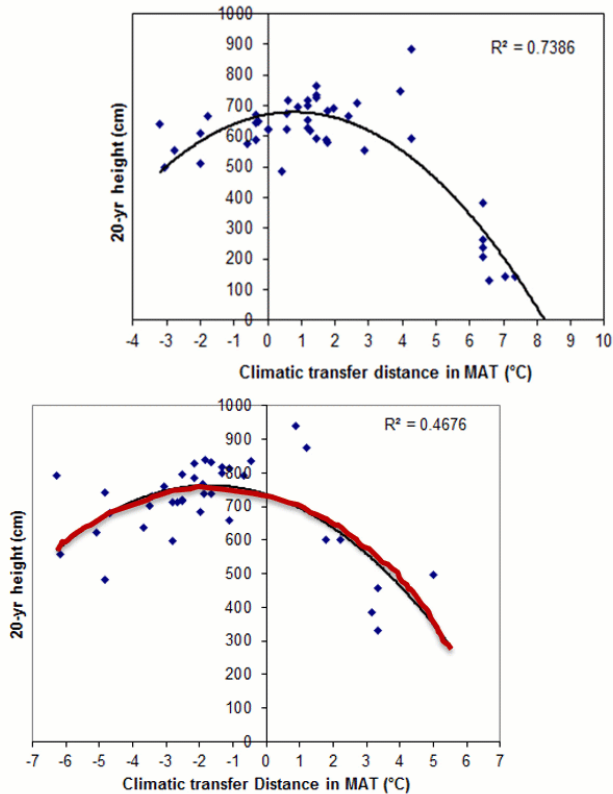


Figure 7. Illustration of difference in transfer effect between two test sites. Image by Tongli Wang under CC BY 4.0.

4.4. Universal transfer function

An improved version of the general transfer function was proposed recently (Leites et al 2012). It can be called “Universal transfer functions”. A universal transfer function also pools data across test sites. The difference is that it includes provenance climate and an interaction term in addition to transfer distance, as well as some random effects. The

function can be written as:

$$\bullet Y = a + bX + cX^2 + dP + eP^2 + fXP + rs$$

Where,

X is transfer distance;

P is provenance climate;

XP is the interaction term;

rs are some random effect terms.

The universal transfer function can overcome the problems listed above, but the random terms cause additional problems. The model can be too complicated and often fails to converge. Thus, it leads to the use of only one climate variable in applications, such as in Douglas-fir in North America ([Leites et al. 2012](#)) and sessile oak in Europe ([Sáenz-Romero et al. 2016](#)), which can compromise the model capability.

5. Genecology functions

5.1. Genecology functions

A genecology function uses the environmental variables of provenances, rather than transfer distance used in both transfer functions and general transfer functions. Thus, a genecology function represents the relationship between the performance of populations and environmental variables of the population origins (or provenances), which exactly reflects the definition of genecology. A quadratic genecology function can be written as:

$$Y = \mathbf{a} + \mathbf{b}x + cx^2$$

Where, Y is the performance of a population; x is an

environmental variable of provenance; and a , b , and c are parameters to be estimated during the model fitting.

Using the same data for building a transfer function in Figure 4, a genecology function is developed, as shown in Figure 8. Although the shapes of the curves are quite different, the R square values are identical between the two functions.

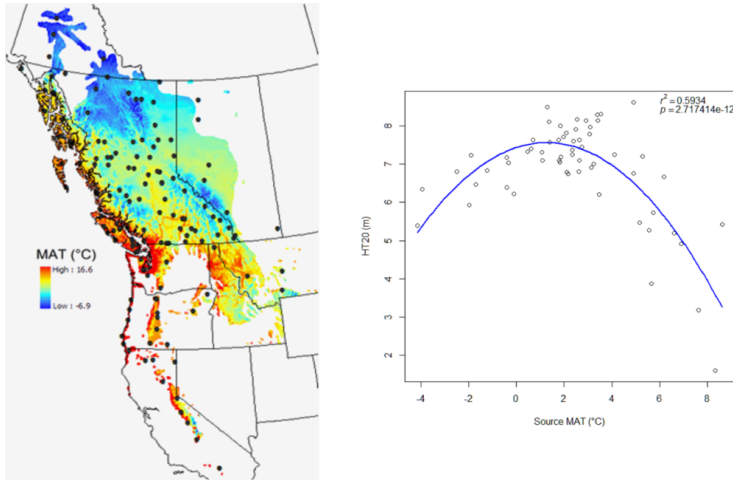


Figure 8. A map of mean annual temperature (MAT) covering the lodgepole pine distribution (left) and a genecology function for a test in the mid of BC (right). Image by Tongli Wang under CC BY 4.0.

5.2. The difference in genecology functions among environments

Genecology functions developed for different environments show different relationships (Figure 9). Thus, a genecology function built based on information from one environment may not represent the relationship

for other environments. This is the same as for the transfer function.

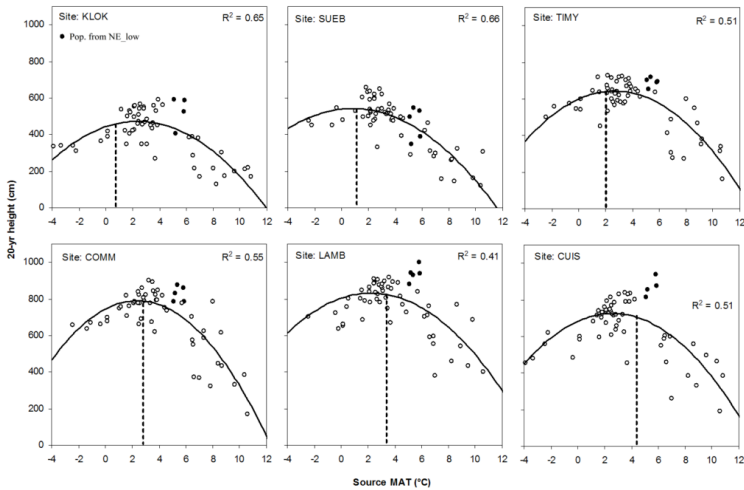


Figure 9. genecology functions developed at six test sites from cold to warm climates. Image by Tongli Wang under CC BY 4.0.

6. Geographic vs. climatic variables

Both transfer functions and genecology functions can be built with either geographic or climatic variables. There are advantages and disadvantages to using geographic or climatic variables as predictors. We will make a detailed comparison between the two.

6.1. Representation of environmental variables

Geographic variables are not direct environmental variables, whereas climatic variables are. However, geographic variables are related to environmental variables. For example, elevation is strongly related to temperature, which affects plant performance, as illustrated in Figure 10 – plants growing well in the valley but not at

the high mountains. Thus, they are often used to substitute environmental variables. However, if we build a greenhouse at the top of the mountain, we can grow any plant that can only grow at the bottom of the valley. It suggests elevation itself does not affect plant growth.



Figure 10. “Little Lakes Valley viewed from about 11,000 feet elevation, hiking along the Mono Pass Trail, Eastern Sierra Mountains, California (mono14xy)” by mlhradio is licensed under CC BY-NC 2.0

6.2. Accessibility

Geographic variables are more accessible than climatic variables. That is why they were more often used for experimental design and data analysis in the past. With increasingly available climate data, accessing climatic variables are becoming as easy as accessing geographic variables. Especially with the scale-free climate models

ClimateNA and ClimateAP, climatic variables can be easily obtained for any specific location within their coverage.

6.3 local vs. universal

A transfer function or a genecology function built with geographic variables only works for the region where the model is built. This is because the effect of a geographic variable changes from region to region, or location to location. For example, an elevation of 2000m in California is suitable for many plants to grow; but the same elevation in BC is beyond the treeline. Also, an increase in elevation by 100 m in California will result in a decrease in temperature by 0.65°C; but the same change in northern BC will only result in a decrease in temperature by about 0.2°C. In contrast, a transfer function or a genecology function built with climatic variables can be applied anywhere.

6.4. Applicability for climate change

A transfer function or a genecology function built with geographic variables cannot be used for assessing the impact of climate change as geographic variables are static and have no future projections. In contrast, climatic variable-based functions can be used to predict the past and future. That is the reason why genecology has been a hot area.

6.5. Non-climatic factors

Geographic variables are still useful to reflect some

important non-climatic factors. Genetic variations among populations are also affected by spatial autocorrelation involving migration and gene flow related to physical barriers and distance. Those factors sometimes need to be included in genecology functions. We will discuss this further in the next topic.



By Tongli Wang, Climate Modelling and Forest Applications is licensed under [Creative Commons <CC BY-NC-ND 4.0> License](https://creativecommons.org/licenses/by-nc-nd/4.0/)

Topic 4.3. Introduction to Climate Response Functions

Tongli Wang



Lecture videos:



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://pressbooks.bccampus.ca/fode010notebook/?p=44#video-44-1>



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://pressbooks.bccampus.ca/fode010notebook/?p=44#video-44-2>



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://pressbooks.bccampus.ca/fode010notebook/?p=44#video-44-3>



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://pressbooks.bccampus.ca/fode010notebook/?p=44#video-44-4>

Contents

[1. Overview](#)

[2. Definition](#)

[3. Response functions with a single variable](#)

[3.1. Objectives](#)

[3.2. Provenance test and climate data](#)

[3.3. Finding populations for future climates](#)

[3.4. Anchor points](#)

[4. Response functions with multiple climatic variables](#)

[4.1. Function](#)

[4.2. Response surface and function parameters](#)

[4.3. Model interpretation](#)

[5. Multiple response functions](#)

[5.1. Objectives](#)

[5.2. Putting multiple response curves on the same plot](#)

[5.3. Comparisons among multiple response curves](#)

[5.4. Limitations](#)

[6. Universal response functions](#)

[6.1. Objectives](#)

[6.2. Function](#)

[6.3. Features](#)

1. Overview

The concept of the population response function is relatively new in ecology. It was first introduced to model forest trees by Dr. Rehfeldt in 1999. It can predict the mismatch at the population level individually (vertically in Figure 1). It can also predict the mismatch at the species level integratively if an adequate number of populations are modeled (both vertically and horizontally in Figure 1). In addition, the climatic niche model by using response functions represents the fundamental niche of the species instead of the realized niche as modeled using presence-absence data.

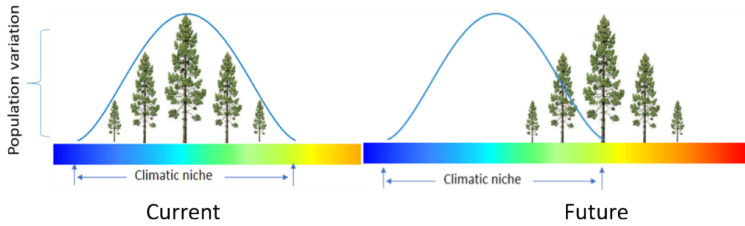


Figure 1. Illustration of mismatches caused by climate change in terms of climatic niche and among-population variation. Image by Tongli Wang under CC BY 4.0.

2. Definition

A population response function describes the response of a population to the change of environmental conditions. It represents the environmental effect on the population. It also reflects the plasticity of the population over a range of environmental conditions.

This is contrasting to genecology functions or transfer functions, which are to model among-population variations in performance along environmental gradients of population origins. They reflect genetic variation in populations associated with environmental factors. A common feature of these two kinds of models is that they both are built based on observations from provenance tests.

A typical linear and quadratic population response functions can be written as:

$$Y = a + bx \quad [1]$$

$$Y = a + bx + cx^2 \quad [2]$$

Where, Y is the performance of a population; x is one (or more) environmental variable of test sites; and a , b , and c are parameters to be estimated to fit the model.

Similar to a genecology function, a response function

can be built with geographical variables (e.g., latitude, longitude, or elevation), or climatic variables (e.g., temperature and precipitation). Climatic variables are increasingly used as they have many advantages over geographical variables, particularly for predicting the impact of climate change on forests. Up to date, all response functions are built with climatic variables.

3. Response functions with a single climate variable

3.1. Objectives

A single climatic variable often plays a dominant role in affecting the performance of populations. For simplicity, we can build a response function with just one climatic variable. With a single climatic variable, a response function can reflect a simple linear relationship with a single climatic variable. It can also be built with a quadratic or a cubic function. A quadratic function showing a bell-shaped response curve is most commonly used as most climatic variables have an optimal range suitable for a plant species to grow.

3.2. Provenance test and climate data

Let's use the same provenance test shown in Figure 2 to demonstrate response functions. In the provenance test, each population was planted at multiple test sites. After 20 years of growth under different environmental conditions, the population showed substantial differences in performance among the test sites.

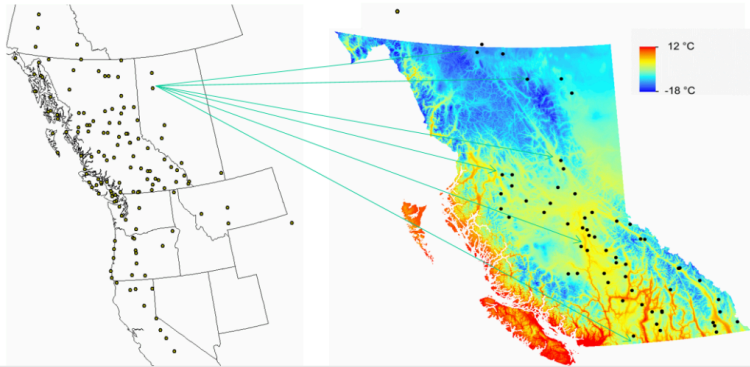
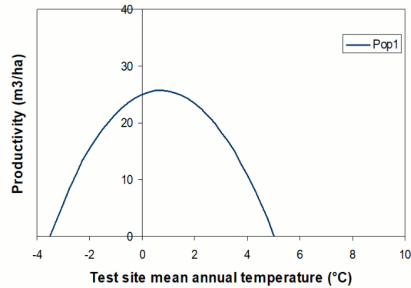


Figure 2. A provenance test comprising 140 provenances of lodgepole pine in western North America and tested in British Columbia. Image by Tongli Wang under CC BY 4.0.

3.3. Finding populations for future climates

After we obtained climate data from ClimateNA, we can develop a response function for this population to represent the effect of temperature on the tree height of the population. As shown in Figure 3, the response function predicts the performance of the population Pop1 in different temperatures with its optimal performance around its local temperature. With an increase in temperature by 2°C due to climate change, the predicted productivity is predicted to drop substantially. We need to find populations that perform better in a warmer climate. Ideally, we may find populations that can perform well over a broader range of temperatures covering both the current and future climates.

- Use provenance test data to develop response functions to climate



$$Y = a + bx + cx^2 \quad [1]$$

x – A climate variable for sites

Figure 3. Illustration of the response of a population to different temperatures and the responses of populations suitable for future climates. Image by Tongli Wang under CC BY 4.0.

3.4. Anchor points

Most provenance tests have a limited number of test sites due to the high cost and availability of planting sites. Also, due to a large number of provenances, there is not enough space to include all provenances. As a result of these situations, some provenances are often not tested at some extreme climatic conditions and make it difficult to build an effective response function. For the latter case, a genealogy function can be used to predict the performance of the missing populations at the site where the genealogy function is built for. The predicted values are called anchor points as they play a role to anchor the

curve down at the end of the species range. Details are described in [Wang et al. \(2006\)](#).

4. Response functions with multiple climatic variables

4.1. Function

A single variable is often not sufficient to explain the variation in the performance of a population across different environmental conditions. In those cases, two or more environmental variables are needed to be included in the model to increase its explanatory power. Then, the formula of a quadratic response function with multiple variables becomes:

$$Y = b_0 + b_1x_1 + b_2cx_1^2 + b_3x_2 + b_4cx_2^2 + \dots$$

Where, Y is the performance of a population; x_1 , x_2 , ..., are the climatic variables; and b_0 , b_1 , b_2 , ..., are parameters to be estimated to fit the model.

4.2. Response surface and function parameters

We use the dataset from the same provenance test for lodgepole pine to illustrate the response surface. A quadratic response function is built with two climatic variables, mean annual temperature (MAT) and annual heat-moisture index (AHM), as predictors. The response surface represents the response function is shown in Figure 4. The parameters estimated for the response function and the significance of the predictors are listed in Table 1.

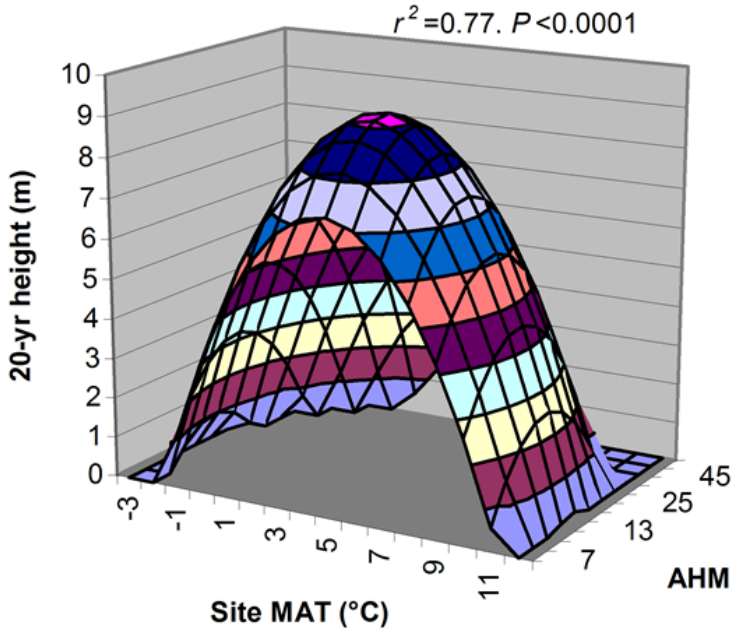


Figure 4. Response surface predicted by a response function built with two climatic variables, mean annual temperature (MAT) and annual heat-moisture index (AHM). Image by Tongli Wang under CC BY 4.0.

Table 1. Parameters estimated and significance of the predictors

	Estimate	Std. Error	t value	Pr(> t)	Significance
(Intercept)	1.731036	1.758753	0.984	0.3298	
MAT	2.231324	0.511851	4.359	6.68E-05	***
AHM	0.303606	0.145037	2.093	0.0415	*
MAT^2	-0.24298	0.05008	-4.852	1.28E-05	***
AHM^2	-0.00755	0.003411	-2.214	0.0315	*
MAT*AHM	-0.0126	0.028548	-0.441	0.6608	

4.3. Model interpretation

The information from the response surface shown in Figure 4 and the parameters estimated in Table 1 show that MAT plays a dominant role in affecting the performance of the population, while AHM also significantly contributes to the model. However, the interaction between MAT and AHM does not significantly contribute to the model. In total, the model explains 77% of the total variation in tree height of the population among all the test sites. The model is highly accurate at $P < 0.0001$.

The response surface also shows: 1) this population can grow within the range between -3°C and 11°C for MAT and between 5 and 45 for AHM, but slightly varying depending on each other (i.e., climatic niche of this population); 2) the optimal climatic condition for this population is a combination between MAT at around 5°C and AHM at about 13; however, 3) the population can grow well between 3°C and 7°C for MAT and between 10 and 20 for AHM.

5. Multiple response functions

5.1. Objectives

A population response function built for a population reflects the characteristics of the given population. In other words, the function does not represent the responsive nature of other populations. To understand the difference among populations in this regard, we need to build response functions for other populations individually and then make comparisons among them. If there are a sufficient number of populations having response functions built, they can provide an overview of the response to climate at the species level.

5.2. Putting multiple response curves on the same plot

After response functions are built for multiple populations, making comparisons among the response functions is important. Putting the response curves on the same plot is the most straightforward approach to making such a comparison. However, it is difficult to do so if the response functions are built with two or more climate variables. A solution for this is only to compare the response against the dominant variable while keeping the other variables constant or at their optimums. Figure 5 shows multiple response curves against MAT on the same plot while keeping AHM at its optimum for lodgepole pine.

Response functions for lodgepole pine populations

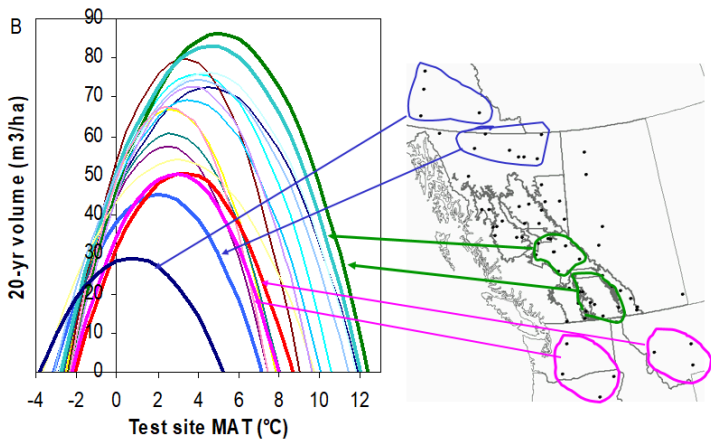


Figure 5. Comparisons in response curves among populations of lodgepole pine in BC. Image by Tongli Wang under CC BY 4.0.

5.3. Comparisons among multiple response curves

Through comparisons among the response curves shown in Figure 5, the following results can be drawn:

- 1) Lodgepole pine can grow between -4°C and 12.5°C in MAT (broader than Pop1).
- 2) Northern populations outperform other populations in a very cold environment at the cold end of the species range, but grow much slower in all other environments.
- 3) Populations from the warmer locations in the south are not performing better than many other populations in warmer conditions.

- 4) The best performing populations are from the center of the distribution in southern BC. Those populations not only outperform other populations in most environments but also have a broader adaptational range.
- 5) An increase in temperature due to climate change will improve the performance of the northern populations to some extent; it will not make them outperform the central populations. To count on the increase in productivity in the north to compensate for the productivity loss in the south due to global warming is not realistic.
- 6) Thus, to mitigate the impact of climate change on the productivity of this species, we need to migrate the central populations to the north when the climate becomes suitable.

These results drawn from the comparison provide a scientific basis for developing adaptive forest management strategies for this species.

5.4. Limitations

Response functions predict the response of populations to the change in environmental conditions, primarily climatic conditions. It is probably the most straightforward approach for understanding the impact of climate change on the performance of populations and for developing forest adaptive strategies. However, it has the following two limitations:

- 1) A response function developed for a

population only reflects the responsive nature of the given population. Thus, it may not be able to represent other populations.

- 2) A credible response function can only be developed if a population is tested at multiple test sites across a wide range of climatic conditions. Most provenance tests comprise only a small number of test sites located in central distribution areas. Genecology functions or transfer functions are a better choice for these cases.

6. Universal response functions

6.1. Objectives

Genecology functions and transfer functions represent variation in the performance of populations caused by their origin (genetic effect). In contrast, response functions explain the variation in performance in response to differences in climatic conditions among test sites (environmental effect). Both have limitations in practice applications; a genecology function or a transfer function built for a test site may not apply to other environments. Similarly, a response function developed for a population may not apply to other populations. A universal response function (URF) is developed by [Wang et al. \(2010\)](#) to overcome these limitations by integrating both the genetic effect and the environmental effect into one model, as illustrated in Figure 6.

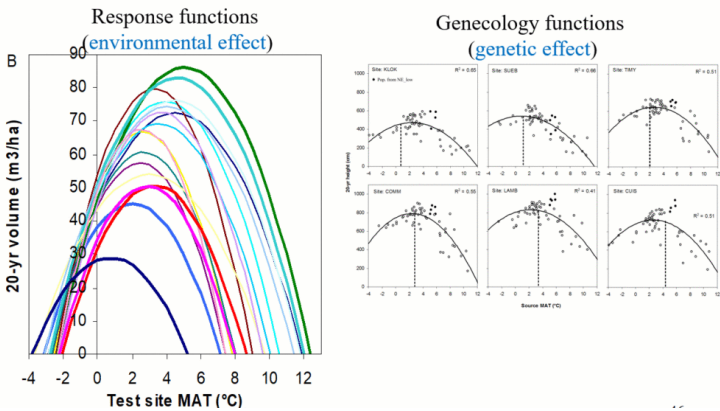


Figure 6. Illustration of integrating genealogy functions and response functions to build a universal response function.

6.2. Function

A universal response function is a combination of a genealogy function and a response function. A genealogy function can be written as

$$Y = b_0 + b_1x_1 + b_2x_1^2$$

Where, Y is the performance of a population; x_1 is one (or more) environmental variables of provenance; and b_0 , b_1 , and b_2 are parameters to be estimated to fit the model.

A response function can be written as

$$Y = b_0 + b_1x_2 + b_2x_2^2$$

Where, Y is the performance of a population; x_2 is one (or more) environmental variable of test site; and b_0 , b_1 , and b_2 are parameters to be estimated to fit the model.

Then, a universal response function can be written as

$$Y = b_0 + b_1x_1 + b_2cx_1^2 + b_3x_2 + b_4cx_2^2 + b_5x_1x_2$$

Where, Y is the performance of a population; x_1 and

x_2 are as defined above; x_1x_2 is the interaction between x_1 and x_2 ; $b_0, b_1, b_2, \dots, b_5$ are parameters to be estimated to fit the model.

Figure 7 shows a universal response function built for lodgepole pine from Wang et al. (2010). As the complete model involves multiple variables, the response surface is produced based on the two most important climate variables for visualization.

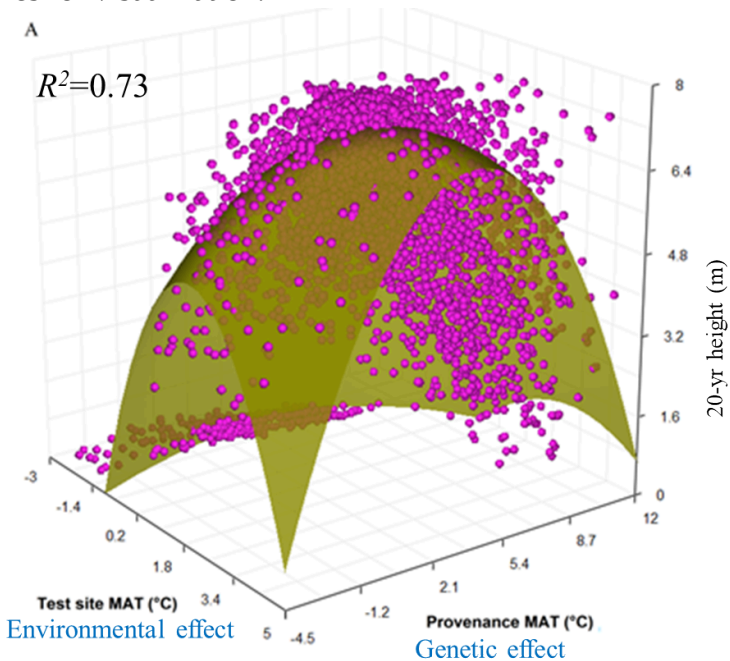


Figure 7. Response surface of a simplified universal function for lodgepole pine. Image by Tongli Wang under CC BY 4.0.

6.3. Features

Through the integration of both genetic and environmental

effects into a single model, a universal response function (URF) has the following unique features:

- 1) A URF model can predict the performance of a population from any climate and planted at any climate. In other words, it can predict the performance of any population planted anywhere.
- 2) Model accuracy is improved. Genecology functions or transfer functions reflect the genetic effect among populations, while response functions describe the environmental effect. The URF model integrates both the environmental and genetic effects, thus improves the model accuracy. The simplified version of the URF shown in Figure 7 explains 73% of the total variation in population performance, while the full model accounts for 83% (Wang et al. 2010).
- 3) Environmental and genetic effects on population performance can be quantified and compared. By taking partial derivatives respect to site climate and provenance climate, respectively, the rate of change along the site climate (environmental effect) and provenance climate gradients (genetic effect) can be calculated (Figure 8).

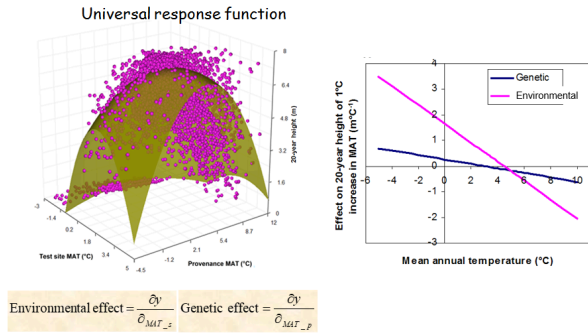


Figure 8. URF response curve and effects of genetic and environmental effects on population performance. Image by Tongli Wang under CC BY 4.0.

- 4) It relaxed the requirement for sample size. The URF takes better use of information obtained from a provenance test by including both site and provenance climates into the model; thus, it can achieve the same model accuracy by using much smaller sample sizes (Figure 9).

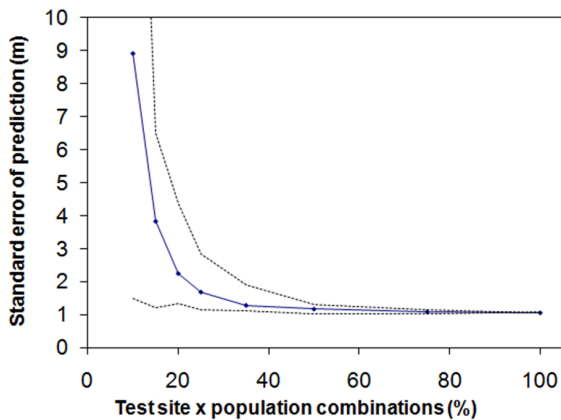


Figure 9. genecology functions

developed at six test sites from cold to warm climates. Image by Tongli Wang under CC BY 4.0.

- 5) A URF can identify optimal provenances for any planting sites. By setting the partial derivative function to zero, the optimal values of the provenance climate (i.e., corresponding to the peak of the response surface) can be determined, which can be used to locate the optimal provenances. Through comparisons of the difference in performance between local and optimal provenances, adaptational lag can be assessed ([Hu et al. 2019](#)).
- 6) A URF can be used to model the climatic niche of a species. The output of a climatic niche model is usually either in binary format (presence and absence) or in probability ($0 \sim 1$), while A URF is usually built with growth traits. It can be considered as absence if growth is zero and the rest as presence. As there is no interspecies interaction in common garden experiments for provenance tests; thus, the climatic niche predicted by a URF can be regarded as a fundamental niche, which is more relevant for the use in assisted migration (to be discussed in later topics). In addition, URF predictions covering the information for both climatic niche and growth potential. Figure 10 shows the climatic niche, growth potentials using local and optimal provenances of lodgepole pine for the current and future periods.

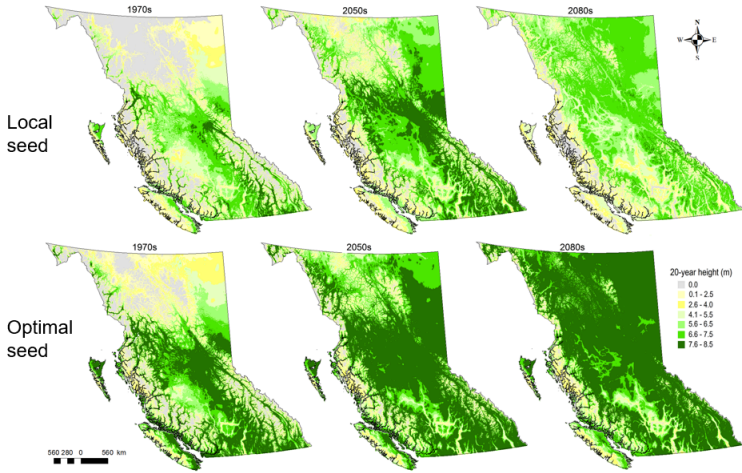
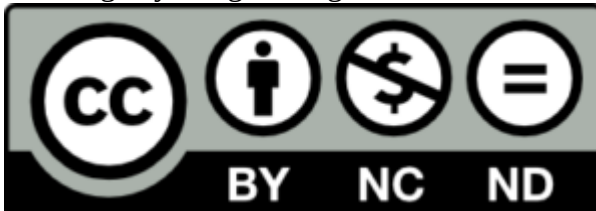


Figure 10. Climatic niche, growth potentials using local (the upper panel), and optimal (the lower panel) provenances of lodgepole pine for the current and future periods. Image by Tongli Wang under CC BY 4.0.



By Tongli Wang, Climate Modelling and Forest Applications is licensed under [Creative Commons <CC BY-NC-ND 4.0> License](https://creativecommons.org/licenses/by-nc-nd/4.0/)

Module V. Model Applications

Topic 5.1. Model Interpretation and Applications

Tongli Wang



Contents

[1 Model interpretation](#)

[1.1 Climatic niche models vs. species distribution models](#)

[1.2 fundamental niche vs. realized niche](#)

[2 Model applications](#)

[2.1 Predictions for the current period](#)

[2.2. Future projections](#)

[2.3 Understanding species responses to climate change](#)

[2.4. Providing a basis for resources management](#)

1. Model interpretation

A niche-based model can be interpreted as a species distribution model (SDM), an ecological niche model, or a climatic niche model. If it is considered as a niche model, it can also be interpreted as a fundamental niche model or

a realized niche model. Different interpretations can lead to different expectations and provide different implications for applications in mitigation and adaptation to climate change. Thus, how to interpret a model prediction is critical for model applications.

1.1 Climatic niche models vs. species distribution models

If a climatic niche model is considered as a SDM, and the model predictions for a future period likely face criticisms for not considering the following factors:

- The rate of migration of the species;
- Interactions with other species at new sites;
- Impacts of pests and diseases
- Interactions with micros in the soil
- Soil conditions;
- ...

Those criticisms are all correct as they all affect the species distribution. Thus, a niche based-model cannot predict the actual species distribution. What the model can predict is the spatial distribution of suitable climate conditions for the species. So, strictly speaking, a climatic niche model should not be called a SDM although it is often called so.

Criticisms on not considering:

- Rate of migration ✓
- Interactions with other plants ✓
- Impacts of pests and diseases ✓
- Interactions with micros in the soil ✓
- Among population variation ✓

Modelers' claim

- Species distribution models (SDM)
 - All criticisms valid

Figure 1. Criticisms are valid if a climatic niche model is called a species distribution model (SDM). Image by Tongli Wang under CC BY 4.0.

You may wonder, the species occurrence data, which are used to build the climate niche model, reflects the factors listed above, why the future predictions of the model do not consider these factors? It is true that these factors are considered in predictions for the current period, in which the predicted geographic areas are local or overlapping with the occurrence data points (Figure 2). For future or historical periods, predicted areas are, at least partially, in new locations, where the above-listed factors can be different. However, as these factors are also affected by climate, the effect of these factors may fade out in the long run.

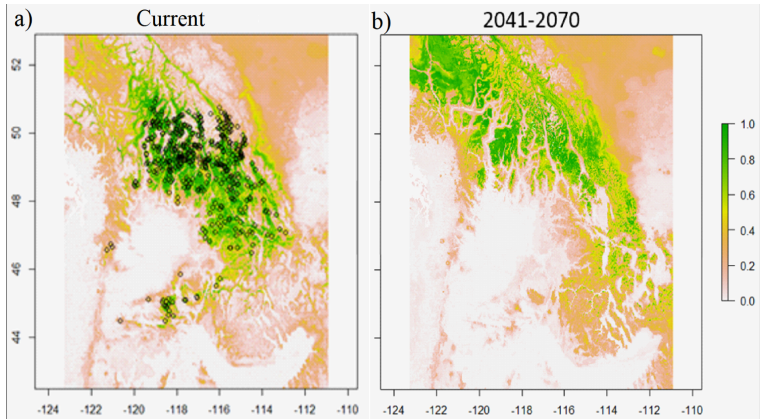


Figure 2. Model predictions for the current (a) and a future (b) period. Black diamonds in (a) are locations of presence data points. Image by Tongli Wang under CC BY 4.0.

Predicted distribution of the climatic niche of a species can serve as a baseline of the potential distribution of the species. Predictions of other factors, such as soil conditions, can be overlaid onto the baseline to refine the potential distribution of a species.

1.2 fundamental niche vs. realized niche

If a climatic niche model for a species is built based on presence and absence data with climate variables for a species, we consider the model predictions to represent the realized climatic niche (RCN) of the species (Figure 3). The RCN reflects the suitable climatic conditions supporting the current species distribution in the presence of interspecific competitions and other factors listed in Figure 1. If an area is predicted to be within the RCN, it indicates that this area has the potential to support the occurrence of the species if the species face the same interspecific competitions and the same other factors.

Those factors may not be available in the area and may have a negative or positive effect on the species. For example, if suitable soil conditions are not available there (e.g., a lake), the species cannot occur for sure. On the other hand, if competing species are absent, such as in plantations, the species may go well beyond its RCN. Again, as climate drives almost all other factors listed in Figure 1 in long run, those factors would eventually come into play. The majority of climatic niche models and so-called SDMs reported in the literature are in this category. Predictions of RCN for future periods indicate the potential occurrence of a species through natural processes. They can also be used to conservatively guide assisted migration of a species through plantations.

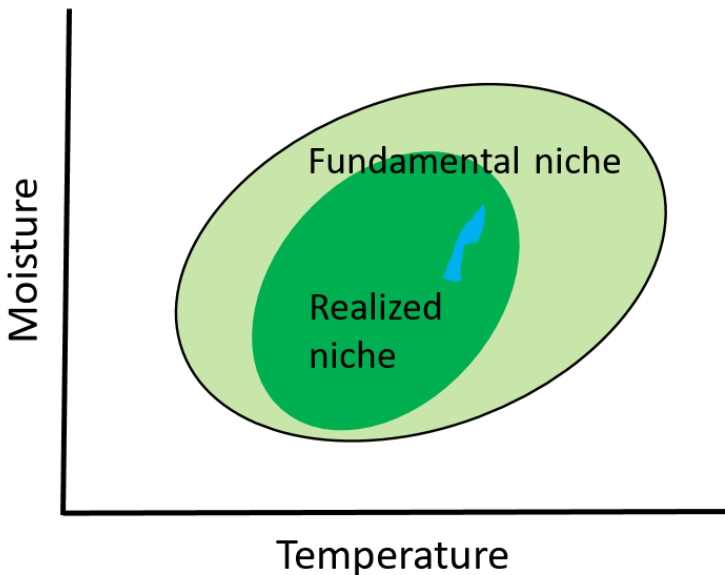


Figure 3. Illustration of fundamental niche and realized niche in two dimensions. Image by Tongli Wang under CC BY 4.0.

For a model that is built based on a comprehensive provenance test, such as a universal response function (URF), we also consider it as a CNM, but it predicts fundamental climatic niche (FCN) as illustrated in Figure 3. A FCN represents the ranges of physical conditions within which a species can survive, grow, and reproduce without the presence of interactions with other organisms. Predictions of FCN for future periods can provide information to guide assisted migration of a species through plantations, which do not involve interspecific competitions.

After a climatic niche model is built for a species, it can be used to predict the spatial distribution of the climatic niche for historical, current, and future periods, and even for paleo periods, depending on the climatic variables to be fed to the model. Predictions for different periods have different implications for forest resources management, including mitigation and adaptation to climate change.

As the future climate variables from the General Circulation Models (GCMs) are projected with uncertainty, the terms “projected” and “projections” are used to describe the output of climate niche models for future periods. On the other hand, “predicted” and “predictions” are used to refer to the output of climatic niche models for historical or the current period.

2.1 Predictions for the current period

Each tree species is adapted to a specific range of climate conditions – the climatic niche of a species. Understanding the climate niche of a species is essential for the resources management of the species, especially for formulating seed transfer guidelines and assisted migration under a rapidly changing climate. Predictions of a climatic niche model

for the current period can be used to map the spatial distribution of the climatic niche under the current climate conditions (Figure 4) and to define the climate niche in terms of major contributing climate variables. However, for most of the forest tree species, such information is still lacking although we know their geographic distributions. Thus, a climatic niche model and its predictions for the current period are important.

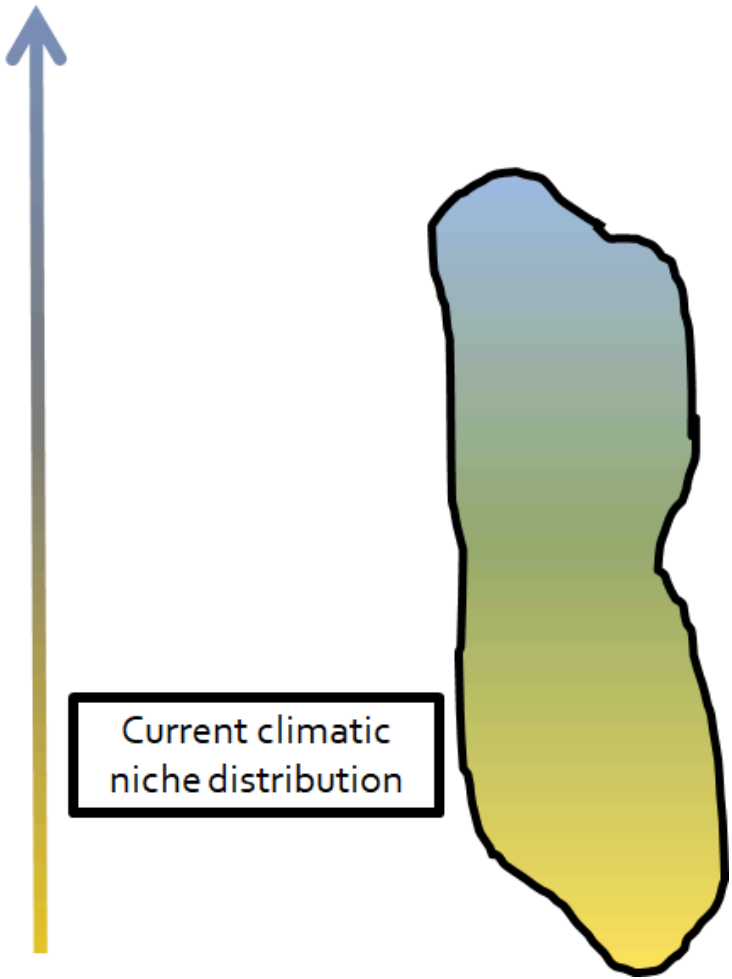


Figure 4. Illustration of geographic distributions of the climatic niche of a species under the current climate conditions. Image by Tongli Wang under CC BY 4.0.

2.2. Future projections

After predictions are mapped for the current period,

projections for a future period can be overlaid on the top of the current one, which allows comparing their changes over the current period (Figure 5). This overlay can provide a framework to: 1) assess the potential shift in the spatial distribution of a climatic niche from the current to a future period; 2) understand the responses of the species to climate change; and 3) develop forest adaptive strategies under a changing climate.

For a species in the north atmosphere, the projected distribution of its climate niche often shows a northward shift, as illustrated in Figure 5. It also always shifts upward in mountains. Changes in geographic coordinates, altitude, and areas from the current to a future period can be estimated. Of course, for future projections, it often needs to project for different future periods with various climate change scenarios.

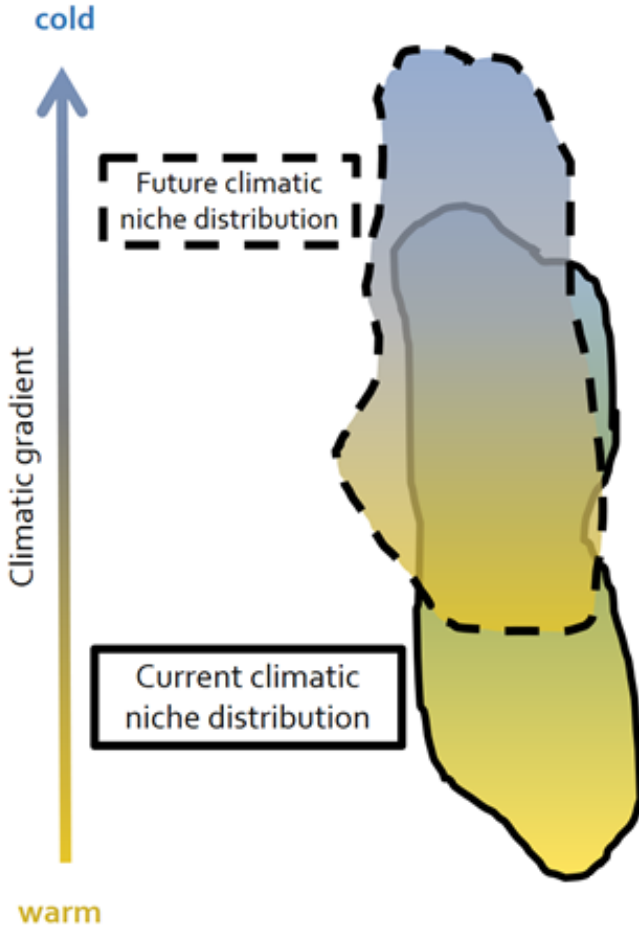


Figure 5. Illustration of geographic distributions of the climatic niche of a species under the current (solid outlines) and future (dashed outlines) climate conditions. Image by Tongli Wang under CC BY 4.0.

2.3 Understanding species responses to climate change

Trees or populations in different parts of a species

distribution may respond differently to climate change. The overlay of the future distribution of a climatic niche over the current distribution, as shown in Figure 5, can help us to understand the responses of a species to climate change in different parts of its distribution.

Trees on the leading edge of the distribution are likely to migrate to areas that are projected to be climatically suitable for the species under a future climate (Figure 6). However, the migration of the species into the newly available habitat will likely be lagging behind the shift in the spatial distribution of the climatic niche. This is because the rate of climate change is predicted to be far greater than the rate of species migration, which is affected by seed dispersal, physical barrier, soil, and other factors.

Trees in the lagging edge (or trailing edge) of the distribution are projected to be suffering from unfavorable climatic conditions as they would be outside of their climatic niche (Figure 6). A climatic mismatch will be a major challenge for these trees. Depending on the level of the mismatch, some populations in the lagging area may eventually be extirpated.

Trees in the areas, where the future distribution of the climatic niche overlaps over the current one (Figure 6), remain within their climatic niche in the future climate at the species level. Trees in these areas are less vulnerable to climate change than that in the lagging edge. However, they will no longer be in their optimal conditions at the population level, as illustrated in Figure 9 in Topic 1 of Module 4, and will be under selection pressure. Within- and among-population variation, as well as gene flow, may facilitate adaptation to a changing climate.

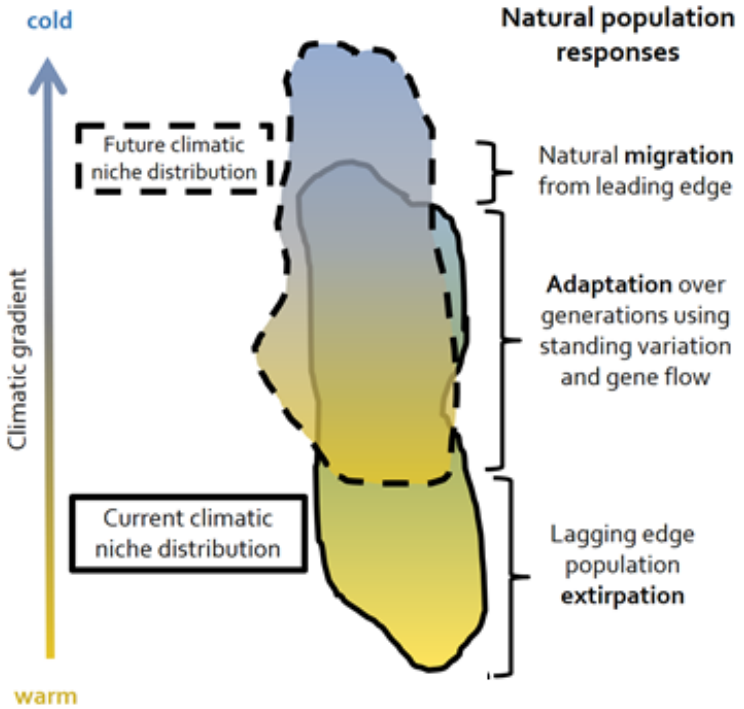


Figure 6. Potential responses of a species to climate change in different regions based on the overlapping status between the current and a future distribution of the climatic niche. Image by Tongli Wang under CC BY 4.0.

2.4. Providing a basis for resources management

A good understanding of species response to climate change based on the framework of climatic niche distributions (Figure 6) provides a scientific basis for forest resource managers to develop adaptive management strategies to mitigate the impact of climate change. Such strategies should reflect the characteristics of the different sections of the species distribution.

For the areas at the leading edge of the distribution, assisted migration is suggested to be a priority (Figure 7). As mentioned above, the rate of climate change is far greater than the rate of the natural migration of almost all tree species. When areas become suitable for a species, particularly for a favorable species, assisted migration may become necessary or beneficial to take the use of these areas.

For areas in the lagging edge of the distribution, *ex situ* conservation is considered important (Figure 7). Genetic variation existing in the populations should be conserved through seed collections in cold storage. Seed transfer to the overlapping areas for plantations is another way to conserve the genetic resources under a changing climate. Some silvicultural measures, such as thinning, may also be helpful for trees to mitigate the impact of climate change in these areas. Meanwhile, the introduction of new species should also be considered for better use of natural resources available in these areas. As trees in these areas are most vulnerable to climate change, monitoring of the stand dynamics might be necessary.

For the overlapping areas (Figure 7), *in situ* conservation is important to ensure that the species can adapt to a changing climate through evolution and local adaptation. As the trees are not in their optimal climatic conditions at the population level, their growth and productivity might be compromised, although survival may not be a concern. Thus, assisted migration at the population level, also referred to as assisted gene flow or seed transfer, may be necessary.

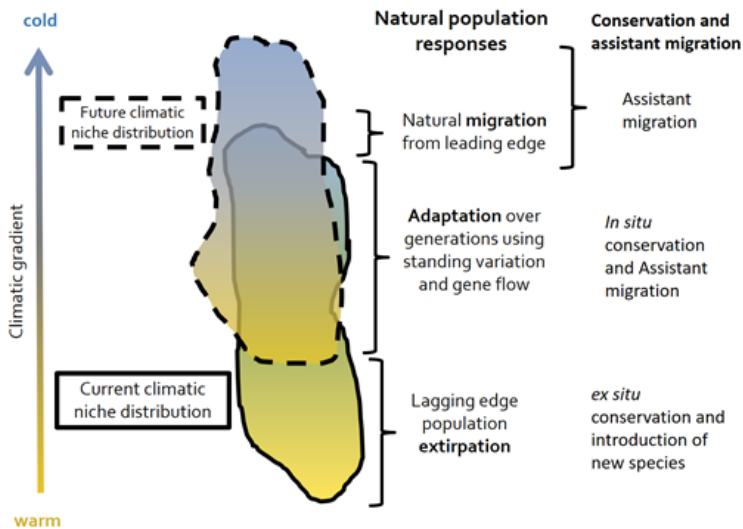


Figure 7. A screenshot of the file-open box to illustrate the selection of a raster input file. Image by Tongli Wang under CC BY 4.0.



By Tongli Wang, Climate Modelling and Forest Applications is licensed under [Creative Commons <CC BY-NC-ND 4.0> License](https://creativecommons.org/licenses/by-nc-nd/4.0/)

Topic 5.2. Case Studies with Forest Tree Species

Tongli Wang



Contents

[1 Douglas-fir](#)

[1.1 Climatic niche and future projections](#)

[1.2 Impact of climate change](#)

[1.3 Management strategies](#)

[2 Chinese fir](#)

[2.1 Climate niche and climate change impact](#)

[2.2 Limiting factors](#)

[2.3 Management strategies](#)

[3 Bluegum](#)

1 Douglas-fir

Douglas-fir (*Pseudotsuga menziesii*) is a major coniferous species in western North America. It is one of the most fast-growing forest tree species with high wood quality in

western North America. It also has a relatively low pest infection.



Figure 1. Douglas-fir trees in Vancouver, BC. Image by Korszun is licensed under CC BY 2.0

1.1 Climatic niche and future projections

Figure 1 shows its realized climatic niche for the current and three future periods. For the current period, predictions

were made from 10 Random Forest models and integrated according to the consensus among the 10 model's predictions.

The ranges of the climate niche for Douglas-fir are $-0.6 \sim 16.0^{\circ}\text{C}$ for mean annual temperature, $-13.8 \sim 10.8^{\circ}\text{C}$ for mean coldest month temperature, and $260 \sim 6040 \text{ mm}$ for mean annual precipitation. The climate niche is not likely to change although its geographical distribution will change in future climates.

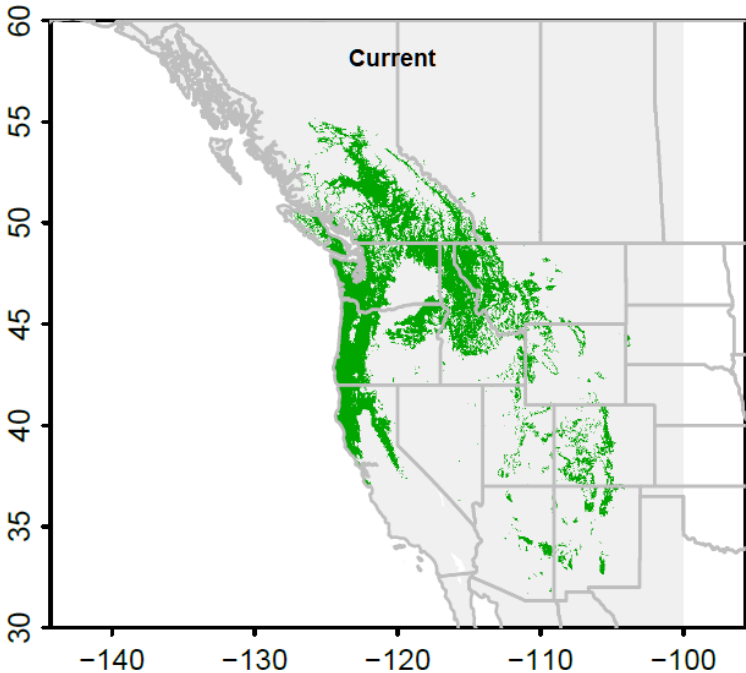


Figure 2. Geographic distributions of climate niche for Douglas-fir based on multi-model predictions for the current (1961-1990). Image by Tongli Wang under CC BY 4.0.

1.2 Impact of climate change

Figure 3 shows the consensus projections for the three future periods 2020s, 2050s, and 2080s. For each of the future periods, each of the 10 models was used to project the futures with 15 GCMs for a moderate climate change scenario RCP4.5, and the final projections were based on the consensus among the 150 (10 x 15) projected layers. The climatic niche is projected to expand northward at the leading edge (Figure 1) and upwards (not easily visible). However, contractions are not apparent at the trailing edge, which is special as most species in this region are projected to shift northward.

Overall, projected changes are 1.4 degrees of increase in latitude and 89 m in elevation of the distribution. The total area of suitable climate for this species will increase by about 10%.

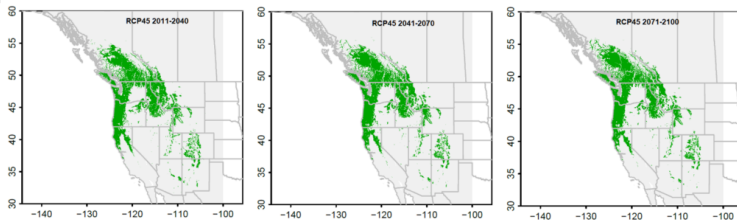


Figure 3. Geographic distributions of climate niche of Douglas-fir for the three future periods 2020s, 2050s, and 2080s using consensus projections based on the climate change scenario RCP4.5. Image by Tongli Wang under CC BY 4.0.

1.3 Management strategies

As Douglas-fir is a favorable forest tree species, it is

desirable to plant this species as much as where its suitable climate conditions become available. Thus, assisted migration of this species through plantation can be an effective way to take the advantage of a warming climate in British Columbia. The climate-based seed transfer system (CBST) recently launched in this province would facilitate the management of genetic resources.

2. Chinese fir

Chinese fir is one of the most socially and economically important coniferous species in China. This species is prized for its durable wood and cultural and ornamental value. It is the species of choice for building temples, furniture, houses, flooring, and boats. Chinese fir trees occupy 30% of all plantations in China and account for 25% of national commercial timber production.



Figure 3. Chinese fir forest and logs. Image by Clive Perrin is licensed under CC BY-SA 2.0.

2.1 Climate niche and climate change impact

The climate niche of Chinese fir spans $9.2 \sim 22.6^{\circ}\text{C}$ in mean annual temperature and $830 \sim 3060 \text{ mm}$ in mean annual precipitation. It is geographically distributed in southern China (Figure 4). Climate change is projected to have a drastic impact on this species.

Future projections of its climate niche suggest that the geographic distribution for Chinese fir may substantially contract under predicted climate change (Figure 4). The contraction of its climate niche distribution in the south is not a surprise as it is common phenomenon in forest trees distributed in the northern hemisphere. However, it is shocking to see the lack of a predicted northward expansion for this widespread species, which is the case for most forest tree species in the northern hemisphere, such as Douglas-fir. As a result, the climate niche distribution of Chinese fir is expected to decrease by 34% by 2050 and more than 50% at the end of this century.

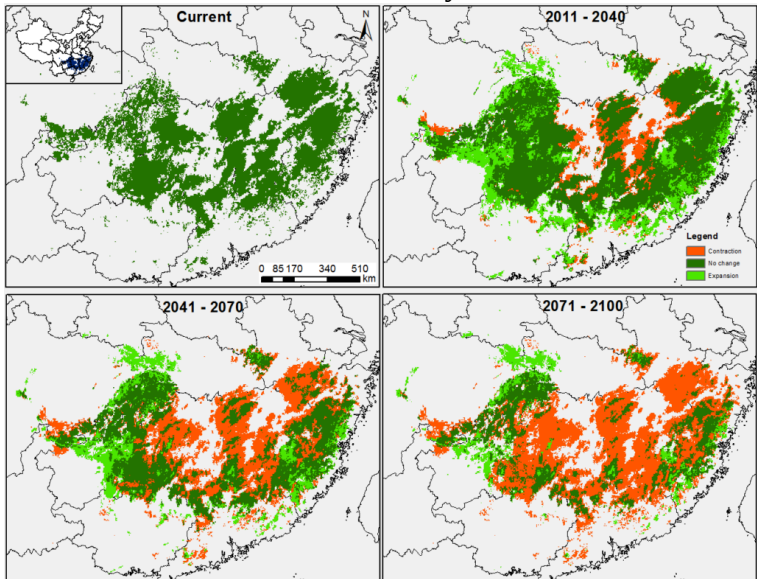


Figure 4. Geographic distributions of climate niche of Chinese fir for the current and three future periods 2020s, 2050s, and 2080s using consensus projections. Image by Tongli Wang under CC BY 4.0.

2.2 Limiting factors

The limited northward expansion is likely attributable to the spatial pattern of precipitation in China. Current species distribution and its predicted current climatic niche occur in areas with mean annual precipitation (MAP) above 830 mm. Under future climates, the geographic distribution of temperatures suitable for this species is projected to move northward; however, the geographic distribution of mean annual precipitation is projected to remain about the same as for the reference period (Figure 5). The areas with MAP above 830 mm are mostly south of the Yangzi River. The sharp contrast in precipitation between northern and southern China appears to play an important role in the direction of future projections and may limit the northward expansion of this species.

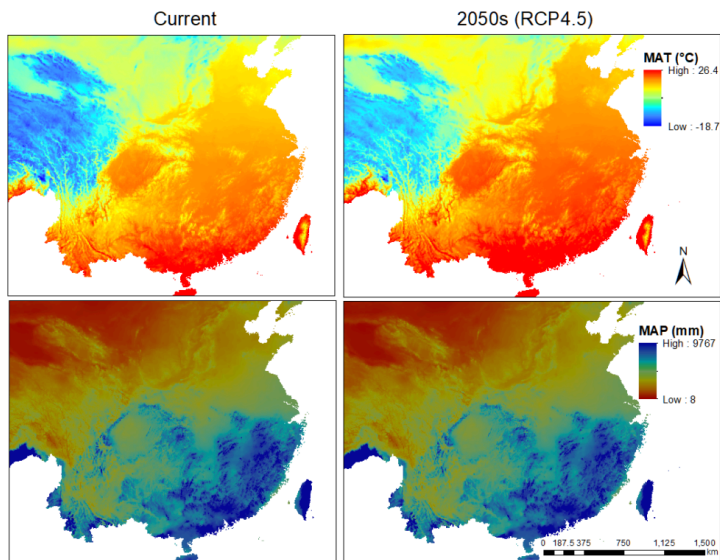


Figure 5. Spatial patterns of mean annual temperature (MAT) and mean annual precipitation (MAP) for the current and the 2050s. Image by Tongli Wang under CC BY 4.0.

2.3 Management strategies

It is worth noting however that the climatic niche modeled in this study (or in the vast majority of other studies) is a realized climatic niche. It reflects the climate conditions currently occupied by the species. Whether or not Chinese fir can grow outside of its realized niche but still within its fundamental niche is unknown and needs to be further explored through field experiments, such as provenance tests.

The dramatic decline in the area suitable for Chinese fir projected in this study is concerning as this species is one of the most important forest species in China in

terms of both its economic value and role in ecosystem functioning. Results shown here should provide an early warning for policymakers and practitioners to develop adaptive strategies in species selection and forest management practices in order to adapt the management of Chinese fir forests/plantations to future climate. Managing this species within its current geographical distribution will become increasingly problematic as its current climatic niche becomes fragmented which may lead to a loss in the economic value of plantations and impair forest ecosystem functioning.

Genetic conservation should be a high priority for this species. *In situ* conservations should be established through parks and reserved areas in areas projected to be climatically suitable for this species in the future. For areas projected not to be suitable for this species, *ex situ* conversation should be planned, including seed storage, clone banks, and assisted migration.

3. Bluegum

Bluegum (*Eucalyptus globulus* Labill.) is a major plantation tree species in Australia. Another reason for including this species as a case study here is that it is distributed in the southern hemisphere. Therefore, the shift in the climatic niche of this species is southward instead of northward.

The geographic distribution of the climatic niche for blue gum is projected to shift southward towards cooler conditions. The poleward shift will be constrained by the Tasman Sea at the leading edge and thus result in a contraction of its total distribution under future climates (by 24% in the 2050s) (Figure 6).

The contraction in the overall climatic niche of blue gum on the mainland of southeast Australia could have negative implications for the productivity and vitality of blue gum plantations into the future but positive outcomes for plantation and forest managers in Tasmania. However, the adaptive capacity of eucalypt plantations is considered high with many eucalypts grown on short rotations (<10 years). This provides managers with opportunities to adapt their silvicultural practices and/or plant different genotypes or species to match changing climatic conditions with relative ease.

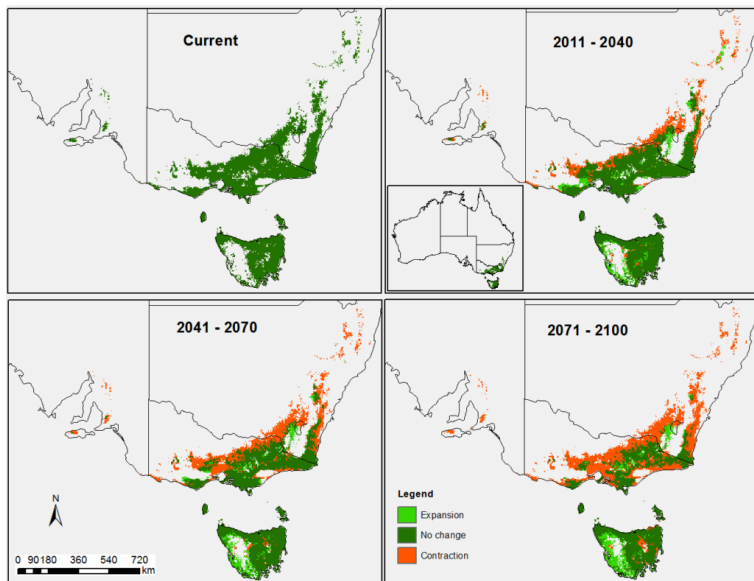


Figure 6. Geographic distributions of climate niche of blue gum for the current and three future periods 2020s, 2050s, and 2080s using consensus projections. Image by Tongli Wang under CC BY 4.0.



By Tongli Wang, Climate Modelling and Forest Applications is licensed under [Creative Commons <CC BY-NC-ND 4.0> License](https://creativecommons.org/licenses/by-nc-nd/4.0/)

Topic 5.3. Case Studies with Forest ecosystems

Tongli Wang



Contents

[1 Biogeoclimatic ecosystems in British Columbia](#)

[1.1 Climate niche modeling](#)

[1.2 Uncertainty of future projections](#)

[1.3 Consensus projections](#)

[1.4 Impacts of climate change on BEC zones](#)

[1.5 Management implications](#)

[2. Forest ecosystem in China](#)

[2.1 Climate niche modeling](#)

[2.2 Future projections](#)

[2.3 Management for adaptation to climate change](#)

1. Biogeoclimatic ecosystems in British Columbia

The forest ecosystem in British Columbia, Canada, is defined through the [Biogeoclimatic Ecosystem Classification](#) (BEC) system. The BEC system divides the province into 16 ecological zones (BEC zones) that reflect terrestrial ecosystem differences along large-scale climate gradients related to changes in altitude, latitude, and continentality (Figure 7). These ecological zones are subdivided into subzones and variants, reflecting plant community composition and structure differences along finer-scale climate gradients. It is probably the best-defined forest ecosystem in the world and serves as the foundation for forest resources management and genetic conservation in the province.

Each ecological unit represents a plant community, thus predictions of the climate niche supporting a plant community may serve broader objectives than predictions of individual species.

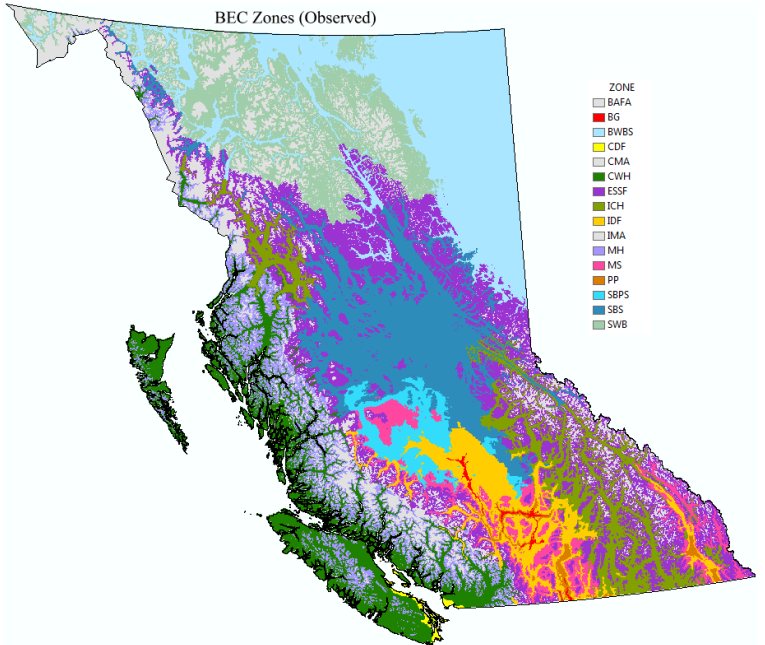


Figure 7. Geographic distributions of ecological zones in British Columbia. Image by Tongli Wang under CC BY 4.0.

1.1 Climate niche modeling

The climate niche of each BEC zone was modeled based on the relationships between the spatial variation of climate variables and the location of the BEC zone (Figure 8). The model was built using the Random Forest algorithm with scale-free climate data generated by ClimateBC.

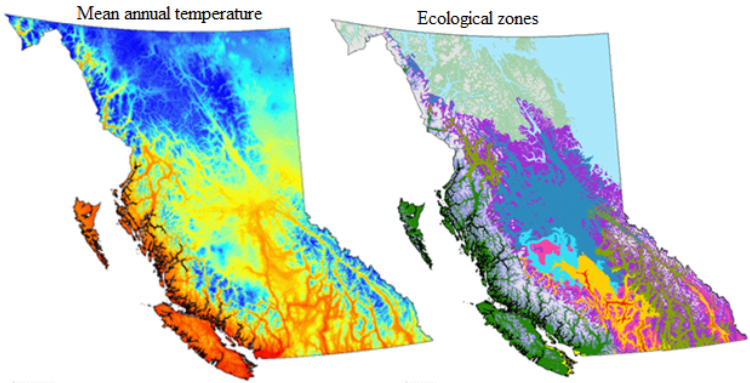


Figure 8. Geographic distributions of mean annual temperature and ecological zones in British Columbia. Image by Tongli Wang under CC BY 4.0.

Predicted climate niches of the ecological zones match well with the mapped zones, indicating high accuracy of the model predictions.

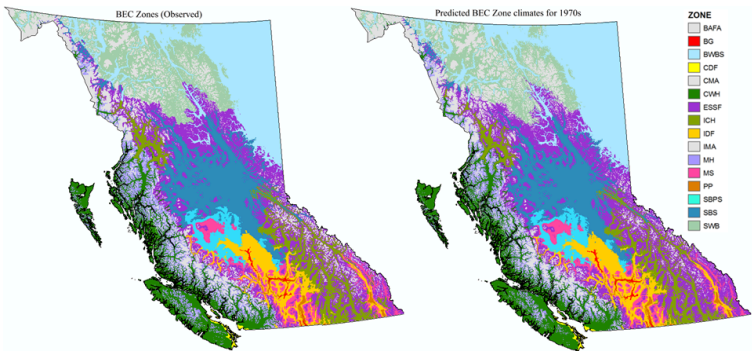


Figure 9. Geographic distributions of mapped ecological zones (left) and predicted climate niches of these zone in British Columbia. Image by Tongli Wang under CC BY 4.0.

1.2 Uncertainty of future projections

Large uncertainty exists in projecting future climates.

Different general circulation models (GCMs) with different climate change scenarios provide a different projection for a future period. For example, 134 GCM-scenarios combinations show a wide range of projected changes by the 2050s (Figure 10). It suggests that random or systematic selection of few GCM-scenario combinations cannot well represent the distribution of future projections. For the projections of BEC zones, 20 combinations (squares) were selected.

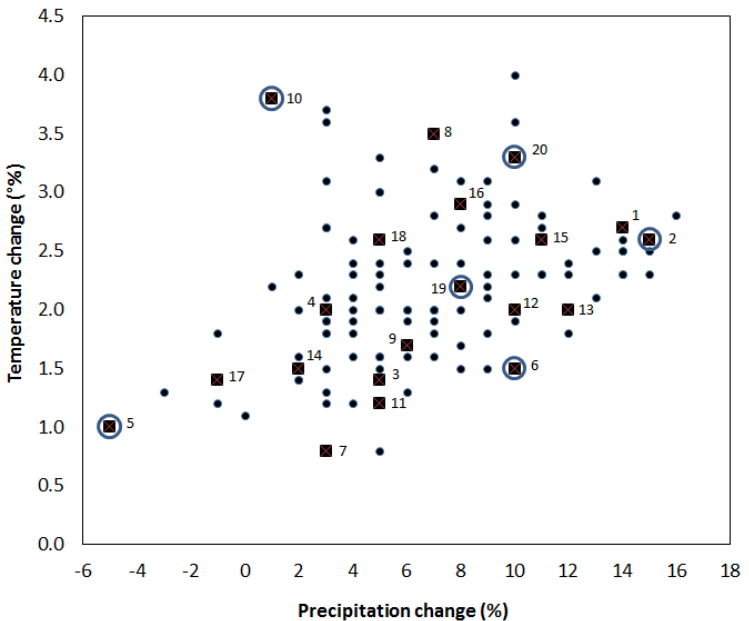


Figure 10. Distributions of projected change in temperature and precipitation from 134 GCM and scenario combinations. Squares are selected for BEC zone projections. Squares with a circle are used to illustrate the impact on BEC zone projections in Figure 11. Image by Tongli Wang under CC BY 4.0.

Six GCM-scenarios are also selected to illustrate the difference in future projections of the BEC zones for the same period (the 2050s) in Figure 11. Differences in the

projected impact of climate change are substantial, which would lead to very different management recommendations.

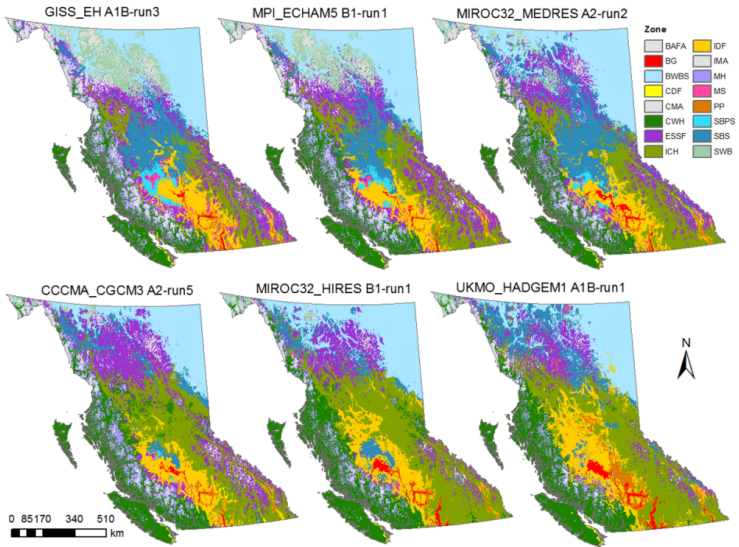


Figure 11. BEC zone projections for the 2050s with six contrasting GCM-scenario combinations, arranged from most conserve to most aggressive climate change scenarios. Image by Tongli Wang under CC BY 4.0.

1.3 Consensus projections

An effective way to consider the uncertainty is to take a consensus projection from a large number of individual projections as shown in Figure 12.

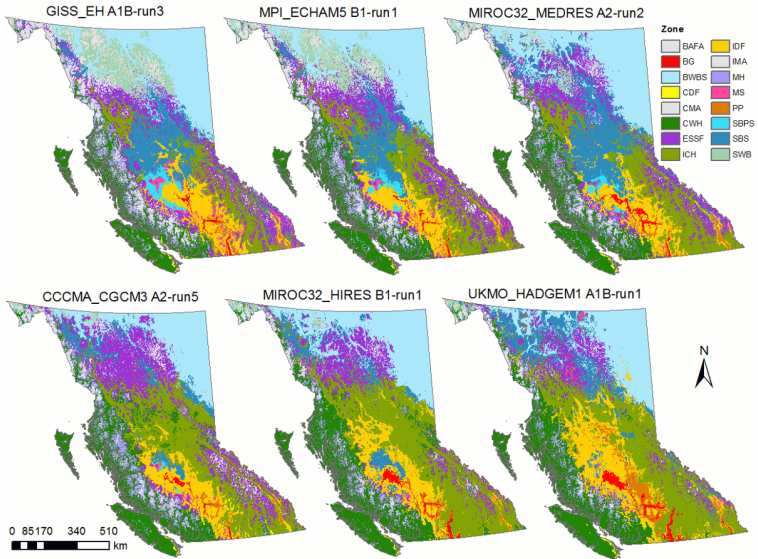


Figure 12. A consensus projection from individual projections. Image by Tongli Wang under CC BY 4.0.

After taking the consensus projections, there is one map, instead of 20 maps, for each period (Figure 13), which best represent the distribution of all projections.

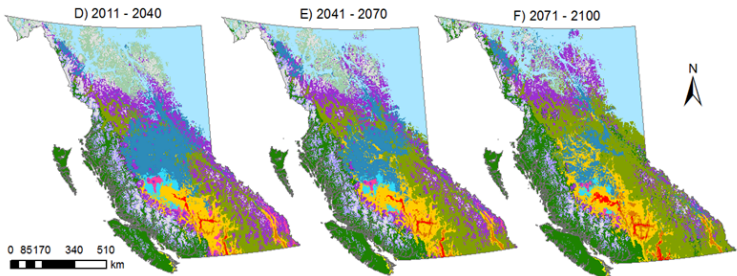


Figure 13. Consensus projections of BC BEC zones for 2011-2040, 2041-2070 and 2071-2100. Image by Tongli Wang under CC BY 4.0.

1.4 Impacts of climate change on BEC zones

Impacts of climate change on BEC zones vary from zone to zone (Figure 14). The ICH, PP, IDF, CWH, BWBS, and BG zone are projected to expand. The expansion of the ICH zone is dramatic, not only in percentage (Figure 14) but also in absolute areas (Figure 13). Forest tree species associated with these zones include western redcedar, western hemlock, interior Douglas-fir, and ponderosa pine.

On the other hand, all the other zones are projected to contract. The contraction is most serious in the MS, SBPS, SWB, and AT zone. Species associated with these zones are interior spruce, lodgepole pine, and subalpine fir.

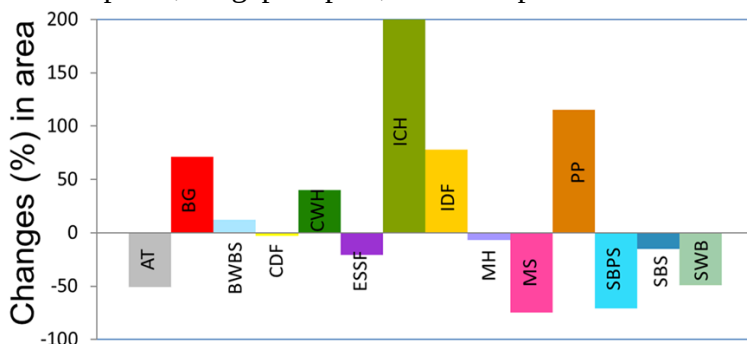


Figure 14. Changes in percent area (%) of each BEC zones for 2041-2070. Image by Tongli Wang under CC BY 4.0.

1.5 Management implications

The ecological units in British Columbia are the foundation of species selection for plantations. Projected shift and expansion/contraction provide a scientific basis for the consideration of climate change in forest resources management. For zones in expansion, more seed orchards

may need to be planned for their associated species, such as western redcedar, interior Douglas-fir, and ponderosa pine. On the other hand, it should be cautious to plan new seed orchards for interior spruce.

In most parts of British Columbia, low temperature is a limit to forest tree growth. Thus, climate change brings in challenges, but also opportunities. The expansion of suitable climate conditions for fast-growing species, including western redcedar, Douglas-fir, western larch, and ponderosa pine, is a clear indication. To take these opportunities, assisted migration is needed to plant these species in their potentially suitable areas progressively.

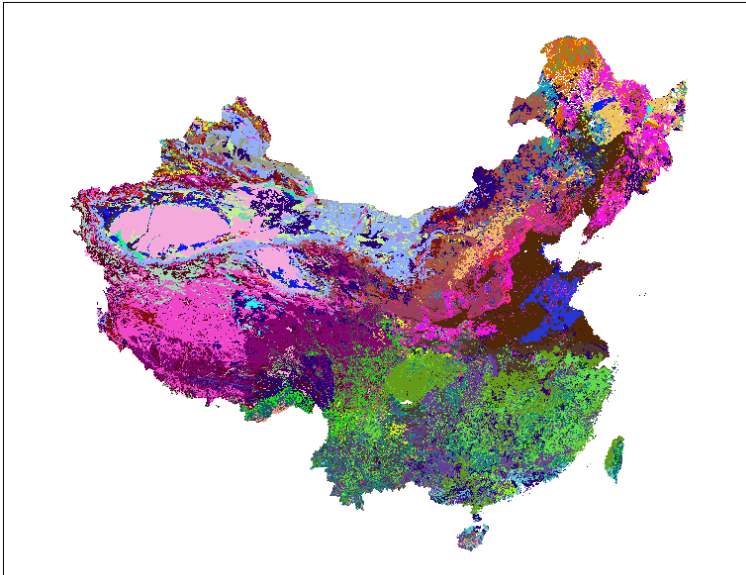
Please note that the above-mentioned projections are at the zone level. Projections at the subzone-variant level are also available at http://www.climatewna.com/ClimateBC_Map.aspx.

2. Forest ecosystem in China

There are about 3,000 tree species in China. The number of commercial tree species is also over hundreds. To model the species distributions of all the species and project their shifts in future climates are unrealistic. To use forest ecotypes as surrogates, as shown for BEC zones, can be particularly effective.

There is no BEC equivalent forest ecosystem classification system in China. Instead, there are four levels of vegetation classifications available in the digital archive of Chinese Vegetation. The level III vegetation classification, with 56 ecotypes, has the closest match to the level of BEC zones in British Columbia, thus was used to represent the forest ecotypes of the country as shown in Figure 9. The 56 ecotypes represent the vegetation

communities and the macroclimate conditions across China. The typical forest ecotypes include broadleaf deciduous forests in the temperate zone (#3), needle-leaf forests in the subtropical zone (#31), broadleaf evergreen and deciduous scrubs in the subtropical and tropical zone (#44). The major ecotypes for none forested areas include: Two years three ripens or one year two ripens grain fields and deciduous orchards (#27, dark brown), One year two ripens or three ripens grain fields; evergreen orchards and subtropical economic tree plantations (#21, dark green), Semi-shrub and dwarf semi-shrub deserts (#24), and grass-carex spp. high-cold steppes (#31).



Ecotype code



Figure 9. The 56 ecotypes at the level of 3 vegetation classification in China. Image by Tongli Wang under CC BY 4.0.

2.1 Climate niche modeling

The climate niche model was developed using a machine-learning algorithm (Random Forest) based on the occurrence of the 56 ecotypes and the climate variables generated using ClimateAP. The vectors of the ecotype distribution map were rasterized into gridded pixels at the spatial resolution of 800 x 800 meters. ClimateAP was then used to generate annual and seasonal climate variables (69 in total) for each of the pixels for the reference period

1961-1990. The ecotype of each pixel was used as the dependent variable and the 69 climate variables of the same pixel were used as the independent variables to develop the climate niche model. The climate niches of the 56 ecotypes predicted by the model (Figure 10) matched well with the original ones shown in Figure 9.

2.2 Future projections

The consensus projections for the forest ecosystem in China are shown in Figure 10. These maps have also been uploaded to a [web platform](#) for spatial visualization. There is a clear trend of northward shifting in general. However, some ecotypes are projected to expand substantially, including the “multiple ripens grain fields and deciduous orchards” (#27, dark brown) in Mideast China, and the “multiple ripens grain fields and evergreen orchards and subtropical economic plantations” (#21, dark green) in the south. On the other hand, the areas of broadleaf deciduous forests in the temperate zone (#3, pink) would decline and shift northward. The areas of the broadleaf evergreen and deciduous scrubs in the subtropical and tropical zone (#44) would also decline and shift westward. Such shifts and expansion may have a dramatic impact on land use and forest resources management in the future.

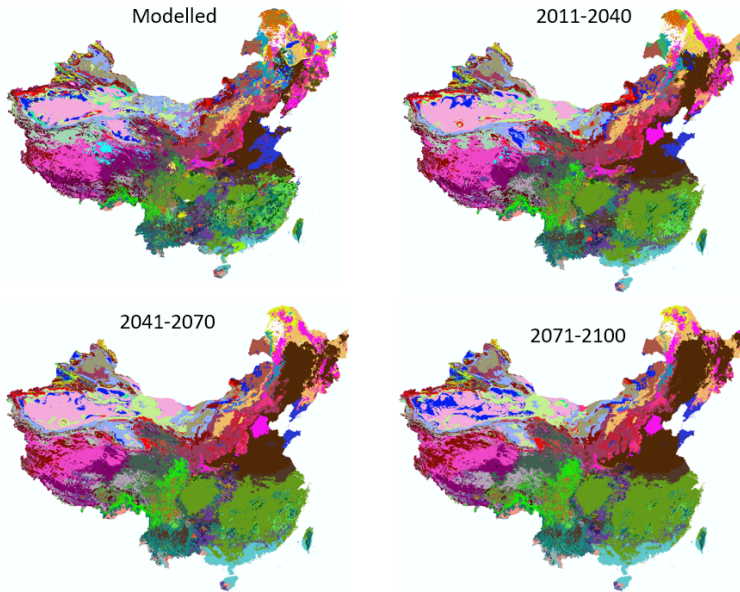


Figure 10. Consensus projections of the forest ecosystem classifications in China based on 15 GCMs and 2 greenhouse emission scenarios (RCP 4.5 and RCP 8.5). Refer to Figure 9 for the color codes. Image by Tongli Wang under CC BY 4.0.

2.3 Management for adaptation to climate change

Ecotypes represent plant communities and have their unique season-dependent management habits. The shifts of major ecotypes in China demonstrated the impact of climate change on macroclimate conditions and potentially the shifts of major vegetation communities. However, not like the BEC system in British Columbia, the ecotypes in China are not directly associated with individual tree species. Thus, to make use of the ecosystem as a vehicle to manage individual species as groups, the association between the ecotypes and species occurrence need to be established. A matrix with ecotypes as rows and species as

columns will serve such an objective. With such a matrix, the projections shown in Figure 10 can be easily translated into species distribution maps for the current and future periods.

The substantial expansions of multiple ripens grain fields favor agriculture if the newly climatically suitable areas are in cultivatable conditions. If these areas fall into mountains only suitable for silviculture, selecting suitable tree species for these areas will be a new challenge as the matrix mentioned above would not apply to these zones. If suitable species are selected, these areas may have a good potential for high productivity of forest products.



By Tongli Wang, *Climate Modelling and Forest Applications* is licensed under [Creative Commons <CC BY-NC-ND 4.0> License](https://creativecommons.org/licenses/by-nc-nd/4.0/)

Module VI. Building Ecological Models in R

Topic 6.1 Building niche-based models - 1/3

Tongli Wang

Contents

1 Overview

2 Some basics of R

[2.1 R programming language](#)

[2.2 Programming features](#)

[2.3 Packages and libraries](#)

[2.4 Download and install R](#)

[2.5 Create a working directory](#)

[2.6 Some basic operations in R](#)

[2.7 Install additional packages](#)

3 Species occurrence data preparation

[3.1 Importing occurrence data](#)

[3.2 Data inspection and cleaning](#)

[3.3 Absence data adjustments](#)

4 Climate variables generation

[4.1 Generating climate variables for specific sample locations](#)

[4.2 Generating spatial climate variables](#)

5 Obtaining species occurrence data

[5.1 Data download](#)

[5.2 Data processing](#)

[5.3 Adding absence data points](#)

1. Overview

All the modeling and mapping work covered in this course can be done in R. However, ArcGIS Map is sometimes easier to use to generate maps with some specific features. As experience in R is not a prerequisite for taking this course, we will show you the basic skills for using R to conduct niche-based modeling and mapping step-by-step in this lecture.

Many modeling algorithms have been used in niche-based modeling. Of those modeling algorithms, machine-learning approaches have stood out in performance. Among the machine-learning algorithms, Random Forest and MaxEnt are most widely used for their model accuracy and easiness to use. We will cover these two modeling algorithms in this lecture.

2. Some basics of R

2.1 R programming language

R is a programming language and free software environment for statistical computing and graphics supported by the R Foundation for Statistical Computing. The R language is widely used among statisticians and data miners for developing statistical software and data analysis.

2.2 Programming features

R supports matrix arithmetic. R's data structures include vectors, matrices, arrays, data frames and lists.

R supports procedural programming with functions. Some functions are specific to a class of object; there are also generic functions that work for almost every class of object, such as *print*.

2.3 Packages and libraries

The capabilities of R are extended through user-created packages, which allow specialized statistical techniques, graphical devices, import/export capabilities, reporting tools, etc. These packages are developed primarily in R, and sometimes in Java, C, C++, and Fortran.

A package is a collection of R functions, data, and compiled code. The location where the packages are stored is called the library. If there is a particular functionality that you require, you can download the package from the appropriate site, and it will be stored in your library.

2.4 Download and install R

- 1) Download R from one of the following links:
<https://mirror.rcg.sfu.ca/mirror/CRAN/bin/windows/base/R-4.1.2-win.exe>
- 2) Click to run the exe file “R-4.1.2-win.exe”. Follow the instruction on the screen to complete the installation.

2.5 Create a working directory

- 1) Create a folder “NicheModels” under “My Document”, and click on it to access the folder.
- 2) Create a new Text Document “0.Rdata” by

right-click inside the folder, as shown in Figure 1. The “0.Rdata” file should have an R logo on the left (Figure 2).

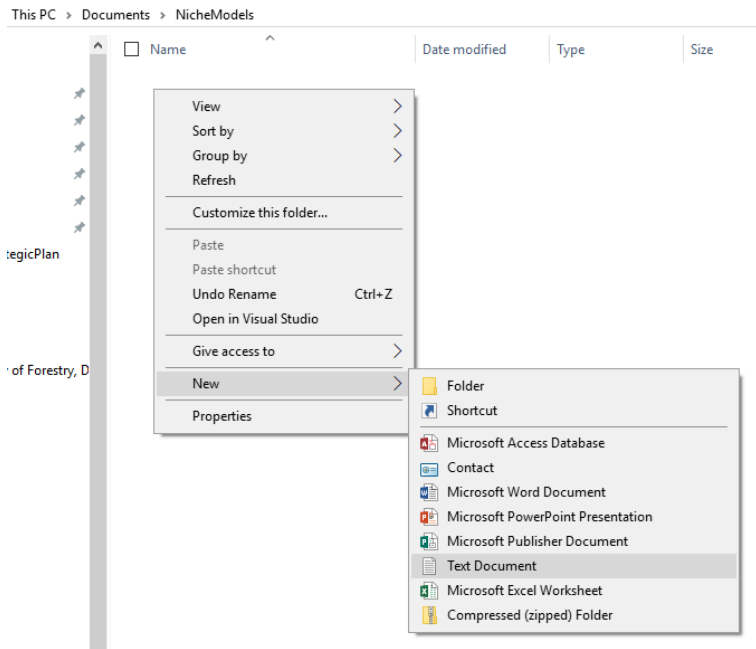


Figure 1. Illustration of creating a Text Document in the folder “NicheModels”. Image by Tongli Wang under CC BY 4.0.

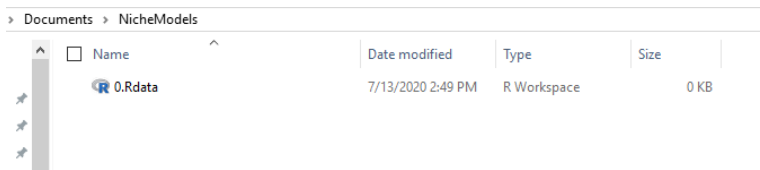
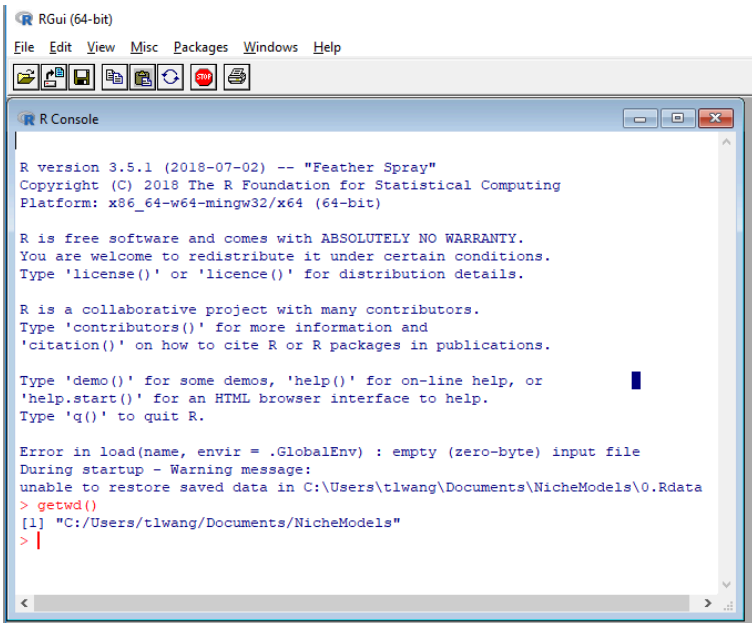


Figure 2. A R starting file with R logo. Image by Tongli Wang under CC BY 4.0.

- 3) Double click “0.Rdata” file to start R with a Graphical User Interface (GUI). The GUI can

also be started from the Start menu or Taskbar. However, using “0.Rdata” makes the current folder be the default working directory in R, as shown in Figure 3, which is convenient.



```

RGui (64-bit)
File Edit View Misc Packages Windows Help

R Console

R version 3.5.1 (2018-07-02) -- "Feather Spray"
Copyright (C) 2018 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

Error in load(name, envir = .GlobalEnv) : empty (zero-byte) input file
During startup - Warning message:
unable to restore saved data in C:\Users\tlwang\Documents\NicheModels\0.Rdata
> getwd()
[1] "C:/Users/tlwang/Documents/NicheModels"
> |

```

Figure 3. The interface of the R console showing the default working directory. “getwd()” is the function to show working directory. Image by Tongli Wang under CC BY 4.0.

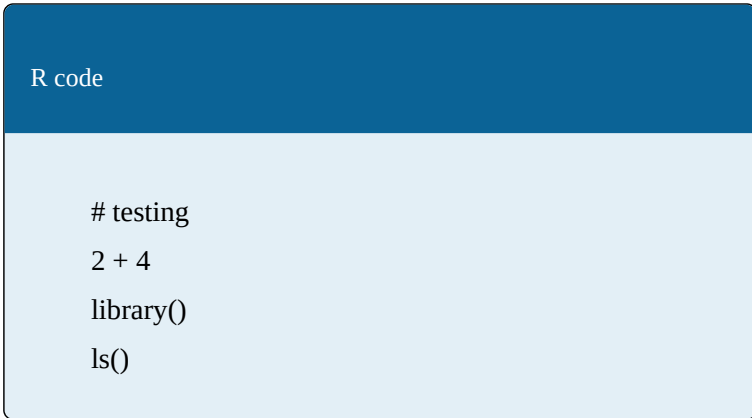
2.6 Some basic operations in R

In the console, type “getwd()” (a function) and hit Return, it displays the current working directory (Figure 3). To change the working directory, use the command “setwd(“your directory”)”

The function “library()” lists the packages included in the R package. Each of them may have many functions.

“ls()” function is to show objectives in the memory. It should be zero for now.

These commands can be used in a script. To open a new script, click “File”, then “New script” to open the R code editor, as shown in Figure 4. Type the following lines in the script window:



```
R code

# testing
2 + 4
library()
ls()
```

Note: # sign is for comments. Any code following # in the same line is not to be executed

Click “File” and “Save as”, and save the file as “test.R”.

Then, click “Edit” and “Run all”.

You can also run a specific line or multiple lines by selecting the lines and hitting “Ctrl-r”.

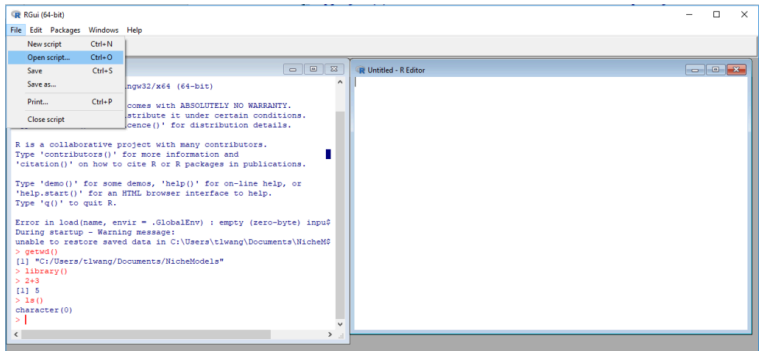


Figure 4. Illustration to open a new script window. Image by Tongli Wang under CC BY 4.0.

2.7 Install additional packages

There are many functions that are included in R. For our niche-based modeling, however, we need to install the following additional packages:

“randomForest” — for model building using Random Forest

“maps” — for showing boundaries of countries or regions.

“mapdata” – for supporting maps.

“raster” – for raster manipulations and mapping. It may require additional packages to run. You can follow the same steps to install them.

To install these packages in the R GUI interface, click “Packages”, then “Install package(s)”, and follow the on-screen guidance to install the packages (Figure 5).

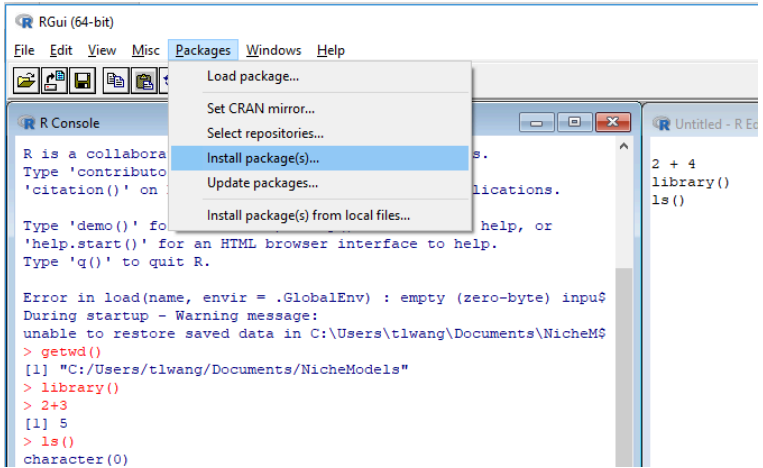


Figure 5. Installing packages through the R GUI interface. Image by Tongli Wang under CC BY 4.0.

Packages can also be installed using a command line in the script:

```
install.packages('raster').
```

To use an installed package "raster", the following command line is required to load the package:

```
library(raster)
```

3. Species occurrence data preparation

Data preparation involves three steps, including 1)

importing species occurrence data into R; 2) data inspection and cleaning; and 3) generating climate variables for model fitting and for spatial predictions.

3.1 Importing occurrence data

We have prepared a sample species occurrence dataset for this lecture. Everyone needs to download this dataset from the following link:

<http://climatena.ca/downloads/courseData/speciesA2.csv>

and put it into your working directory “NicheModels”. Then follow the steps below:

1. Click “0.Rdata” to start R.
2. Start a new script and save it as “1.dataSteps.R”.
3. Type the following lines in the script (please note that there is no additional package is needed):

```
## data steps for niche modeling
getwd() # to check the working directory

#import data —
dat <- read.csv('speciesA2.csv'); #import data
head(dat);dim(dat); #display the first 5 rows and
dimensions
```

It shows that the dataset has 2003 rows and 4 columns with coordinates (x=longitude, y=latitude), elevation, and species occurrence (presence=1 and absence=0).

3.2 Data inspection and cleaning

Errors often occur in species occurrence datasets. It is necessary to inspect the dataset before moving to the next step. Spatial visualization is probably the most effective way to identify errors. We will do this in R with the following code:

Exercises

```
#data inspection —  
library(maps)  
library(mapdata)  
  
#seperate presence and absence data —  
pr <- subset(dat,speciesA==1);head(pr);dim(pr)  
ab <- subset(dat,speciesA==0);head(ab);dim(ab)  
  
#plot presence data points on the map —  
plot(y~x,data=pr,col='blue',pch=19)  
map('world',add=TRUE)
```

The locations of the presence data points are shown in

Figure 6. Now we can see that the four points on the west end of the map are likely errors as three of them are in the ocean, and one is far apart from the other sample points.

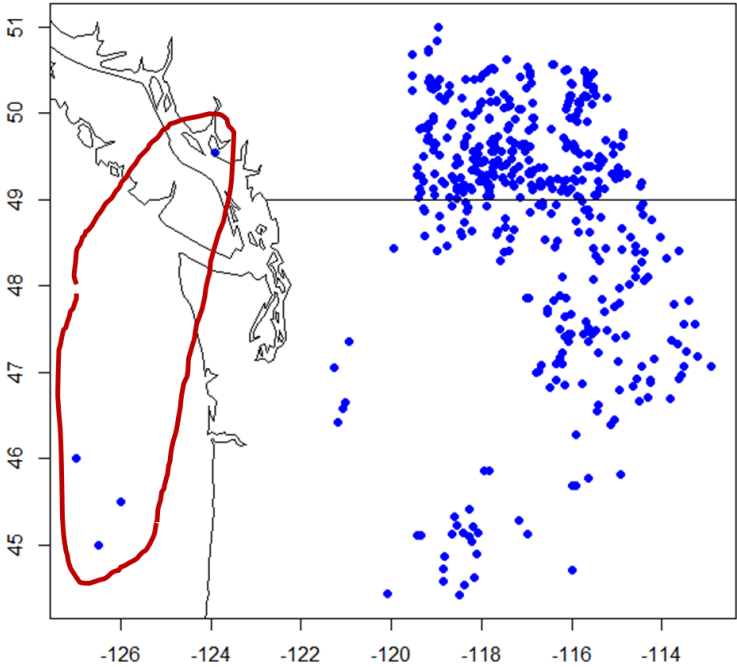


Figure 6. Sample locations of the presence data points of species A on the map with four points being identified as possible errors. Image by Tongli Wang under CC BY 4.0.

Thus, the next step is to remove these data points from the dataset. Here is the code to do so and to plot the remaining sample locations on the map for confirmation:

```
#remove the 4 data points from the dataset —
pr2 <- subset(pr, x > -123); #keep samples with x >
-123
```

```
head(pr2);dim(pr2)
points(y~x,data=pr2,col='darkgreen',pch=19)
```

As shown in Figure 7, the four data points are successfully removed from the dataset.

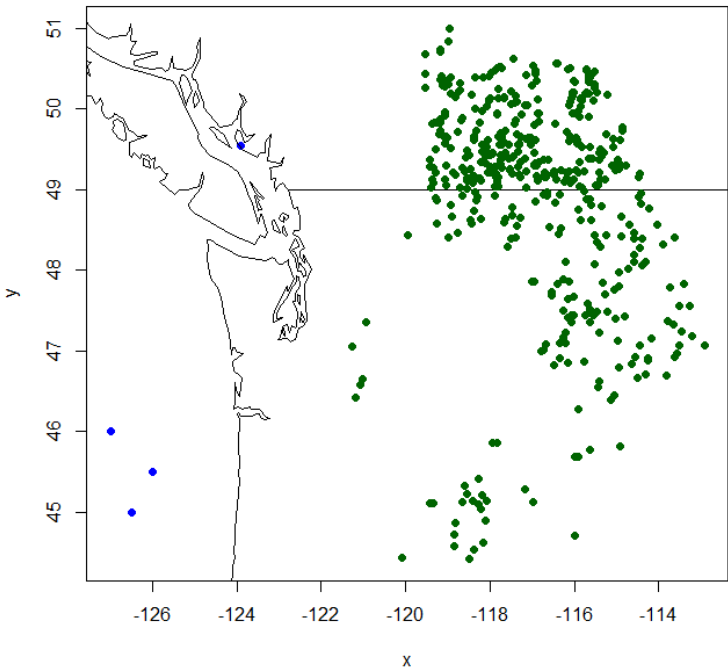


Figure 7. Sample locations of the presence data points of species A on the map before (blue+green) and after (green only) the four data points with errors are removed. Image by Tongli Wang under CC BY 4.0.

3.3 Absence data adjustments

Absence data points can be as useful as presence data points in affecting a model's accuracy and predictions. However, misuse of absence data points can also mislead a model and cause a false model accuracy.

In most modeling approaches, including Random Forest (RF), every data point (regardless of presence or absence) has the same contribution to the model. If there are too many absence data points in a dataset, which is often the case, the model will be more weighed by the absence data points, and the overall model accuracy will also be biased towards the absence data points. As the prediction accuracy is, in most cases, higher for absence data points than for the presence data points, which often results in an overestimate of the model accuracy. We will further demonstrate these issues in our modeling practice later.

One way to adjust the absence data points is to restrain the scope of the absence data points. Limiting the absence data points with 200 km from the edge of the species distribution has been widely applied. The species distribution is represented by the presence data points in practice. Here is the steps:

```
##limiting the scope of absence data points =====  
#display all data points —  
plot(y~x,data=ab,col='grey',pch=1,cex=0.75) #show  
the absence data points  
map('world',add=TRUE) # add the map
```

```

points(y~x,data=pr2,col='darkgreen',pch=19,cex=0.75
) #add presence data points to the map

```

The above code generates the map showing all presence and absence data points as shown in Figure 7A. We can see that range of the absence data points is far beyond the presence data points. The following code is to restrain the absence data points not beyond 2 degree in longitude and latitude, which is about 200 km.

```

#calculate the scope boundaries —
x_max = max(pr2$x)+2;x_max
x_min = min(pr2$x)-2;x_min
y_max = max(pr2$y)+2;y_max
y_min = min(pr2$y)-2;y_min

#drop absence data points outside the scope —
ab2 <- subset(ab,x < x_max & x > x_min & y <
y_max & y > y_min);dim(ab2) #only keep the absence data
points within the defined range.

points(y~x,data=ab2,col='brown',pch=1,cex=0.75)
#show the absence data points within the range

```

These lines of code result in an absence dataset keeping only the absence data points in red color as shown in Figure 7B.

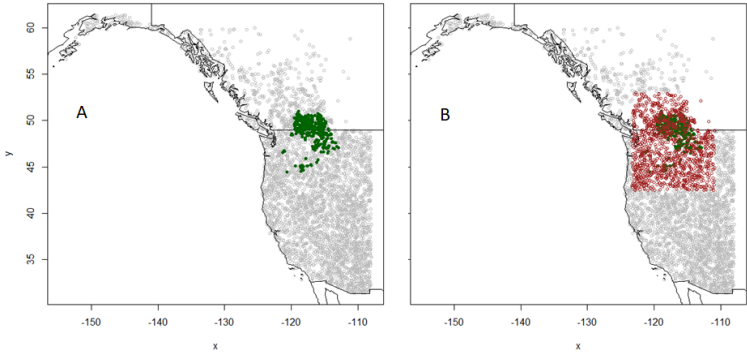


Figure 8. Distributions of presence (green) and absence data points before (gray) and after (red) restrictions. Image by Tongli Wang under CC BY 4.0.

There are 1266 data points remained for absence in the dataset. This number is much bigger than the number of the data points for presence. We should make the numbers of the two categories about the same. As the absence data points represent a larger area, let's make it's number slightly bigger, say by 10%. It can be down using the following code:

```
#drop more absence data points through random
sampling —
n_ab = nrow(pr2)*1.10;n_ab #get the targeted number
of absence data points.
ab3 <-
ab2[sample(1:nrow(ab2),n_ab,replace=FALSE),]
#randomly sample the absence data points
head(ab3)
```

Results:

		x	y		Elevation	SpeciesA
8	196	-12	51.6	7	102	0
		0.241	2250			
4	321	-12	44.5		762	0
		1.354	6834			
5	425	-119	46.8		213	0
		.32	1712			
1	358	-111	42.4	8	240	0
		.114	9286			
3	304	-115	45.4	1	219	0
		.157	5664			
8	182	-117	47.8		823	0
		.722	9664			

We have 550 absence data points left, while there are 500 presence data points. We can now construct the final dataset for model building using the following code.

```
#construct the final dataset for model building —
dat2 <- rbind(pr2,ab3);head(dat2);dim(dat2) #combine
the presence and sampled absence data points
dat3 <- transform(dat2,
```

```

pa=ifelse(speciesA==1,'y','n') #convert "1" to "y" and
"0" to "n"

head(dat3);tail(dat3)

dat4 <- dat3[,-4];head(dat4);dim(dat4) #drop
"speciesA" column

```

The first line is to combine the presence and absence datasets, and the second line is to add a new column for presence “y” and absence “n”. The last line is to drop the speciesA column.

[Results]

	x	y	Elevation	pa
1	-118.7137	50.61465	588	y
2	-114.9559	47.11819	1771	y
3	-119.0610	49.91130	1090	y
4	-116.9345	48.80038	874	y
5	-114.7050	48.45857	1554	y
6	-115.0479	49.25934	1025	y
[1]	1050	4		

4. Climate variables generation

4.1 Generating climate variables for specific sample locations

Climate variables for the dataset for model building should represent the climatic conditions of the sample locations as accurately as possible. Climate data in a grid format, which is provided from most climate data sources, should be avoided if possible. If a model is built with poor climate data, its predictions with high-quality climate data will not help to overcome the problem. Therefore, it is important to use scale-free climate data from ClimateNA or ClimateAP for model building.

To generate climate variables for our dataset, we first need to construct an input file for the climate models (ClimateNA or ClimateAP). Here is the code to do so:

```
#construct an input file for ClimateNA —
ID <- row(dat4)[,1]; #generate an ID for each row
pa_coord <-
data.frame(ID1=ID,dat4[,c(4,2,1,3)]);head(pa_coord)
write.csv(pa_coord,'pa_coord.csv',row.names=FALSE
)
```

Here are the first 5 rows of the input dataset look like:

	ID1	p	y	x	Ele vation
1	1	y	50.6 1465	-118 .7137	588
2	2	y	47.1 1819	-114 .9559	177 1
3	3	y	49.9 113	-119 .0610	109 0
4	4	y	48.8 0038	-116 .9345	874
5	5	y	48.4 5857	-114 .7050	155 4
6	6	y	49.2 5934	-115 .0479	102 5

We can then use ClimateNA to generate climate variables for all these locations. For how to use ClimateNA, please refer to Lecture 3.2. Generating point and spatial climate data using ClimateNA and ClimateAP. We will use the default filename of ClimateNA output “**pa_coord_Normal_1961_1990Y.csv**”. This file will be used for model building later.

4.2 Generating spatial climate variables

We will need spatial climate data in raster (or grid) format for building MaxEnt models and for spatial predictions. Although the scale-free point climate data are more accurate, it is not feasible for model prediction spatially.

Also, MaxEnt requires climate data in a raster stack for model building. Spatial climate data for the entire North America can be downloaded at <http://ClimateNA.ca>. For our practice, we will generate them by ourselves. Here are the steps:

- 1) **Define the extent of the area and obtain a digital elevation model (DEM) covering the targeted area.** We will use GTOPO30 at 800 m resolution DEM. The DEM for western North America can be downloaded from [here](#) on Canvas and put it into our working directory “NicheModels”. We can then import it into R and obtain a DEM in a raster for our targeted area:

```
#define the extent and obtain a DEM —
library(raster)
ext <- extent(dat4); ext; #we use extent of the absence
data points
wna <- raster('wna800.tif');# DEM
dem <- crop(wna,ext);#crop the wna by the extent
plot(dem, xlab='Longitude',ylab='Latitude');#plot
DEM on map
```

The DEM is shown in the figure file is a raster as follows:

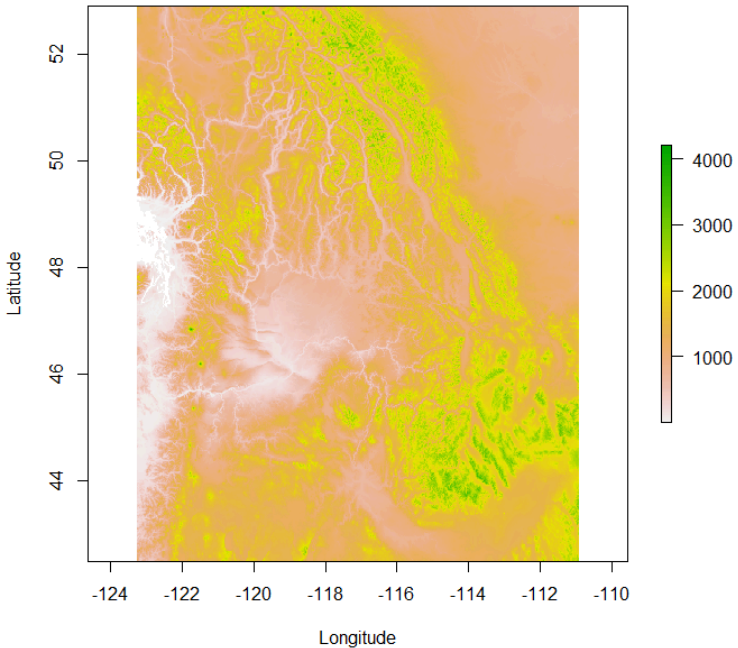


Figure 9. The map of the DEM covering the area of our interest. Image by Tongli Wang under CC BY 4.0.

- **2) Covert and save the DEM to an ASCII file for ClimateNA.** Here is the R code:

```
#Export the DEM to an ASCII file for ClimateNA
NAvalue(dem) <- -9999; # replace NAs with -9999
writeRaster(dem, 'dem800.asc', overwrite=T)
```

- **3) Generate climate variables using**

ClimateNA. With the DEM ASCII file, you can generate climate variables using ClimateNA referring to Lecture 3.2. For this practice, we need annual climate variables for the reference period 1961-1990 and a future period “13GCMs_ensemble_ssp245_2041-2070Y”. The ASCII output files need to be saved to the following sub-directories:

```
“ClimateData\Normal_1961_1990Y”
  “ClimateData\
  13GCMs_ensemble_ssp245_2041-2070Y”
```

- 4) **Generate raster stacks.** We need to compile the climate variables into stacks for predictions and for MaxEnt model building. It can be done with the following code:

```
#stacking rasters —
clmDir <- ‘ClimateData/Normal_1961_1990Y/’
#specify a subdirectory
rlist <- list.files(clmDir,pattern='.asc',full.name=T);
rlist[1:5]; #create a file list of climate variables
for(lyr in rlist){
  r <- raster(lyr)
  if(grepl(‘MAT|MCMT|MWMT|TD|AHM|SHM|EMT|
EXT|MAR’,lyr)){r=r/10} #convert integer to decimal for
temperature variables
  if(lyr==rlist[1]){stk6190=r}
else{stk6190=stack(stk6190,r)} # “6190” representing
1961_1990
```

```

    }; stk6190

    writeRaster(stk6190,'ClimateData/
Normal_1961_1990Y/stk6190.grd')

    clmDir <- 'ClimateData/
13GCMs_ensemble_ssp245_2041-2070Y/'

    rlist <- list.files(clmDir,pattern='.asc',full.name=T);
rlist[1:5]

    for(lyr in rlist){
    r <- raster(lyr)

    if(grepl('MAT|MCMT|MWMT|TD|AHM|SHM|EMT|
EXT|MAR',lyr)){r=r/10}

    if(lyr==rlist[1]){stk2055=r}
else{stk2055=stack(stk2055,r)} # "2055" representing
2041_2070

    }; stk2055

    writeRaster(stk2055,'ClimateData/
13GCMs_ensemble_ssp245_2041-2070Y/stk2055.grd')

```

The above code save a raster stack to each of the above sub-directories.

Now, we can save and close the script file “1. dataSteps.R”.

5. To obtain species occurrence data

Species occurrence data can be obtained from various sources. GBIF is one of the major data sources. GBIF (the Global Biodiversity Information Facility) is an

international network and data infrastructure funded by the world's governments and aimed at providing anyone, anywhere, open access to data about all types of life on Earth. The following will show you how to obtain the occurrence data for species and how to prepare the data for building niche models. This section can be useful for your group project.

5.1 Data download

The occurrence data of a species can be downloaded at: <https://www.gbif.org/occurrence/search>. It is better to use the scientific name to search the database. For example, if we want to get the data for Garry oak, we need to use the name *Quercus garryana* Douglas ex Hook. After hitting the search button, it shows 7,495 results (this number changes with time).

The next step is to click the **Download** button and choose the **Simple** option to download the data. After the downloading is complete, rename the file to “garryOak.csv”.

5.2 Data processing in R

The dataset downloaded has 50 columns and a lot of missing information. Such a dataset is required to be reorganized and cleaned. These steps can be done in R with the following code:

```
#Process species occurrence data from GBIF  
library(maps)
```

```

library(mapdata)
library(data.table)
#load the data —
csv <- fread('garryOak.csv'); head(csv); dim(csv)
#clean the data —
csv2 <-
csv[,c('decimalLatitude', 'decimalLongitude', 'elevation')];
head(csv2) #only keep lat, long and elevation
csv3 <-
csv2[!is.na(csv2$decimalLatitude),]; head(csv3) # drop the
rows missing lat
csv4 <- csv3[!is.na(csv3$elevation),]; head(csv4);
dim(csv4) # drop the rows missing elevation
names(csv4) <- c('lat', 'lon', 'elev'); head(csv4)
#rename the columns
csv4$species <- 'y'; head(csv4)
#inspect the data —
plot(lat~lon, csv4)
map('world', add=TRUE)
#remove outliers —
csv5 <- subset(csv4, lon < -115); head(csv5);
dim(csv5)
plot(lat~lon, csv5)
map('world', add=TRUE)

```

5.3 Add absence data points

The dataset obtained has presence data points only. Thus, we need to add absence data points, which can be done using the DEM file through the following steps in R.

```

## add absence data points —————
library(raster)

#load DEM—
wna <- raster('wna4000.asc');wna
xyv <- rasterToPoints(wna);head(xyv) #convert raster to
points
xyv2 <-
xyv[sample(1:nrow(xyv),20000,replace=FALSE),];head(xy
v2);dim(xyv2) #sampling
dat <- data.frame(xyv2);head(dat) #make it as a dataframe
dat$species <- 'n';head(dat) #add a new variable
names(dat) <- c('lon','lat','elev','species');head(dat)
#rename the columns to be consistent to presence data

# combine the presence and absence data, and remove
overlapping absence data points—
cmb <- rbind(csv5, dat); head(cmb);dim(cmb)
cmb$x2 <- round(cmb$lon,1); #make a new variable x2
from rounded lon
cmb$y2 <- round(cmb$lat,1);head(cmb) #make a new
variable x2 from rounded lat

#drop the data points falling into the same grid at 0.1
degree (10 km)
cmb2 <-
cmb[!(duplicated(cmb[,c('x2','y2')])|duplicated(cmb[,c('x2
','y2')])), ];head(cmb2)
xy_n <- subset(cmb2,species=='n');head(xy_n);dim(xy_n)
#re-generate the absence data
cmb3 <- rbind(csv5,xy_n[,1:4]);head(cmb3); #combine the
presence data with the new absence data

```

Now you have a dataset comprising both presence and absence data points. To build a niche model with this

dataset, you will need to follow the steps in Section 3 above.

Topic 6.2 Building niche-based models - 2/3

Tongli Wang

Contents

[5 Random Forest Model](#)

[5.1 Overview](#)

[5.2 Model building](#)

[5.3 Importance values](#)

[5.4 Predictions for specific locations](#)

[5.5 Predictions for geographic areas](#)

[5.6 Overlaying sample data points](#)

[6 MaxEnt model](#)

[6.1 Overview](#)

[6.2 Installations of required packages](#)

[6.3 Model building](#)

[6.4 Model evaluation and cutoff criterion](#)

[6.5 Predictions for specific locations](#)

[6.6 Predictions for geographic areas](#)

5. Random Forest Model

5.1 Overview

Random Forest (RF) is an ensemble learning method for classification, regression and other tasks. A RF model is

constructed with many decision trees, collectively called a 'forest'. The output of the model is the mode of the classes (classification) or mean prediction (regression) of the individual trees.

Each of these decision trees in the forest is constructed using a bootstrap sample of the input data (i.e., a random sample with replacement) so that the resulting dataset ('bagged sample') contains about 64% of the original observations, and the remaining observations comprise the 'out-of-bag' (OOB) sample.

Using the trees grown from a bootstrap sample, each of the independent observations in the OOB sample is classified (assigned to either presence or absence) and a model prediction error, called the OOB error (% of incorrectly classed observations), is calculated.

RF is designed for overcoming collinearity and overfitting problems and is considered one of the most credible statistical methods for climate niche model building. Thus, RF can handle a large number of independent variables, and a prescreen of these variables is not required.

A Random Forest model has many decision trees. Each tree can be considered as a trained expert from a random university. With many trained experts from many universities together is likely to make the right decision from their consensus, as illustrated in Figure 1.

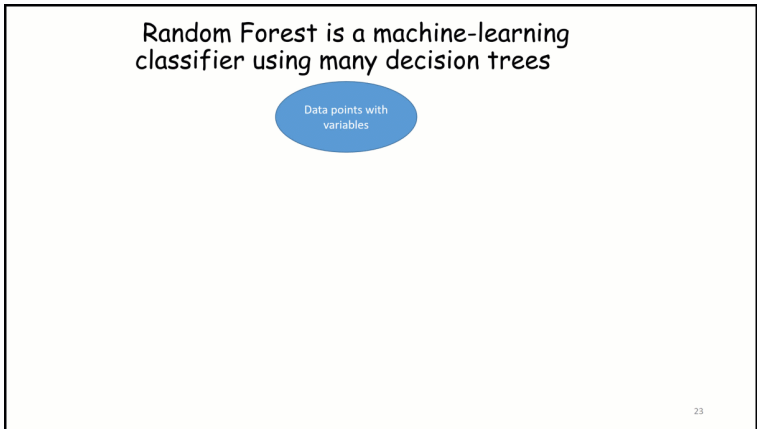


Figure 1. Illustration of how Random Forest work to make a right decision. Image by Tongli Wang under CC BY 4.0.

5.2 Model building

The model building or model fitting is technically quite similar across the modeling methods that exist in R. a formula is required to identify the dependent and independent variables. A simple formula could look like this:

$$y \sim x1 + x2$$

where, y is a function of $x1$ and $x2$. For RF model building, it can be written as:

```
formula01 <- pa ~ MAT + MWMT + MCMT
randomForest(formula01, data = dat4)
```

The formula can also be written as:

```
randomForest(x,y)
```

where, x is the matrix for all independent variables, and y is the vector for the dependent variable. This format is simpler and easier to use.

We will build a Random Forest model with the dataset

that we have prepared. We need to start a new script file called “rf_model.R”, and type the following code:

R code:

```
## building an RF model
library(randomForest)
#import the dataset generated by ClimateNA —
dat4 <- read.csv('ClimateData/
pa_coord_Normal_1961_1990Y.csv');head(dat4)
dat4[dat4 < -9999] <- NA # replace all -9999
with NA in all cols.
dat5 <- na.omit(dat4) #drop all rows with NA
```

Here are the first five rows of the dataset look like:

	D1	I a	P	Latitude	Longitude	El
1		1	y	50.61465	-118.714	58
2		2	y	47.11819	-114.956	17
3		3	y	49.9113	-119.061	10
4		4	y	48.80038	-116.935	87
5		5	y	48.45857	-114.705	15
6		6	y	49.25934	-115.048	10

Next, we need to construct x matrix comprising all the climate variables, and y vector with the dependent variable.

```
# construct x matrix and y vector —
names(dat5); #showing the variable names with
column numbers
x <- dat5[,6:29]; # all climate variables
y <- as.factor(dat5$pa); #y vector to be a factor
variable
```

Now, we are ready to do the model fitting:

```
#RF model fitting —  
rf <- randomForest(x,y)  
print(rf)
```

Here is the result of the model fitting:

Output

```
Call:  
  
  randomForest(x = x, y = y)  
Type of random forest: classification  
Number of trees: 500  
No. of variables tried at each split: 4  
OOB estimate of error rate: 14.38%  
  
Confusion matrix:  
n y class.error  
n 455 95 0.1727273  
y 56 444 0.1120000  
  
>
```

Please note, the default number of trees (ntree) is 500. It can be changed, but it should not be too small. However, it may take a long time to build the model if the ntree is too large.

The number of variables tried at each split is 4, which varies depending on the total number of variables in the dataset. In most cases, there is no need to change this value.

OOB estimate of error rate is 14.38%. In other words, the accuracy of the model prediction is 85.62%. This is an independent model evaluation, as it is based on OOB samples not involved in model building.

Let's re-build the model with 200 trees and make the importance available.

```
rf2 <- randomForest(x,y, ntree=200)
print(rf2)
```

It shows a similar result:

```
Call:
randomForest(x = x, y = y, ntree = 200)

Type of random forest: classification
Number of trees: 200
No. of variables tried at each split: 4
OOB estimate of error rate: 14.48%

Confusion matrix:
n y class.error
```

```
n 460 90 0.1636364
y 62 438 0.1240000importance(rf2)
```

It indicates that 500 trees may not be necessary, but it does not hurt if it does not take too long to run.

5.3 Importance values

RF also generates an importance value for each independent variable. It can be used to rank the variables and to eliminate the least important variables from the model. It is often used to identify important variables driving the differences among the observations.

Here is the code to make the importance values available:

R code

```
rf2 <- randomForest(x,y, ntree=200, importance=T)
imp <- importance(rf2) #get importance values
imp[order(-imp[,3]),] # sort by the importance values
```

It displays the importance values for all independent variables:

Output

Variable	Variance	Mean	Standard Deviation	MeanDecreaseAccuracy	MeanDecreaseGini
AR	8.874124	24.50032	23.99387	48.51724	
AS	5.305395	17.43046	18.23334	35.03652	
D	2.922773	17.44136	17.8561	29.08314	
CMT	3.835332	16.25158	16.69499	31.04835	
WMT	3.197735	14.58981	15.84366	20.67795	
MT	4.18638	13.06808	15.59201	17.19945	
AP	4.296944	13.66554	15.57938	22.54782	
HM	-0.73502	14.89259	15.50967	21.25916	
SP	-2.41199	15.23768	15.01118	17.77178	
D_0	5.074101	12.19935	14.60607	32.12419	
MI	3.189601	13.02122	14.48578	21.63149	
ref	5.506015	10.91403	13.97705	16.25914	
FP	6.18945	10.57743	13.8959	20.85242	

D_18	D	5.3	11.	12.81475	23.28081
		45905	29515		
MD	C	1.4	12.	12.57828	19.16087
		46897	32909		
HM	A	3.2	11.	12.3261	20.86442
		66295	21948		
D18	D	3.2	10.	12.21824	15.34795
		06686	25491		
XT	E	3.7	10.	12.20974	13.68602
		32476	48977		
D5	D	5.6	10.	12.13998	20.64162
		06495	12526		
FFD	N	6.1	9.9	12.02255	18.0316
		65801	92309		
FP	F	5.9	9.4	11.94414	15.85869
		65675	87863		
AT	M	5.3	11.	11.8743	18.40527
		76307	21463		
H	R	4.2	10.	11.75021	12.41577
		9294	11635		
FP	eF	2.8	8.9	9.473222	11.51494
		53226	41361		

MeanDecreaseAccuracy is the amount of decrease in model accuracy if the given independent variable is removed from the model. This is an often-used parameter. MeanDecreaseGini represents the number of splits involved by that climate variable. It also indicates the importance of the variable.

So far, we have completed building a RF model, “rf” or “rf2”. Let’s use “rf2” as the final model for predictions later. Let’s save this model for later use:

```
#save the model —  
save(rf2, file='rf2.Rdata')
```

5.4 Predictions for specific locations

We first need to generate climate variables for the locations to be predicted. In this practice, we will use sample locations from the dataset that we used to build the models. Let’s first open a new script file called “rf_predict.R” and type in the following R code:

```
## predictions  
library(randomForest)  
  
#import the sample coordinate dataset —  
dat0 <-  
read.csv('pa_coord_Normal_1961_1990Y.csv');head(dat0)  
dat2 <-  
dat0[sample(1:nrow(dat0),10,replace=F),];#take 10 random  
samples
```

```
#random forest =====  
load('rf2.Rdata'); #load the RF model  
p <- predict(rf2,dat2);p #make predicitions  
  
#combine predictions with locations —  
p2 <- cbind(dat2[,1:5],p);p2
```

Here is the final result – the last column shows the result of predicitions:

Output

		ID1	pa	Lati tude	Lon gitude	Ele vation	P
	356	356	y	49.0 8627	-116 .11	910	y
	14	14	y	50.3 1583	-119 .149	750	y
5	100	100	n	45.0 6805	-123 .108	116	n
9	100	100	n	42.5 3559	-112 .344	160	n
	629	629	n	44.9 7478	-119 .847	143	n
	254	254	y	49.7 8399	-118 .297	136	y
	690	690	n	42.9 695	-118 .965	149	n
	162	162	y	50.0 8556	-116 .841	142	y
	267	267	y	50.1 8432	-116 .892	135	y
	604	604	n	49.0 2889	-114 .267	233	n

In this example, we assume that the 10 locations are not involved in the model building. Otherwise, the predictions can be biased. If we want to get unbiased predictions for the locations involved in the training dataset, we can also obtain independent predictions that are produced using

the OBB samples, which are stored within the RF model. Here is the code to retrieve the unbiased predictions:

```
p0 <- predict(rf2)
```

5.5 Predictions for geographic areas

Steps for spatial predictions require spatial climate data in raster stacks, which have been prepared in 4.2. We will continue using the script file called “rf_predict.R” and type the following code:

```
...  
#spatial predictions —  
library(raster)  
  
# load climate raster stacks —  
stk6190 <- stack('ClimateData/Normal_1961_1990Y/  
stk6190.grd')  
stk2055 <- stack('ClimateData/13GCM-  
Ensemble_ssp245_2055Y/stk2055.grd')  
  
# rf predictions —  
ps_rf <- predict(stk6190,rf2)
```

```
ps_rf = ps_rf - 1  
plot(ps_rf)
```

Please note that the order of the rf model (rf2) and the climate data (stk6190) is different in this predict function from the one for specific locations. Also, the direct output is “2” for presence and “1” for absence. The above code converted it to “1” and “0”. The result of RF spatial predictions is shown in Figure 2.

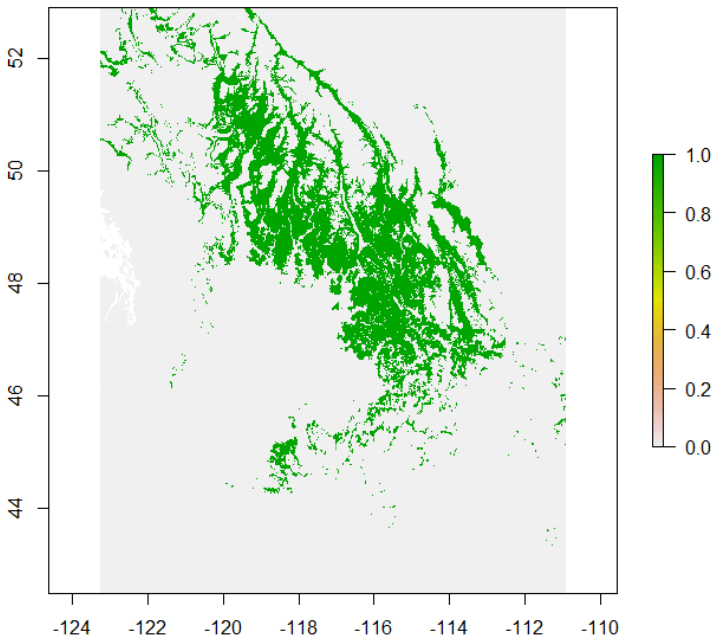


Figure 2. Spatial predictions from a RF model showing the occurrence of the species. Image by Tongli Wang under CC BY 4.0.

RF model output can also be in probability, which can be useful for adjusting the threshold for ‘presence’ vs.

‘absence’ or for the prediction of species abundance. This can be done by specifying: type= “prob” and index=2, using the following code:

```
ps_rf <- predict(stk6190,rf2, type='prob',index=2)
plot(ps_rf)
```

We can also predict the probability for the future scenario “Ensemble_ssp245_2055Y” as follows.

```
ps_rf_2055 <- predict(stk2055,rf2, type='prob',
index=2)
plot(ps_rf_2055)
```

Predicted maps are shown in Figure 3.

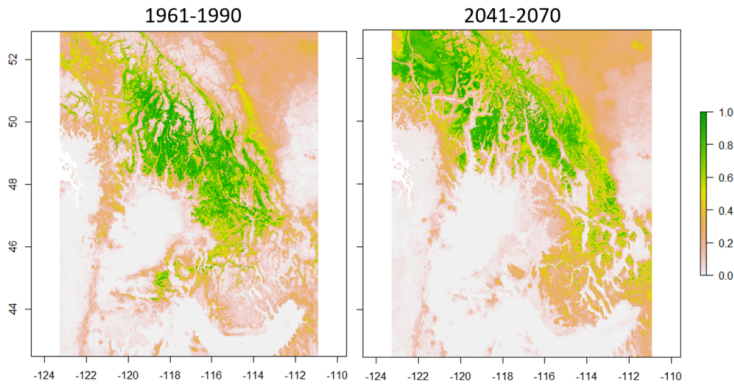


Figure 3. Spatial predictions from a RF model showing the occurrence of the species in probability for the periods of 1961-1990 and 2041-2070. Image by Tongli Wang under CC BY 4.0.

5.6 Overlaying sample data points

Overlaying presence data points over spatial predictions for the reference period is an effective way to visualize the model performance. In addition to see the over performance, we can also see the spatial pattern of the model performance. Here is the code:

```
#visualization of model performance
pr_coord <-
subset(dat0,pa=='y')[,c(4,3)];head(pr_coord) # presence
points
plot(ps_rf) # rf model predictions
points(Latitude~Longitude,pr_coord)
```

Results are shown in Figure 4. We can see that the presence data points are somewhat lagging behind niche predictions in the north, the leading edge. The opposite is shown at the trailing edge.

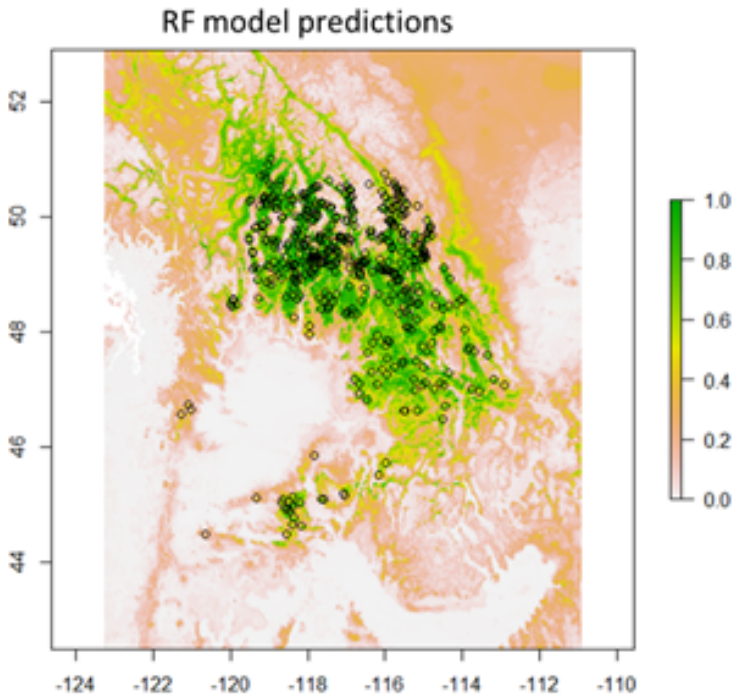


Figure 4. Presence data points overlaying over spatial predictions from a RF model. The color gradient shows the occurrence of the species in probability and circles indicate the locations of presence data points. Image by Tongli Wang under CC BY 4.0.

Now, we can save and close the script file “rf_predict.R”.

6. MaxEnt model

6.1 Overview

Maxent is a general-purpose machine learning method with a simple and precise mathematical formulation, and it has a number of aspects that make it well-suited for species distribution modeling. Maxent software is based on the maximum-entropy approach for modeling species niches and distributions. From a set of environmental (e.g., climatic) grids and georeferenced occurrence localities (e.g. mediated by GBIF), the model expresses a probability distribution where each grid cell has a predicted suitability of conditions for the species. Under particular assumptions about the input data and biological sampling efforts that led to occurrence records, the output can be interpreted as predicted probability of presence, or as predicted local abundance (raw exponential output).

6.2 Installations of required packages

MaxEnt (Maximum Entropy) is the most widely used niche-based modeling algorithm. MaxEnt is available as a standalone Java program. An R package “dismo” has a function “maxent” that communicates with the Java program. Thus, to use MaxEnt, we need to install “dismo”, “maxent” and “java”.

Here are the steps to create an environment to use this function:

- Install “dismo” package in R.
- Download the MaxEnt java program from: <https://biodiversityinformatics.amnh.org/>

[open source/maxent/](#)

Copy the file ‘maxent.jar’ from the downloaded package and put it in the ‘java’ folder of the ‘dismo’ package in the R library, like: “C:\Users\tlwang\Documents\R\win-library\3.5\dismo\java\”

- Download Java software from:
<https://www.java.com/en/download/manual.jsp>.
Please pay attention to the system versions of your computer: 32-bit vs. 64-bit.
- Install Java following the on-screen instructions.

6.3 Model building

We are now ready to use the maxent function in “dismo” package. Let’s open a new script file called “maxent.R” and type the following code:

```
## building a MaxEnt model
library(dismo)
library(raster)

#import the sample coordinate dataset —
dat4 <- read.csv(‘pa_coord.csv’);head(dat4);
dat4[dat4 < -9999] <- NA # replace all -9999 with NA
in all cols.
dat5 <- na.omit(dat4) #drop all rows with NA
```

```
# prepare a coordinate file for presence data only —
pr_coord <-
subset(dat5,pa=='y')[,c(4,3)];head(pr_coord);dim(pr_coord
)

#import climate raster stack —
Stk6190 <- stack('ClimateData/Normal_1961_1990Y/
stk6190.grd')

# maxent model building —
mx <- maxent(stk6190,pr_coord)
save(mx, file='mx.Rdata') # save the maxent model
print(mx)
```

The print function produces receiver operating characteristic (ROC) curve (Figure 9) and variable contributions (Table 1). They show the model performance and contributions of climate variables to the model. The higher the area under the curve (AUC), the better the model performance. The AUC results were considered excellent for AUC values between 0.9-1, good for AUC values between 0.8-0.9, fair for AUC values between 0.7-0.8, poor for AUC values between 0.6-0.7, and failed for AUC values between 0.5-0.6.



Figure 5. Receiver operating characteristic (ROC) curve for the training data. Specificity is the true-positive rate and 1-specificity is the false-positive rate. Image by Tongli Wang under CC BY 4.0.

Table 1. Contributions of climate variables to the model. Both Percent contribution and Permutation importance are in %. They are often not in agreement, and the former is more often used.

Variable	Percent contribution	Permutation importance
PAS	34.4	19.8
MAR	13	5.7
NFFD	8.4	0.1
MCMT	7.2	1.4
EMT	6.9	2.1
TD	5.5	1.6
DD5	4.2	0.1
RH	3.6	17.5
AHM	3.5	6.5
bFFP	2.6	0.6
DD_18	2.6	2.1
MAP	1.8	2.6
SHM	1.2	1
MSP	1	5.6
MWMT	0.9	0.9
DD_0	0.8	0.7
EXT	0.7	17.6
FFP	0.6	0.7
CMD	0.6	3
DD18	0.2	8.3
Eref	0.2	0.6
eFFP	0.1	0.9
MAT	0	0.5

6.4 Model evaluation and cutoff criterion

As the model output is the probability of occurrence, we

need to define a cutoff criteria to separate presence and absence using the following code:

```
#evaluate to obtain cutoff criterion —  
bg <- randomPoints(stk6190,500);# get 500 random  
absence points  
evaluate(mx,p=pr_coord,a=bg,x=stk6190)
```

It shows:

```
class : ModelEvaluation  
n presences : 494  
n absences : 500  
AUC : 0.9338057  
cor : 0.7918147  
max TPR+TNR at : 0.3299249
```

We can see that the optimal cutoff criterion is 0.33, and the correlation coefficient is 0.79, which is lower than the amount of variance explained in RF model (85%).

6.5 Predictions for specific locations

For MaxEnt model, the code is similar to that for RF

models. Let's continue with R script "maxent.R" and type the following lines:

```
#import the sample coordinate dataset —
dat0 <-
read.csv('pa_coord_Normal_1961_1990Y.csv');head(dat0)
dat2 <-
dat0[sample(1:nrow(dat0),10,replace=F),,];#take 10 random
samples

# load MaxEnt model —
load('mx.Rdata') #load MaxEnt model

# make predictions —
p_mx <- predict(mx,dat2)
p_mx2 <- cbind(dat2[,1:5],p_mx);p_mx2 #combine to
RF predicitions
```

Here is the results:

		ID1	pa	Lati tude	Lon gitude	Ele vation	x	p_n
	356	356	y	49.0 8627	-116 .11	910	3076	0.12
	14	14	y	50.3 1583	-119 .149	750	5297	0.12
5	100	100	n	45.0 6805	-123 .108	116	8223	0.04
9	100	100	n	42.5 3559	-112 .344	160	7115	0.00
	629	629	n	44.9 7478	-119 .847	143	5729	0.00
	254	254	y	49.7 8399	-118 .297	136	7268	0.07
	690	690	n	42.9 695	-118 .965	149	2601	0.02
	162	162	y	50.0 8556	-116 .841	142	4913	0.04
	267	267	y	50.1 8432	-116 .892	135	738	0.00
	604	604	n	49.0 2889	-114 .267	233	0759	0.03

MaxEnt model predictions give the probability of occurrence. Users can define a criterion to convert it to presence/absence classes.

6.6 Predictions for geographic areas

Steps for spatial predictions are similar but require spatial climate data in raster stacks, which have been prepared in 4.2. We will continue using the above script file and type the following code:

Code for spatial predictions

```
...
#spatial predictions —
library(raster)
# load climate raster stacks —
stk6190 <- stack('ClimateData/Normal_1961_1990Y/
stk6190.grd')
stk2055 <- stack('ClimateData/15GCM-
Ensemble_rcp45_2055Y/stk2055.grd')

# spatial predicitions —
ps_mx <- predict(stk6190,mx)
plot (ps_mx)

#overlying presence data points —
pr_coord <-
subset(dat0,pa=='y')[,c(4,3)];head(pr_coord)#presence
points
points(Latitude~Longitude,pr_coord)
```

Here is the result shown in Figure 6.

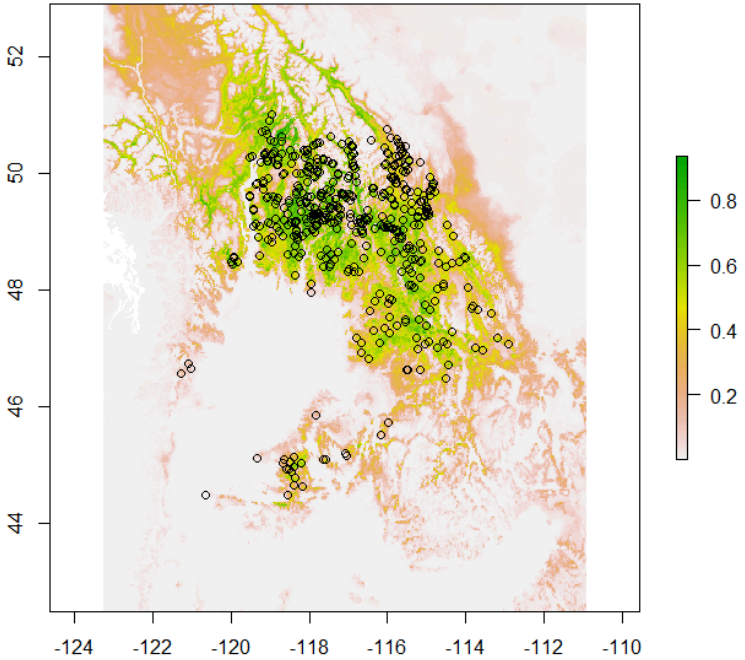


Figure 6. Spatial predictions from a MaxEnt model showing the occurrence of the species in probability and locations of presence data points. Image by Tongli Wang under CC BY 4.0.

The following code is for prediction for the period 2041-2070 (or 2055),

```
# future predictions —  
ps_mx_2055 <- predict(stk2055, mx)  
plot(ps_mx_2055)
```

The predicted map is shown in Figure 7.

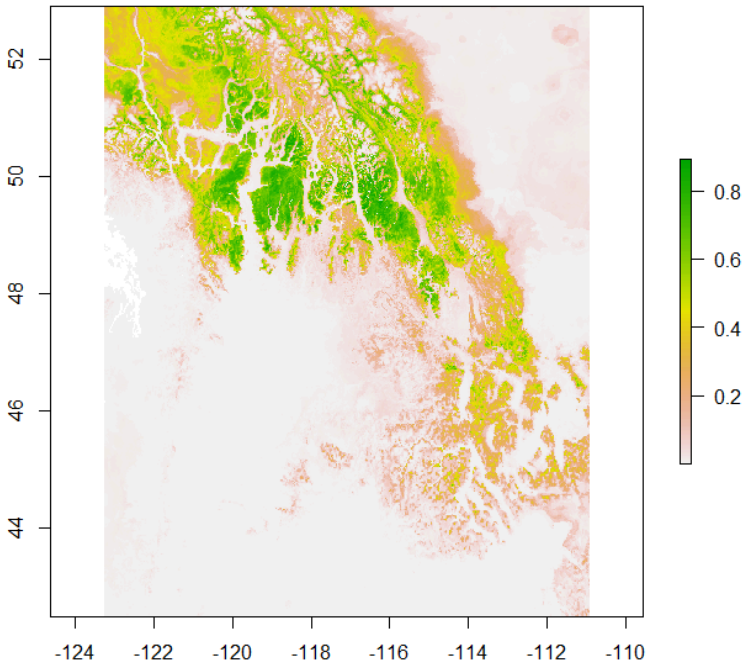


Figure 7. Spatial predictions from a MaxEnt model showing the occurrence of the species in probability for the period 2041-2070. Image by Tongli Wang under CC BY 4.0.

Topic 6.3 Building transfer functions and response functions - 3/3

Tongli Wang

Contents

[1 Overview](#)

[2 Dataset preparation](#)

[2.1 Provenance data](#)

[2.2 Climate data](#)

[2.3 Data integration](#)

[2.4 Data visualization](#)

[3 Transfer functions](#)

[3.1 Individual transfer functions](#)

[3.2 General transfer functions](#)

[3.3 Relative performance-based general transfer functions](#)

[4 Genecology functions](#)

[5 Response functions](#)

[6 Universal response functions](#)

[6.1 Model development](#)

[6.2 Mapping](#)

1. Overview

This lecture will show, step by step, how to build transfer functions, genecology functions, response functions, and universal response functions. These functions can be built

through regression analysis between a trait and climatic variables with linear, quadratic, or non-linear functions. It is quite straightforward in most cases. However, it is more sophisticated to model a fundamental niche with a universal response function. Nevertheless, all of these can be performed in R. Some [basic knowledge](#) has been introduced in 5.1. It is required to go through that part before continuing on this lecture.

2. Dataset preparation

2.1 Provenance data

We use a subset of the provenance-test dataset from British Columbia to demonstrate the methodologies to build these models. AS the dataset is not a complete one, please do not use it for other purposes. Here are the steps:

- Create a folder “TranRespFun”;
- Download the dataset “provDat.csv” [here](#) and put it into the folder “TranRespFun”;
- Create a new file “0.Rdata” in the same folder and double click it to start R;
- Create a script file in R called “dataSteps.R”;
- Type the following code in the script file to import and display the dataset:
-

```
## load the provenance dataset
```

```
=====
```

```
dat <- read.csv('provDat.csv');head(dat);dim(dat)
```

Output:

	site	s	lat_	s	lon_	el_s	v	pro	p	lat_
1	TE	BA	53.8	98	121.	730	61	6667	53.8	
2	AP	CH	54.9	68	126.	790	103	1667	54.9	
3	EE	FR	49.0	85	115.	152	68	1667	51.0	
4	MM	CO	50.9	07	120.	137	42	8333	49.9	
5	SK	BO	52.1	8	120.	110	23	1667	53.0	
6	UI	EQ	50.3	6	119.	137	56	8333	49.9	

```
[1] 2000 9
```

2.2 Climate data

After we have got the provenance coordinates, elevation, and height data growth data, we need to generate climate data for test sites and provenances. Here is the code to generate input files for ClimateNA:

```
## input dataset for sites —  
sites <-  
data.frame(ID=row(dat)[,1],dat[,1:4]);head(sites)  
write.csv(sites, 'sites.csv',row.names=F)
```

Output

	ID	site	lat_	lon_	el_s
1	1	TE BA	53.8	121.	730
2	2	AP CH	54.9	126.	790
3	3	EE FR	49.0	115.	152
4	4	MM CO	50.9	120.	137
5	5	SK BO	52.1	120.	110
6	6	UI EQ	50.3	119.	137

input dataset for provs —

```
provs <-
data.frame(ID=row(dat)[,1],dat[,5:8]);head(provs)
write.csv(provs, 'provs.csv',row.names=F)
```

The input files are saved as “sites.csv” and “provs.csv” in the same folder “TranRespFun”. Now we can run ClimateNA to generate annual climate variables using these two input files, respectively.

2.3 Data integration

The output files of ClimateNA can be imported to R:

```
## load climate data —
clmS <-
read.csv('sites_Normal_1961_1990Y.csv');head(clmS)

clmP <-
read.csv('provs_Normal_1961_1990Y.csv');head(clmP)
```

```
> head(clmP)
```

	ID	v	pro	Lati tude	Lon gitude	Ele vation	T	MA
1	1	61	6667	53.8	-121.8	838		3
2	2	103	1667	54.8	-124.267	970		1
3	3	68	1667	51.0	-115.033	150	1	1.5
4	4	42	8333	49.1	-117.583	998		6
5	5	23	1667	53.0	-123.233	110	0	1.6
6	6	56	8333	49.9	-114.917	128	0	2.9

We then attached the site MAT and MAP (column 6 and 10) to the original dataset “dat” and rename them as ‘mat_s’ and ‘map_s’, respectively:

```
## attached to original dataset —
dat2 <- cbind(dat, clmS[c(6,10)]); head(dat2)
names(dat2)[10:11] <- c('mat_s', 'map_s'); head(dat2)
```

and to do the same for provenance climate data, but renamed as ‘mat_p’ and ‘map_p’. We also need to replace “-9999” (missing values in ClimateNA) to NA (missing values in R), and save the final dataset:

```
dat3 <- cbind(dat2, clmP[c(6,10)]); head(dat3)
names(dat3)[12:13] <- c('mat_p', 'map_p'); head(dat3)

#data cleaning —
dat3[dat3 < -9000] <- NA
dat4 <- na.omit(dat3) #drop the rows with NA
```

2.4 Data visualization

The final dataset has the following columns:

```
> head(dat4)
```

		site	s	lat_	s	lon_		el_s	v	pro	p	lat_
1	TE	BA	8	53.8	98	121.		730		61	6667	53.8
2	AP	CH	2	54.9	68	126.		790		103	1667	54.8
3	EE	FR	8	49.0	85	115.	0	152		68	1667	51.0
4	MM	CO	2	50.9	07	120.	0	137		42	8333	49.9
5	SK	BO	8	52.1	8	120.	0	110		23	1667	53.0
6	UI	EQ	7	50.3	6	119.	0	137		56	8333	49.9

It is a good idea to also visualize the distribution of the data points on a plot and a map.

```
# spatial visualization —
plot(lat_s~lon_s,dat4) #test sites
```

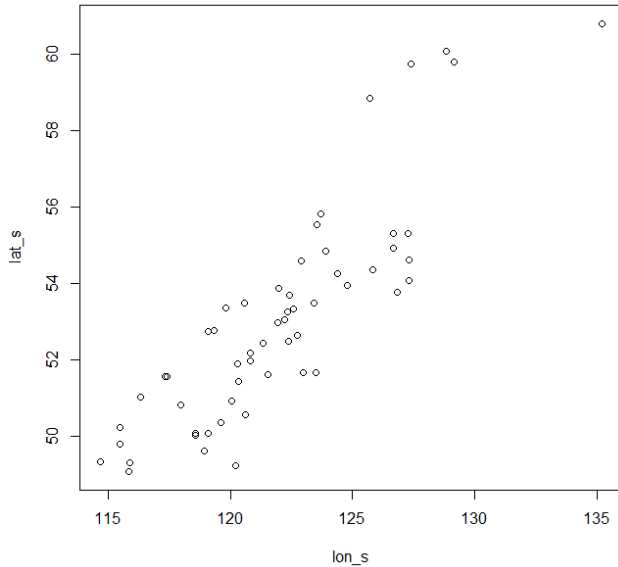


Figure 1. Distribution of test sites along latitude and longitude. Image by Tongli Wang under CC BY 4.0.

From Figure 1, we can see that the longitude is positive values, while it is supposed to be all negative values. So, we can adjust it and plot it again by:

```
#adjust lon_s and plot again —
dat4 <- transform(dat4,lon_s=lon_s * -1)
plot(lat_s~lon_s,dat4) #test sites
```

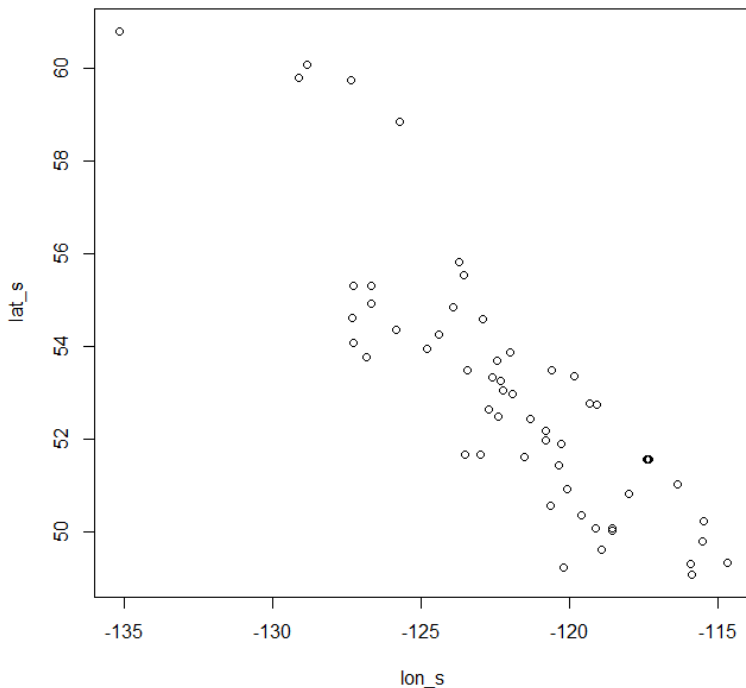


Figure 2. Distribution of test sites along latitude and longitude. Image by Tongli Wang under CC BY 4.0.

Now, let's plot provenance locations

```
plot(lat_s~lon_s,dat4) #test sites
```

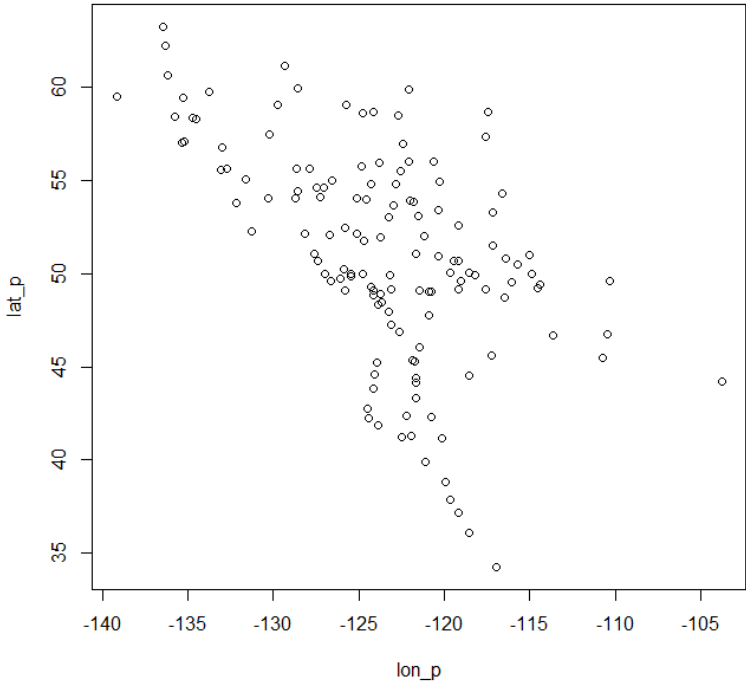
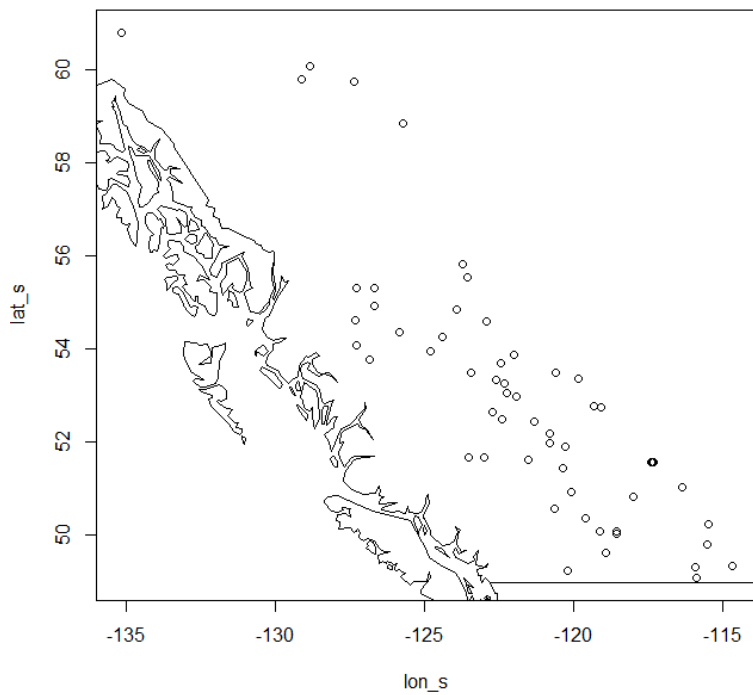


Figure 3. Distribution of provenances along latitude and longitude. Image by Tongli Wang under CC BY 4.0.

We can also plot the locations on a map to make the visualization more effective.

```
# mapping test sites —  
library(maps)  
plot(lat_s~lon_s,dat4) #test sites  
map('world',add=TRUE)  
plot(lat_p~lon_p,dat4) #test sites  
map('world',add=TRUE)
```

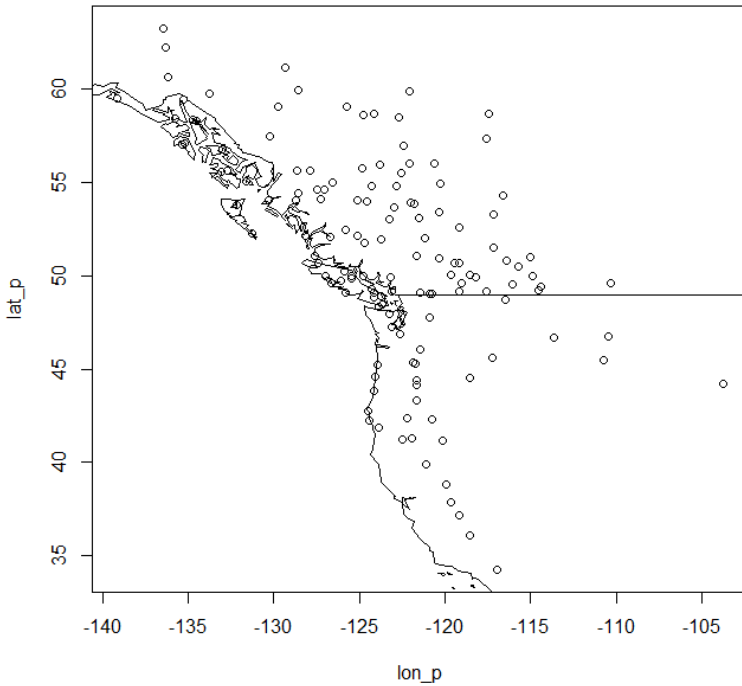


Figure 4. Distribution of test sites and provenances along latitude and longitude. Image by Tongli Wang under CC BY 4.0.

```
#save the final data file —  
write.csv(dat4,'htclm.csv')
```

Now we can save and close the script file.

3. Transfer functions

A transfer function is to reflect the relationship between

the performance of populations and their transfer distances. It can be built for a test site with multiple populations originated from a range of seed sources (i.e., provenances) or for multiple tests combined.

Let's create a new script file in R called "transfer_function.R" and type the following code to import the dataset and calculate population transfer distance:

```
### Transfer functions =====  
## import data and calculate transfer distance—  
dat <-  
read.csv('htclm.csv',stringsAsFactors=T);head(dat)  
dat$mat_td <- dat$mat_s - dat$mat_p;head(dat)  
#calculating trasfer distance in MAT
```

```
> head(dat)
```

	X	site	s	lat_	s	lon_	el_s	v	pro
1	1	TE	8	53.8	98	121.	730		61
2	2	AP	2	54.9	68	126.	790		103
3	3	EE	8	49.0	85	115.	152		68
4	4	MM	2	50.9	07	120.	137		42
5	5	SK	8	52.1	8	120.	110		23
6	6	UI	7	50.3	6	119.	137		56

The calculated transfer distances are attached to the dataset as the last column as shown above.

3.1 Individual transfer functions

A transfer function built for a single site is called an individual transfer function or simply a transfer function. For the dataset that we just prepared, we need to extract the data for a test site.

```
## individual transfer function _____
#extract data for the site 'CHAP' —
```

```
chap <-
subset(dat,site=='CHAP');head(chap);dim(chap)
```

We will start with a simple linear regression to build a transfer function.

```
#perform a linear regression —
lm0 <- lm(HT20~mat_td,data=chap)
print(lm0)
```

```
> print(lm0)
```

Call:

```
lm(formula = HT20 ~ mat_td, data = chap)
```

Coefficients:

```
(Intercept) mat_td
```

```
7.2219    0.2777
```

So, the linear transfer function is: $Y = 7.219 + 0.2777 \cdot \text{mat_td}$.

Where `mat_td` is the transfer distance in mean annual temperature. We can use `'summary(lm0)'` to see the R square and P values, but it is to show these results in a plot.

Let's plot the results:

```
#plot the results —
r2 <- format(summary(lm0)$adj.r.squared,digits=4);r2
p <-
format(summary(lm0)$coefficients[2,4],digits=4);p
legend_r2 <- paste0('r2=',r2)
legend_p <- paste0('P=',p)
plot(HT20~mat_td,data=chap,xlab='Transfer distance
in MAT (°C)',ylab='HT20 (m)') #scatter plot
abline(lm1) #add regression line
legend('bottomright',c(legend_r2,legend_p)) #add a
legend
```

The result of the regression analysis is shown in Figure 5.

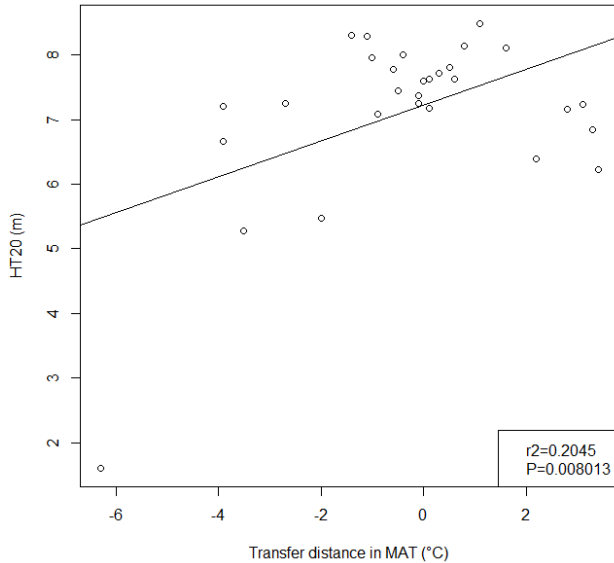


Figure 5. Plot showing the results of the linear regression between tree height (HT20) and population transfer distance. Image by Tongli Wang under CC BY 4.0.

From Figure 5 we can see that the linear relationship is not strong, and the data points are not well represented by the regression line. Let's try the quadratic function, which is more complicated.

```
#perform a quadratic regression —
lm2 <- lm(HT20~mat_td + I(mat_td^2),data=chap)
print(lm2)
```

```
> print(lm2)
```

Call:

```
lm(formula = HT20 ~ mat_td + I(mat_td^2), data =
chap)
```

Coefficients:

```
(Intercept)  mat_td  I(mat_td^2)
  7.81584    0.05265   -0.12818
```

So, the quadratic transfer function is: $Y = 7.82 + 0.05265 \cdot \text{mat_td} - 0.12818 \cdot \text{mat_td}^2$

Let's plot the results.

```
#more steps are required to extract P value —
r2 <- format(summary(lm2)$adj.r.squared,digits=4);r2
fs <- summary(lm2)$fstatistic
p <- as.numeric(pf(fs[1], fs[2], fs[3],lower.tail =
FALSE));p
legend_r2 <- paste0('r2=',r2)
legend_p <- paste0('P=',p)
```

```
#to generate predicted values to draw the regression
curve
x0 <-
data.frame(mat_td=seq(min(chap$mat_td),max(chap$mat_t
d),len=100));
pr <- predict(lm2,x0);
px <- data.frame(x0,pr);
plot(HT20~mat_td,data=chap,xlab='Transfer distance
in MAT (°C)',ylab='HT20 (m)');
lines(px$pr~px$mat_td, col='blue',lwd=2)
```

The results are in Figure 6. The relationship is much stronger than that from linear regression. Thus, a quadratic function is more frequently used.

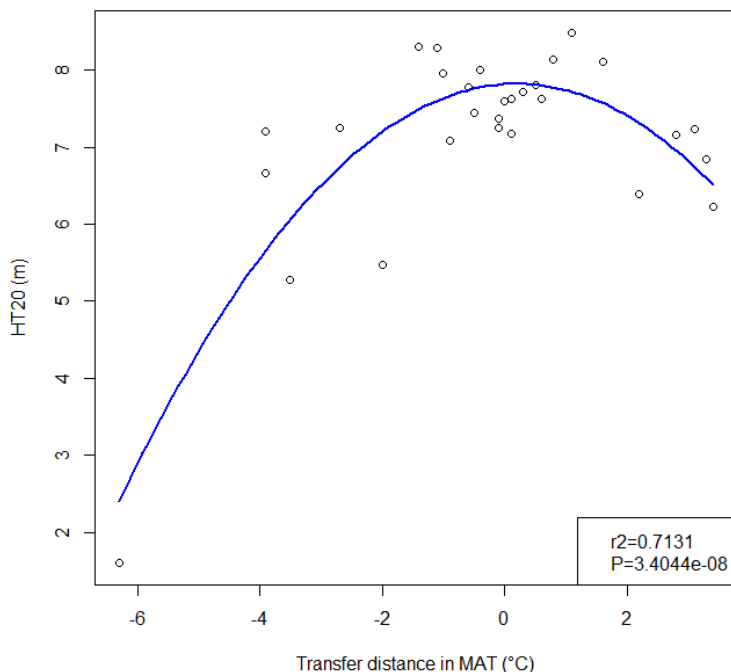


Figure 6. Plot showing the results of quadratic regression between tree height (HT20) and population transfer distance. Image by Tongli Wang under CC BY 4.0.

We have successfully built a transfer function for one test site. We can build a transfer function for other test site using the above procedure and code.

3.2 General transfer functions

A general transfer function is built on a pooled dataset across all test sites. Let's directly start with a quadratic function. The only difference in the code is that we will use the complete dataset "dat", instead of "chap".

```
## general transfer function —————
lm3 <- lm(HT20~mat_td + I(mat_td^2),data=dat);
print(lm3);
```

```
> print(lm3)
```

Call:

```
lm(formula = HT20 ~ mat_td + I(mat_td^2), data =
dat)
```

Coefficients:

```
(Intercept) mat_td I(mat_td^2)
7.07235 0.18546 -0.06375
```

The general transfer function is: $Y = 7.07 + 0.18546 \cdot \text{mat_td} - 0.06375 \cdot \text{mat_td}^2$

Next is to plot the results:

```
#plot results —
r2 <- format(summary(lm3)$adj.r.squared,digits=4);r2
fs <- summary(lm3)$fstatistic
```

```

p <- as.numeric(pf(fs[1], fs[2], fs[3],lower.tail =
FALSE));p
p <- format(p,digits=4);p
legend_r2 <- paste0('r2=',r2)
legend_p <- paste0('P=',p)
x0 <-
data.frame(mat_td=seq(min(dat$mat_td),max(dat$mat_td),l
en=100))
pr <- predict(lm2,x0)
px <- data.frame(x0,pr)
plot(HT20~mat_td,data=dat,xlab='Transfer distance in
MAT (°C)',ylab='HT20 (m)')
lines(px$pr~px$mat_td, col='blue',lwd=2)
legend('bottomright',c(legend_r2,legend_p))

```

Here is the plot (Figure 7). The function explains about 30% of the total variation ($r^2=0.2778$) although P value is reduced a lot due to an increase in sample size. This is usual for a general transfer function and is not very encouraging due to the issues we discussed in Topic 4.2 of Module IV.

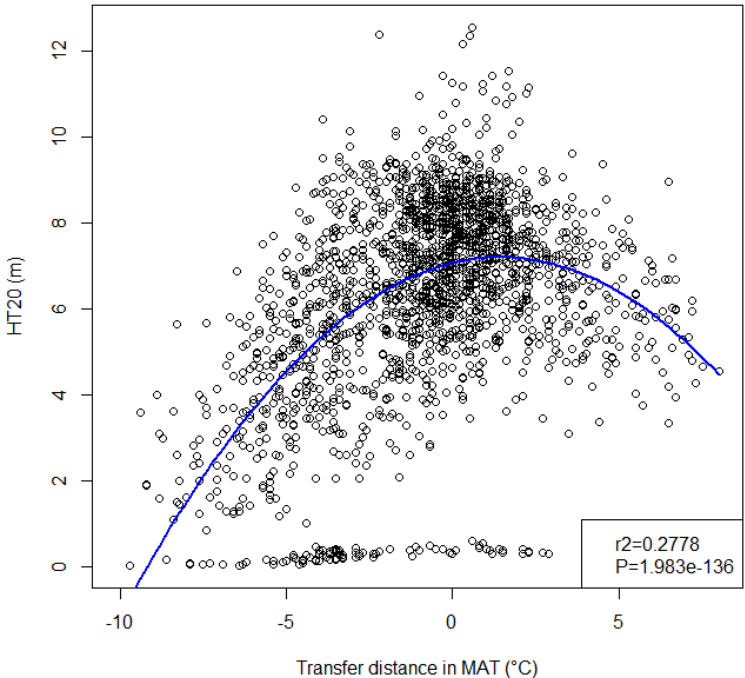


Figure 7. Plot showing the results of quadratic regression between tree height (HT20) and population transfer distance with pooled data across all test sites. Image by Tongli Wang under CC BY 4.0.

3.3 Relative performance-based general transfer functions

Please note that there are a number of data points at the bottom with HT20 close to zero showing as outliers. They are populations growing at several cold sites in the north. Differences in growth at these sites are not well represented in this plot and in the general transfer function. A relative performance-based general transfer function can overcome this problem. We need first to convert HT20 of populations

to relative height against local populations or site mean. It is often difficult to define local populations, thus we will use the site mean for the conversion in this practice.

```
## relative General transfer function —
# calculate site means —
siteHt <- dat[,c('site','HT20')];head(siteHt);
siteHtM <-
aggregate(siteHt,by=list(site=siteHt$site),FUN=mean,na.rm=T)[,c(1,3)]
names(siteHtM)[2] <- 'siteHt';
head(siteHtM);

#merge the site mean to the “dat” and calculate
relative ht
dat2 <- merge(dat,siteHtM,by='site');hd(dat2)
dat2$rht <- dat2$HT20/dat2$siteHt;hd(dat2)

# quadratic regression —
lm4 <- lm(rht~mat_td +
I(mat_td^2),data=dat2);summary(lm4)
print(lm4)
```

```
> print(lm4)
```

Call:

```
lm(formula = rht ~ mat_td + I(mat_td^2), data = dat2)
```

Coefficients:

```
(Intercept) mat_td I(mat_td^2)
```

```
1.101426 0.011685 -0.009999
```

So, the relative general transfer function is: $Y = 1.10 + 0.011685 * \text{mat_td} - 0.009999 * \text{mat_td}^2$

```
#ploting —
r2 <- format(summary(lm4)$adj.r.squared,digits=4);r2
fs <- summary(lm4)$fstatistic
p <- as.numeric(pf(fs[1], fs[2], fs[3],lower.tail =
FALSE));p
p <- format(p,digits=4);p
legend_r2 <- paste0('r2=',r2)
legend_p <- paste0('P=',p)

x0 <-
data.frame(mat_td=seq(min(dat2$mat_td,na.rm=T),max(dat
2$mat_td,na.rm=T),len=100))
pr <- predict(lm4,x0)
px <- data.frame(x0,pr)
```

```
plot(rht~mat_td,data=dat2,xlab='Transfer distance in  
MAT (°C)',ylab='HT20 (m)')  
lines(px$pr~px$mat_td, col='blue',lwd=2)  
legend('bottomright',c(legend_r2,legend_p))
```

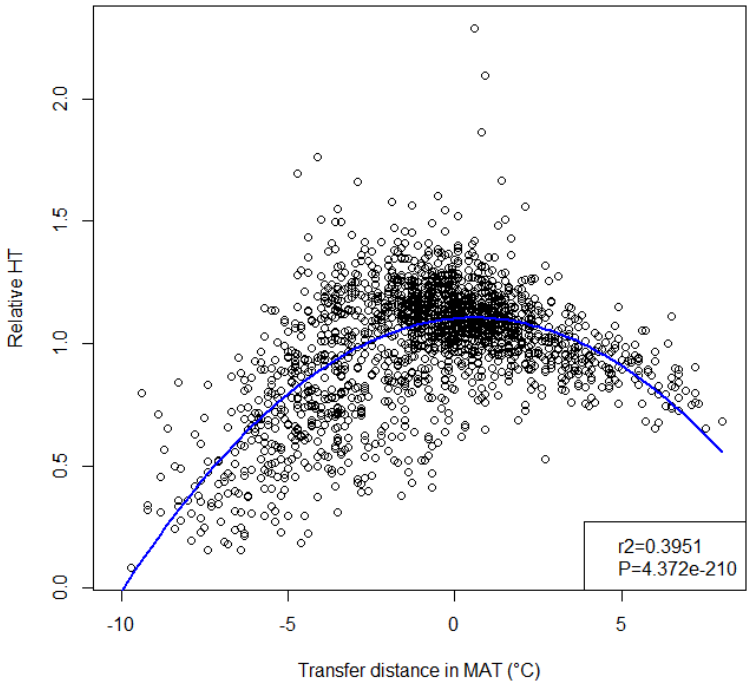


Figure 8. A general transfer function using relative performance. Image by Tongli Wang under CC BY 4.0.

Please keep in mind that, in the applications of the relative general transfer function, it may be necessary to convert the relative height back to actual height.

4. Genecology functions

The algorithms to build a genecology function are about the same as to build a transfer function. The only difference is that the independent variable is a climate variable of the provenances instead of the transfer distance. Also, no general genecology function has been reported. Thus, we will only go through the methods for an individual genecology function.

```
## genecology functions—————  
lm5 <- lm(HT20~mat_p +  
I(mat_p^2),data=chap);summary(lm5)  
print(lm5)
```

```
> print(lm5)
```

Call:

```
lm(formula = HT20 ~ mat_p + I(mat_p^2), data =  
chap)
```

Coefficients:

```
(Intercept)  mat_p  I(mat_p^2)  
7.3611      0.4857     -0.1282
```

So, the genecology function is: $Y = 7.36 + 0.4857 * \text{mat_p} - 0.1282 * \text{mat_p}^2$

Where mat_p is the mean annual temperature of provenances.

- #plot the results —
- `r2 <-
format(summary(lm5)$adj.r.squared,digits=4);r
2`
- `fs <- summary(lm5)$fstatistic`
- `p <- as.numeric(pf(fs[1], fs[2],
fs[3],lower.tail = FALSE));p`
- `p <- format(p,4);p`
- `legend_r2 <- paste0('r2=',r2)`
- `legend_p <- paste0('P=',p)`
-
- `x0 <-
data.frame(mat_p=seq(min(chap$mat_p),max(c
hap$mat_p),len=100))`
- `pr <- predict(lm5,x0)`
- `px <- data.frame(x0,pr)`
-
- `plot(HT20~mat_p,data=chap,xlab='Seed
source MAT (°C)',ylab='HT20 (m)')`
- `lines(px$pr~px$mat_p, col='blue',lwd=2)`
- `legend('bottomleft',c(legend_r2,legend_p
)`

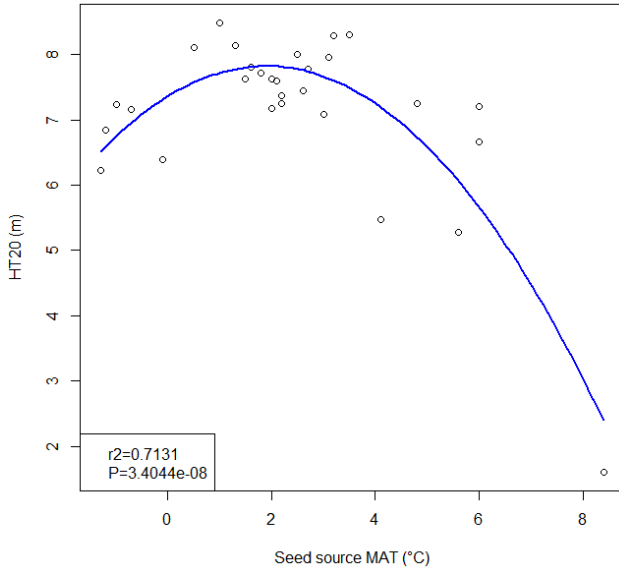


Figure 9. Plot showing the results of quadratic regression between tree height (HT20) and seed source (provenance) mean annual temperature (MAT). Image by Tongli Wang under CC BY 4.0.

Please note that the plot and the regression curve are very different from the transfer function shown in Figure 6. However, the R^2 and the P values remain the same.

Let's save and close the script file "2.transferFunction.R"

5. Response functions

A response function is built based on the relationship between the performance of a population and planting site environment, which is climate in our case. Thus, to build a response function, we first need to pull out the data for a population, $prov = 1$.

Let's start a new script file "responseFunction.R" and type the following code (blue):

```
## import data ——  
dat <-  
read.csv('htclm.csv',stringsAsFactors=T);head(dat);str(dat)  
  
# response function for pop#1 —————  
p1 <- subset(dat,prov==1);head(p1)  
  
# quadratic function —  
lm6 <- lm(HT20~mat_s +  
I(mat_s^2),data=p1);summary(lm6)  
print(lm6)
```

```
> print(lm6)
```

Call:

```
lm(formula = HT20 ~ mat_s + I(mat_s^2), data = p1)
```

Coefficients:

```
(Intercept) mat_s I(mat_s^2)
```

```
4.8525 2.1517 -0.3376
```

So, the response function for prov=1 is: $Y = 4.85 + 2.1517 * \text{mat_s} - 0.3376 * \text{mat_s}^2$

Where mat_s is the mean annual temperature of test sites.

```
# plotting the results —
r2 <- format(summary(lm6)$adj.r.squared,digits=4);r2
fs <- summary(lm6)$fstatistic
p <- as.numeric(pf(fs[1], fs[2], fs[3],lower.tail =
FALSE));p
p <- format(p,4);p
legend_r2 <- paste0('r2=',r2)
legend_p <- paste0('P=',p)

x0 <-
data.frame(mat_s=seq(min(p1$mat_s),max(p1$mat_s),len=
100))
pr <- predict(lm6,x0)
px <- data.frame(x0,pr)

plot(HT20~mat_s,data=p1,xlab='Test site MAT
(°C)',ylab='HT20 (m)')
lines(px$pr~px$mat_s, col='blue',lwd=2)
legend('bottomright',c(legend_r2,legend_p))
```

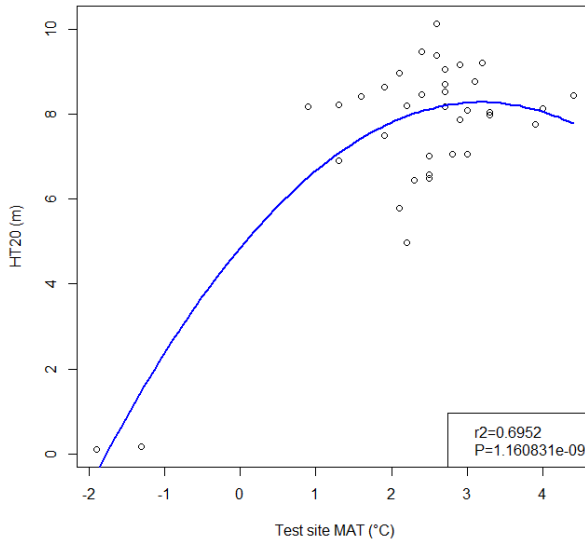


Figure 10. Plot showing the results of quadratic regression between tree height (HT20) and test site mean annual temperature (MAT). Image by Tongli Wang under CC BY 4.0.

The response function represent a strong relationship between the performance of the population and test site climate (Figure 10). However, the plot shows that the test sites are not well distributed along the climate gradient, concentrated within the range of 2-3°C.

6. Universal response functions

6.1 Model development

A universal response function (URF) is an integration of

genecolgy functions of all sites and response functions of all populations into a single response function. It is simply to build, we can simply include test site climate and provenance climate as independent variables.

```
## universal response function MAT only—
lm7 <- lm(HT20~mat_s+mat_p+ I(mat_s^2) +
I(mat_p^2)+mat_s*mat_p,data=dat);
summary(lm7)
```

```
> summary(lm7)
```

Coefficients:

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	5.092728	0.093999	54.178	< 2e-16 ***
mat_s	1.277098	0.042905	29.766	< 2e-16 ***
mat_p	0.129457	0.030607	4.230	2.46e-05 ***
I(mat_s^2)	-0.172534	0.013718	-12.578	< 2e-16 ***
I(mat_p^2)	-0.068583	0.003115	-22.018	< 2e-16 ***
mat_s:mat_p	0.046901	0.010716	4.377	1.27e-05 ***

—

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.57 on 1794 degrees of freedom
 Multiple R-squared: 0.525, Adjusted R-squared: 0.5237
 F-statistic: 396.6 on 5 and 1794 DF, p-value: < 2.2e-16

So, the URF is: $Y = 5.09 + 1.2770 * \text{mat_s} + 0.1295 * \text{mat_p} - 0.172534 * \text{mat_s}^2 - 0.06858 * \text{mat_p}^2 + 0.04690 * \text{mat_s} * \text{mat_p}$

Next, let's plot the response surface. For that, we need to install the package "RSA".

```
Install.packages('RSA')
library(RSA)
cf <- summary(lm7)$coefficients;cf
plotRSA(x=cf[2], y=cf[3], x2=cf[4],
y2=cf[5],xy=cf[6], b0= cf[1],
xlim=c(-5,8), ylim=c(-5,10),xlab='mat_s (°C)',
ylab='mat_p (°C)', zlab='HT20 (m)')
```

Here is the response surface in Figure 11:

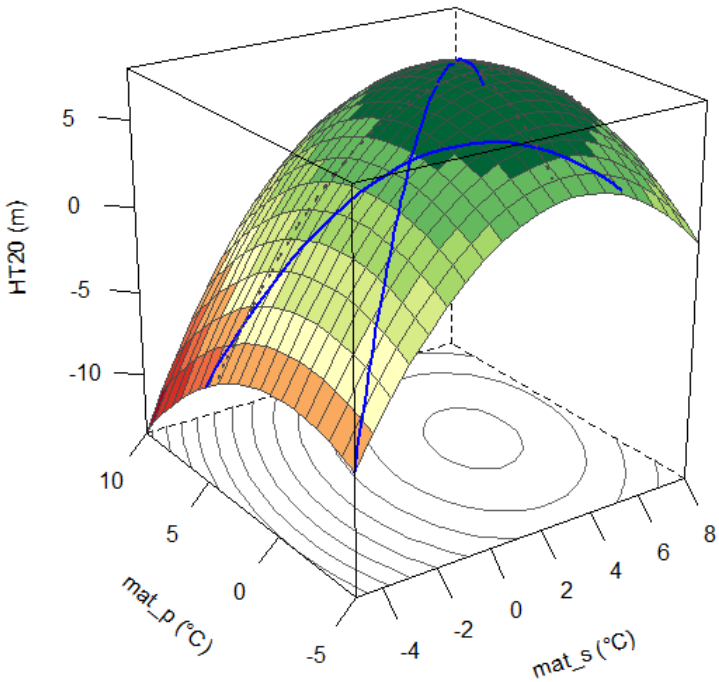


Figure 11. Response surface tree height (HT20) of populations against test site mean annual temperature (mat_s) and seed source temperature (mat_p). Image by Tongli Wang under CC BY 4.0.

6.2 Mapping

A URF can predict the performance of any population planted at any site. If we assume that we plant local seed provenances, we can map their performance over a region.

First, we need to copy the folder “ClimateData” under “NicheModels” that we created in 5.1, and paste to our current working folder “TranRespFunc”. Then we can type the following code in the script panel:

```
#import climate raster stack —
library(raster)

# import the climate data in raster stack prepared in
5.1 —
Stk6190 <- stack('ClimateData/Normal_1961_1990Y/
stk6190.grd')

#extract MAT —
mat_s <- subset(Stk6190,'MAT');mat_s

#rename the layer to match the variables in the URF
names(mat_s) <- 'mat_s';mat_s
mat_p <- mat_s
names(mat_p) <- 'mat_p';mat_p

#make a new stack for prediction —
stk2 <- stack(mat_s,mat_p);stk2

#prediction —
pre <- predict(stk2,lm7);pre
pre[pre <0 ] <- 0 #eliminate negative values

#plot the map —
plot(pre, xlab='Longitude',ylab='Latitude')
```

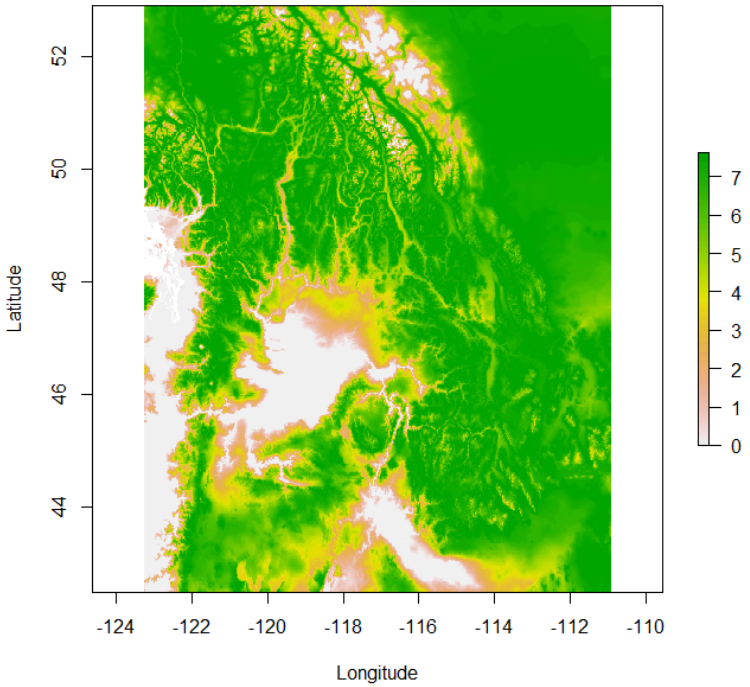


Figure 12. Map of predicted HT20 over a region assuming local provenances are used. Image by Tongli Wang under CC BY 4.0.

Now we can save and close the script file.

Appendix

This is where you can add appendices or other back matter.