Ryan Patrick McGehee
Dr. Puneet Srivastava

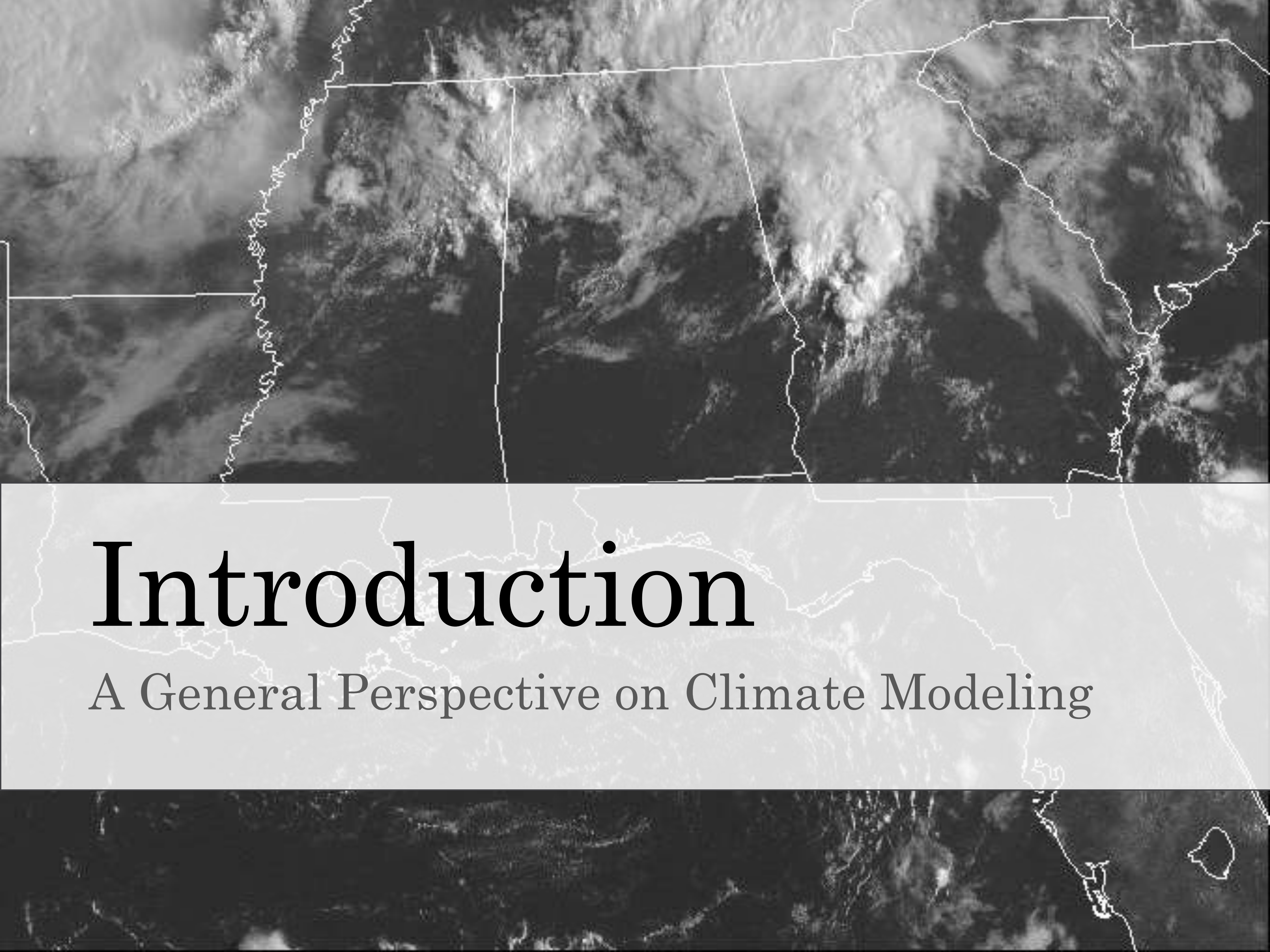
Artificial Neural Network Prediction of Future Rainfall Intensity

A Precursor to Understanding Climate Change Outcomes for the Southeastern United States

Why Rainfall Intensity?

- **Basis for Design and Development**
  - Bridges, Roads, Dams, Structures, etc.
  - Peak Discharge (TR-55) Method
  - Rational Method
  - IDF Curves and Hydrology
- **Basis for Conservation Efforts**
  - Soil Loss Tolerances (Ag, Construction, etc.)
  - USLE, RUSLE1, RUSLE2
- **Basis for Determining a Philosophical Limit for Sustainable Interaction with Our Environment and Our Climate**

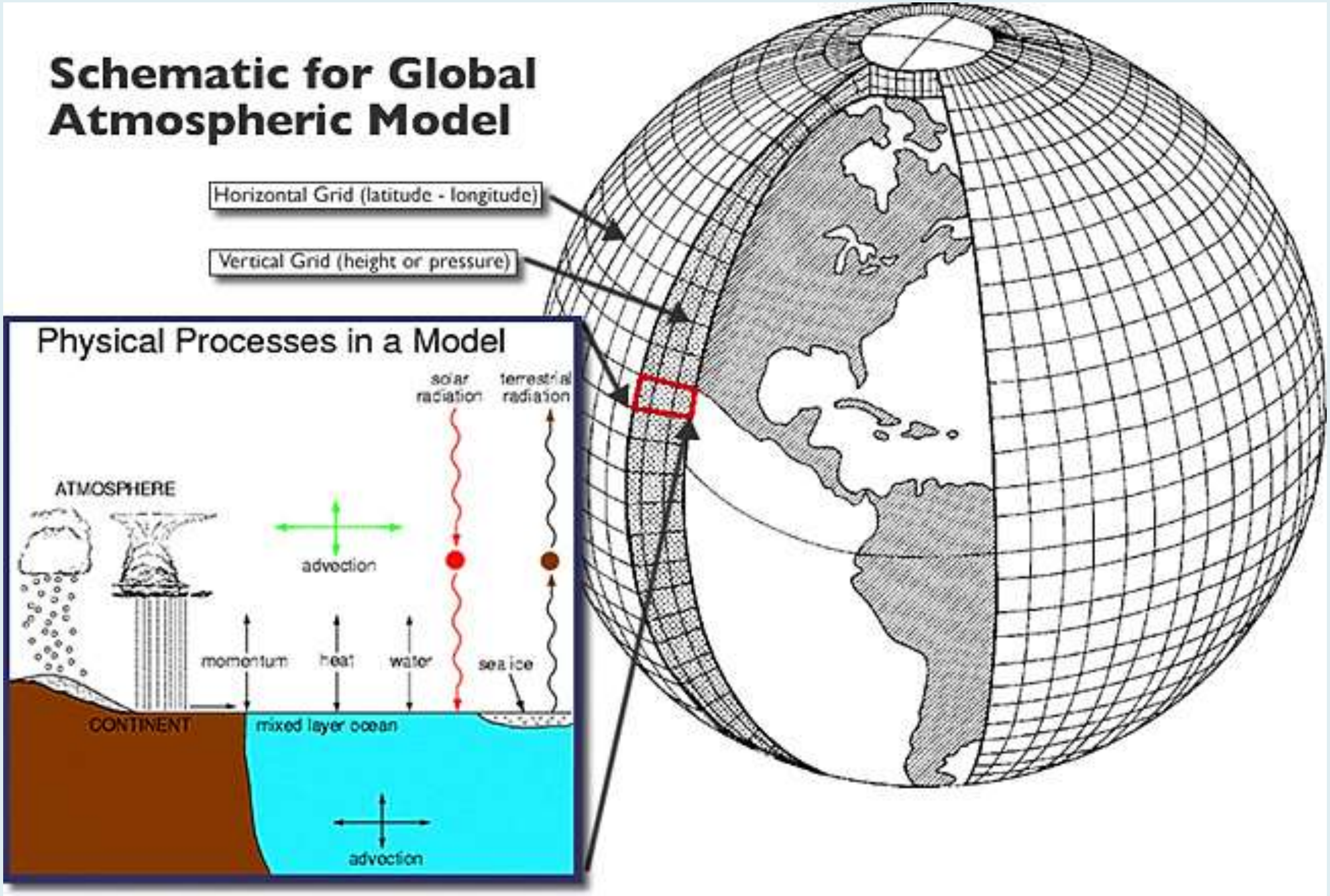
AUBURN UNIVERSITY



Introduction

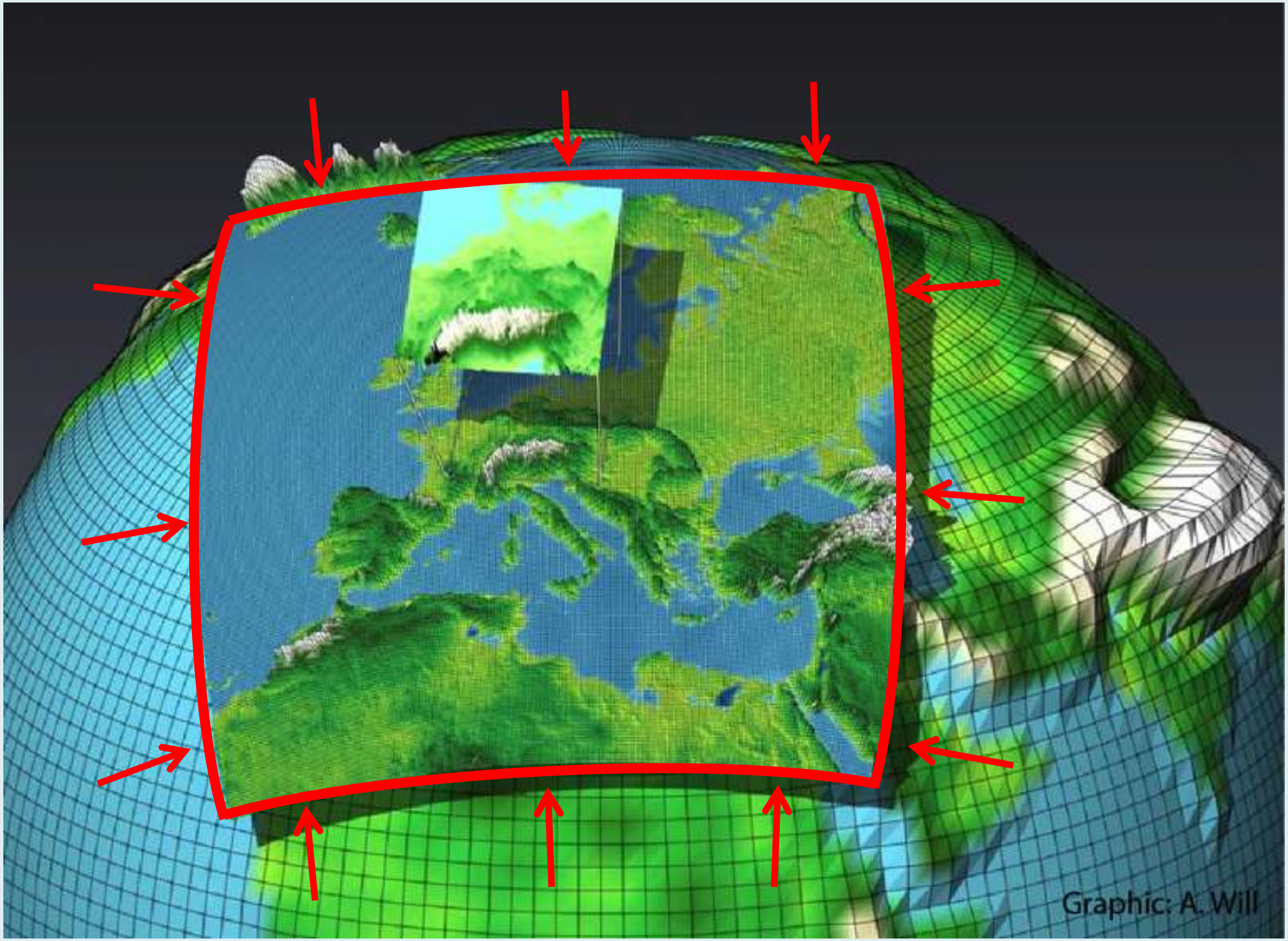
A General Perspective on Climate Modeling

AOGCMs (Global Models)



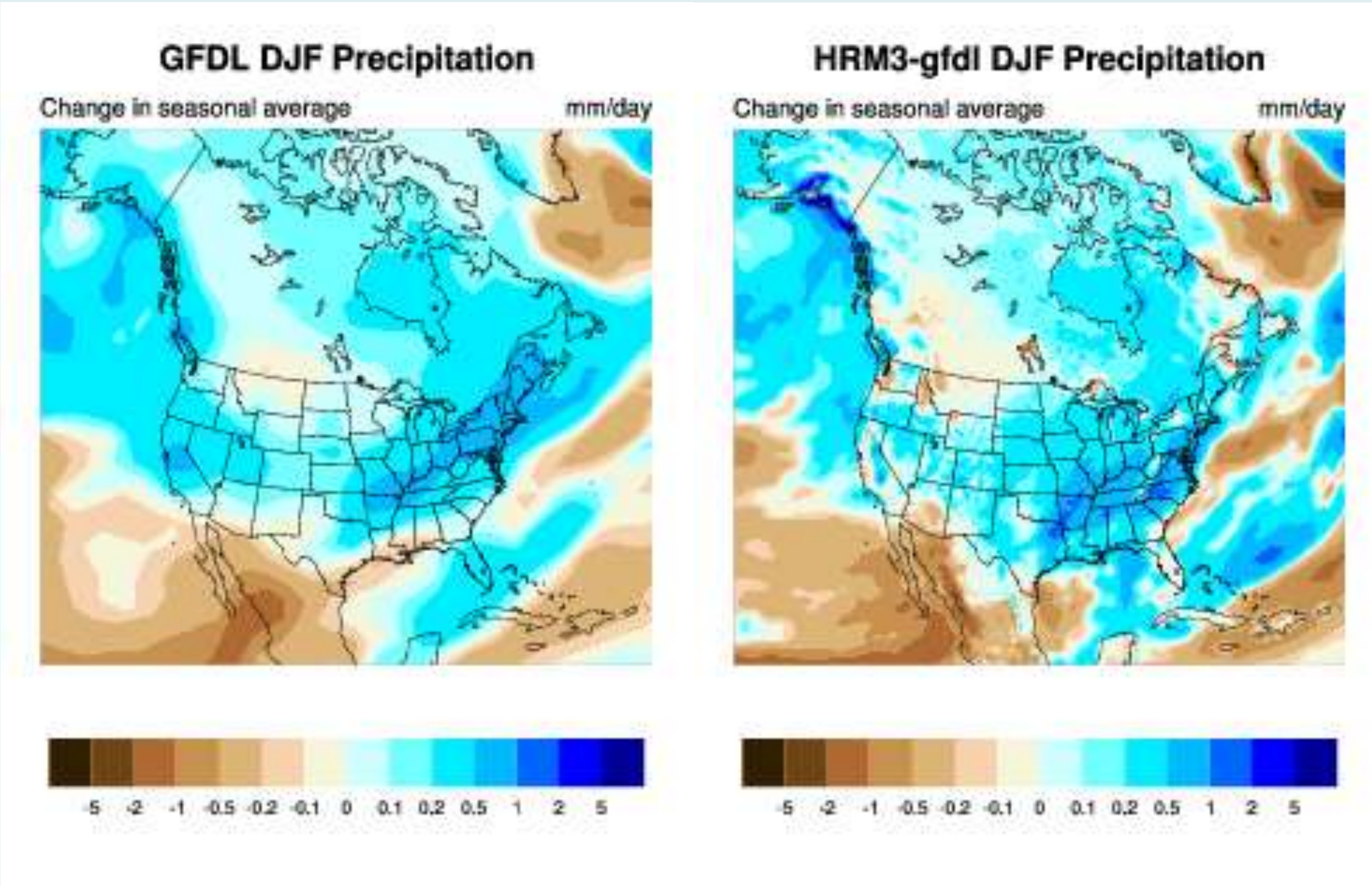
http://www.climatewizard.org/Climate_Models_and_Scenarios/Climate%20Models%20-%20Documentation_img_0.jpg

RCMs (Regional Models)



http://www.clm-community.eu/images/13_Picture1_1403507274.jpg

The Influence of Scale



http://www.narccap.ucar.edu/results/downscaling-hrm3.html

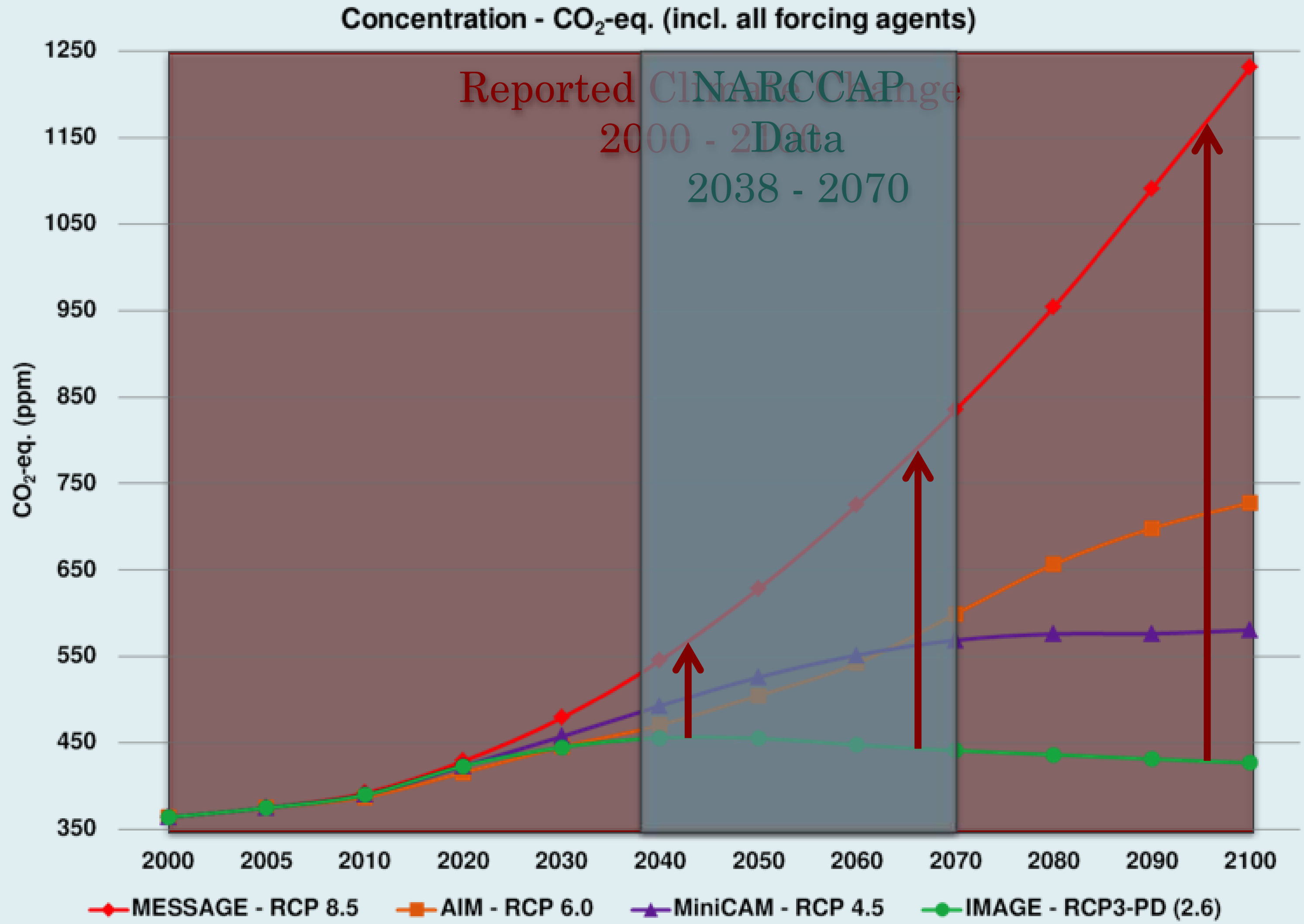
Spatial Perspective

- **Newest Hadley Center AOGCM Model (HADGEM3 — Under Development)**
  - Doubled Ocean Layers (20 to 40)
  - Doubled Atmospheric Layers (19 to 38)
  - Halved Grid Size (270 km to 135 km)
- **Regional Models (i.e. HRM3, MM5I)**
  - 50 km Resolution
- **RCMs Require Less Computation for Areas of Greater Interest (Continents)**

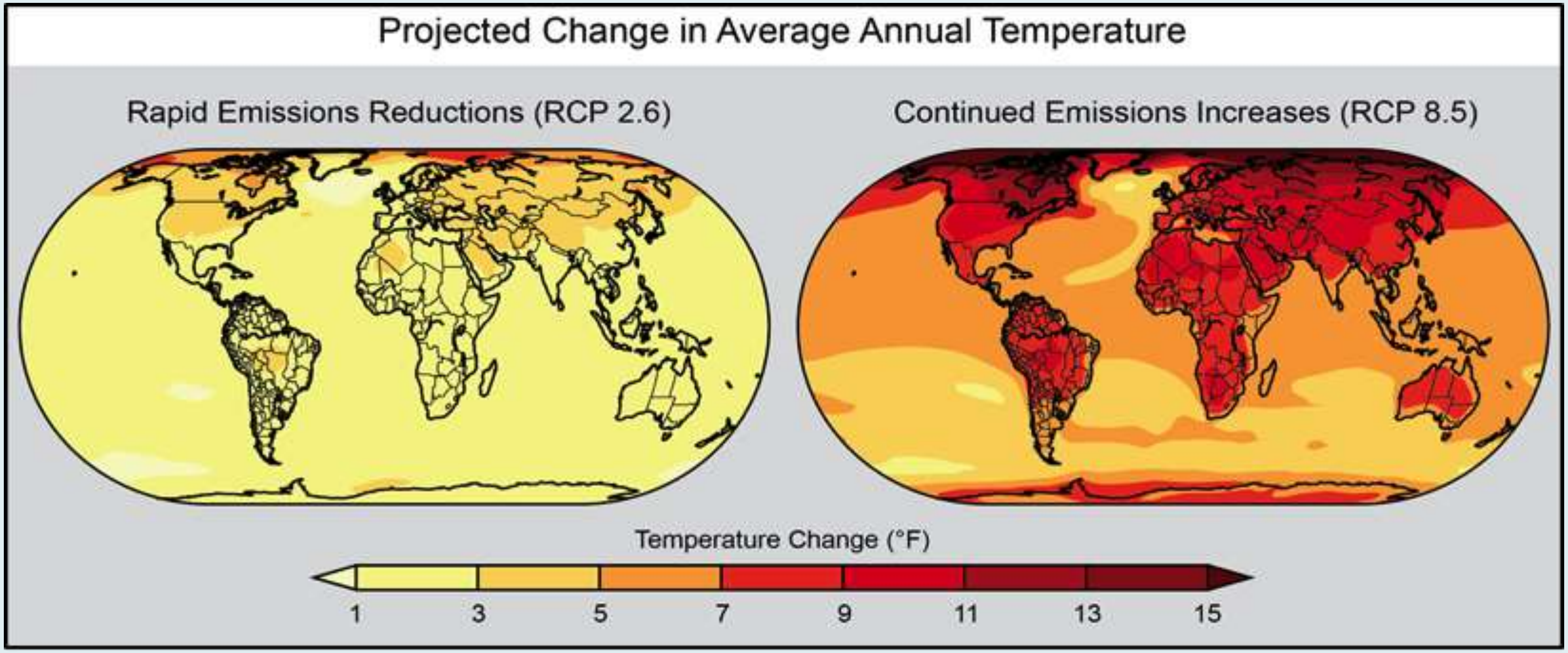
Temporal Perspective

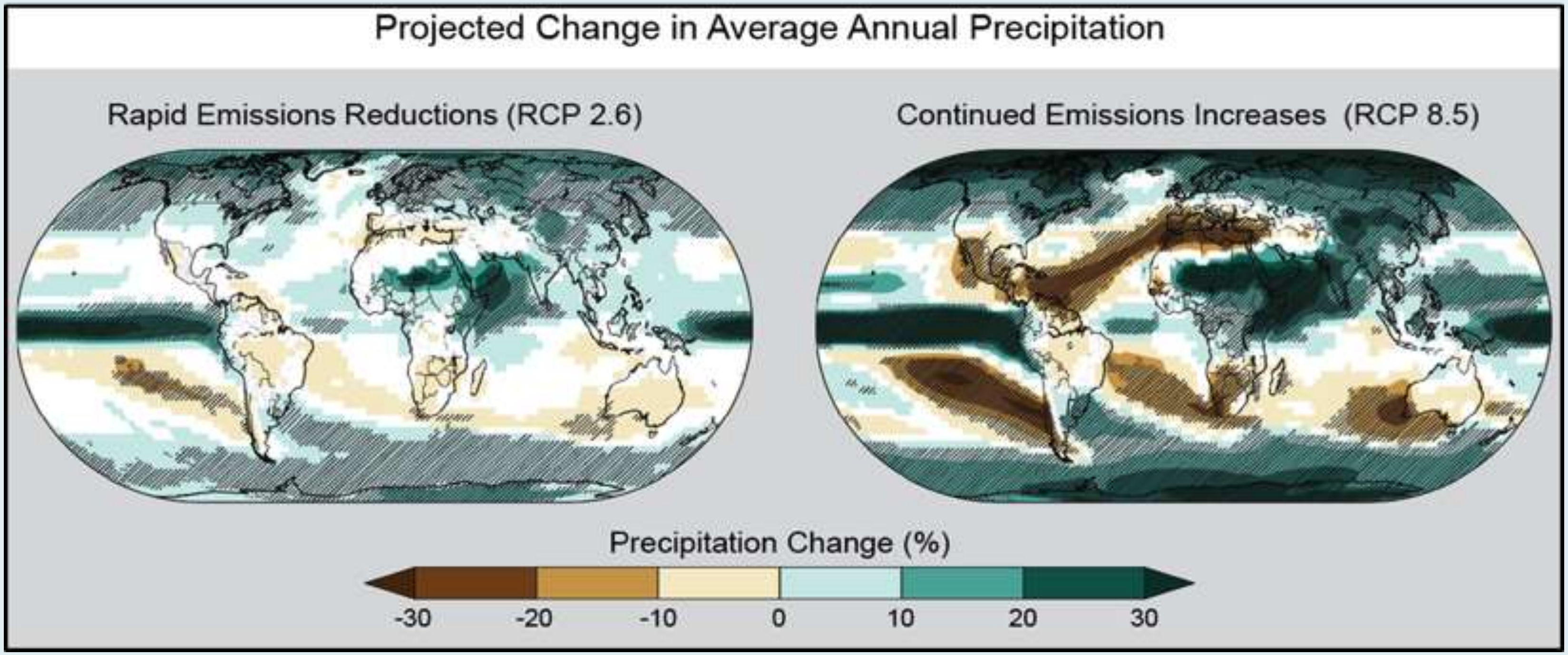
- **Models Run at Various Timesteps**
  - Usually (100 - 300 Seconds)
  - Converted to 3-Hourly Values
  - Sometimes Summed to 6-Hourly Values
- **Models Run for About 30 Years**
  - Roughly 90,000 Timesteps
  - Huge Data / Computation Limitations
- **Timestep has a Significant Effect on Precipitation Modeling (Intensity)**

Forcing Scenarios

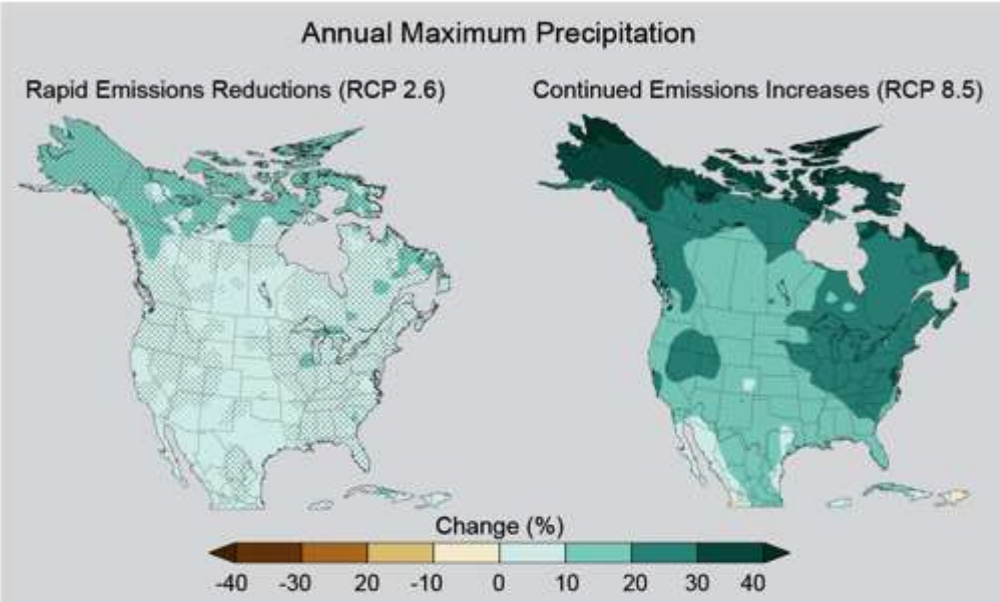


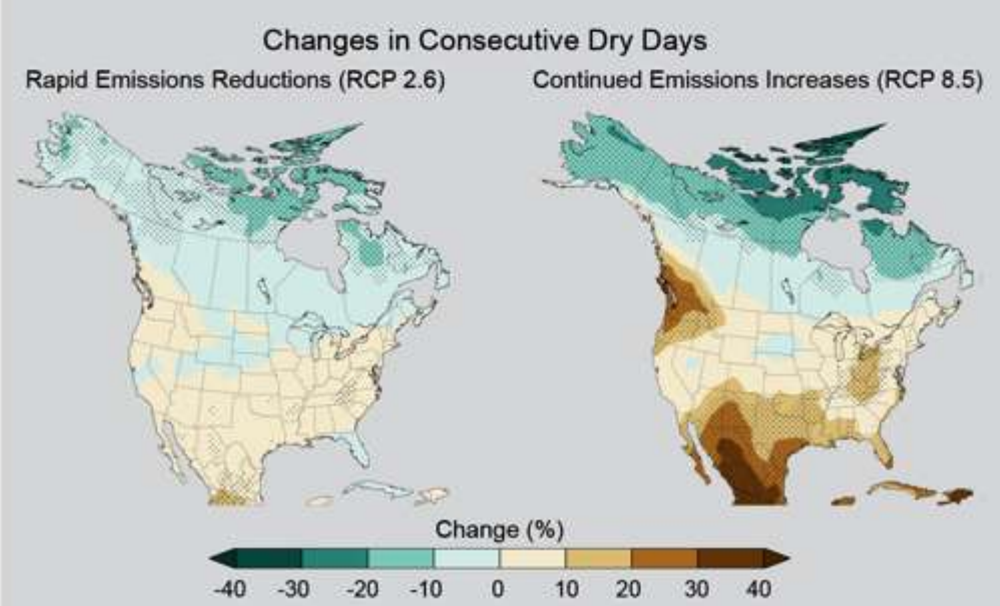
https://en.wikipedia.org/wiki/Representative_Concentration_Pathways#/media/File:All_forcing_agents_CO2_equivalent_concentration.png

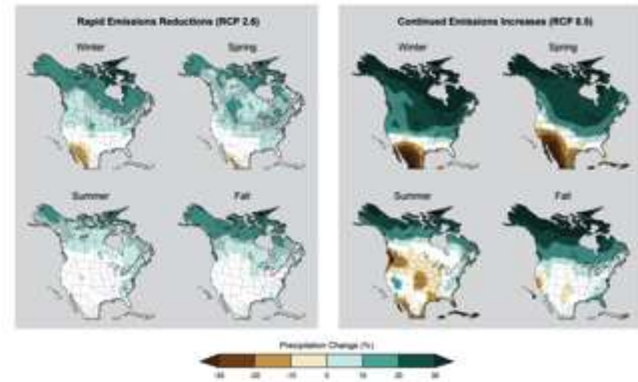


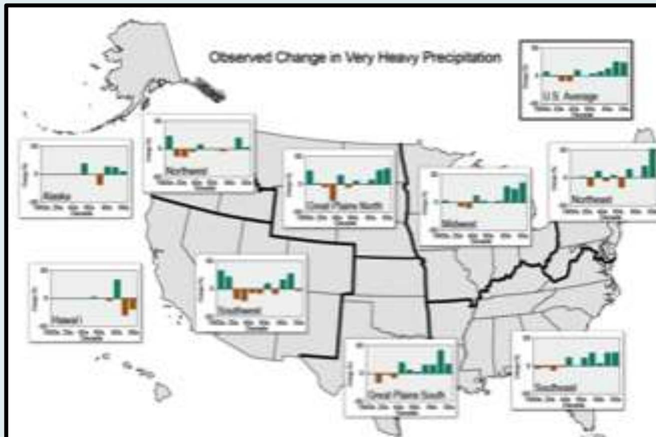


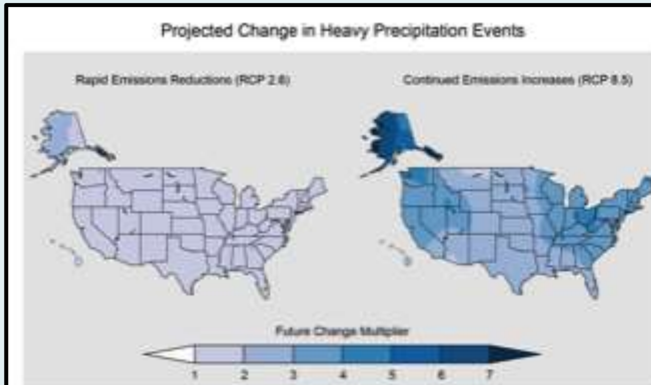
AUBURN UNIVERSITY


Annual Maximum Precipitation
Rapid Emissions Reductions (RCP 2.6)
Continued Emissions Increases (RCP 8.5)
Change (%)
-40 -30 20 -10 0 10 20 30 40


Changes in Consecutive Dry Days
Rapid Emissions Reductions (RCP 2.6)
Continued Emissions Increases (RCP 8.5)
Change (%)
-40 -30 -20 -10 0 10 20 30 40


Rapid Emissions Reductions (RCP 2.6)
Continued Emissions Increases (RCP 8.5)
Winter
Spring
Summer
Fall


Observed Change in Very Heavy Precipitation
U.S. Average
Alaska
Northwest
Great Plains North
Midwest
Northeast
Hawaii
Southwest
Great Plains South
Southeast


Projected Change in Heavy Precipitation Events
Rapid Emissions Reductions (RCP 2.6)
Continued Emissions Increases (RCP 8.5)
Future Change Multiplier
1 2 3 4 5 6 7

Characteristics of Rain

- **From One Precipitation Variable (pr)**
  - Annual Average Rainfall
  - Annual Maximum Rainfall
  - Consecutive Dry Days
  - Seasonal / Monthly Rainfall
  - Size of Heavy Downpours
  - Return Period of Heavy Downpours
- **Many Measures of Climate Are Not Direct Model Outputs**
- **Climate Measures Are Harvested From Outputs in Sometimes Complex Procedures**

Motivation

"As climate changes, the main changes in precipitation will likely be in the intensity, frequency, and duration of events, but these **characteristics are seldom analyzed** in observations or models."

Bulletin of the American Meteorological Society
September 2003

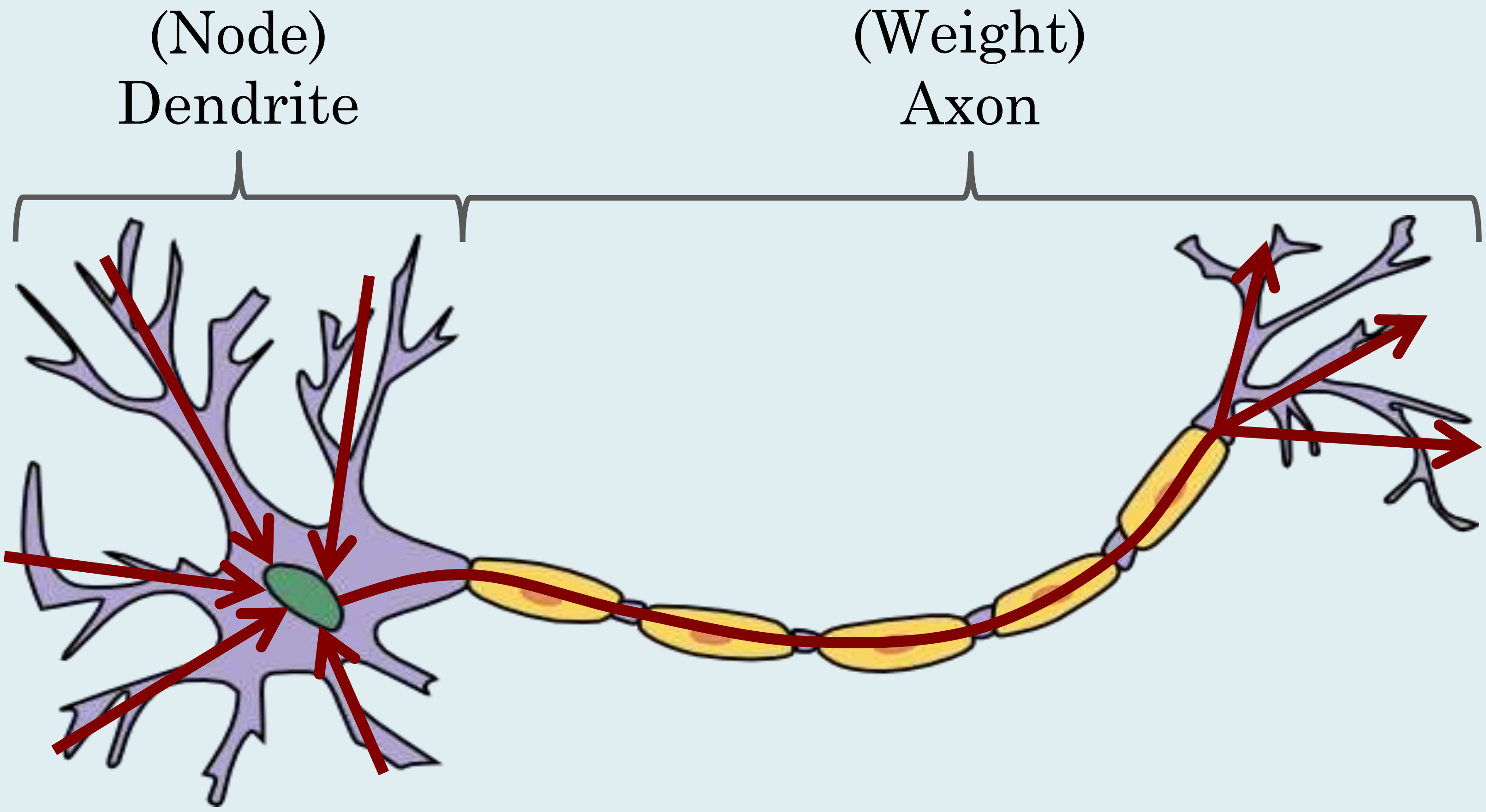
AUBURN UNIVERSITY

Background

The Foundation of Artificial Neural Networks

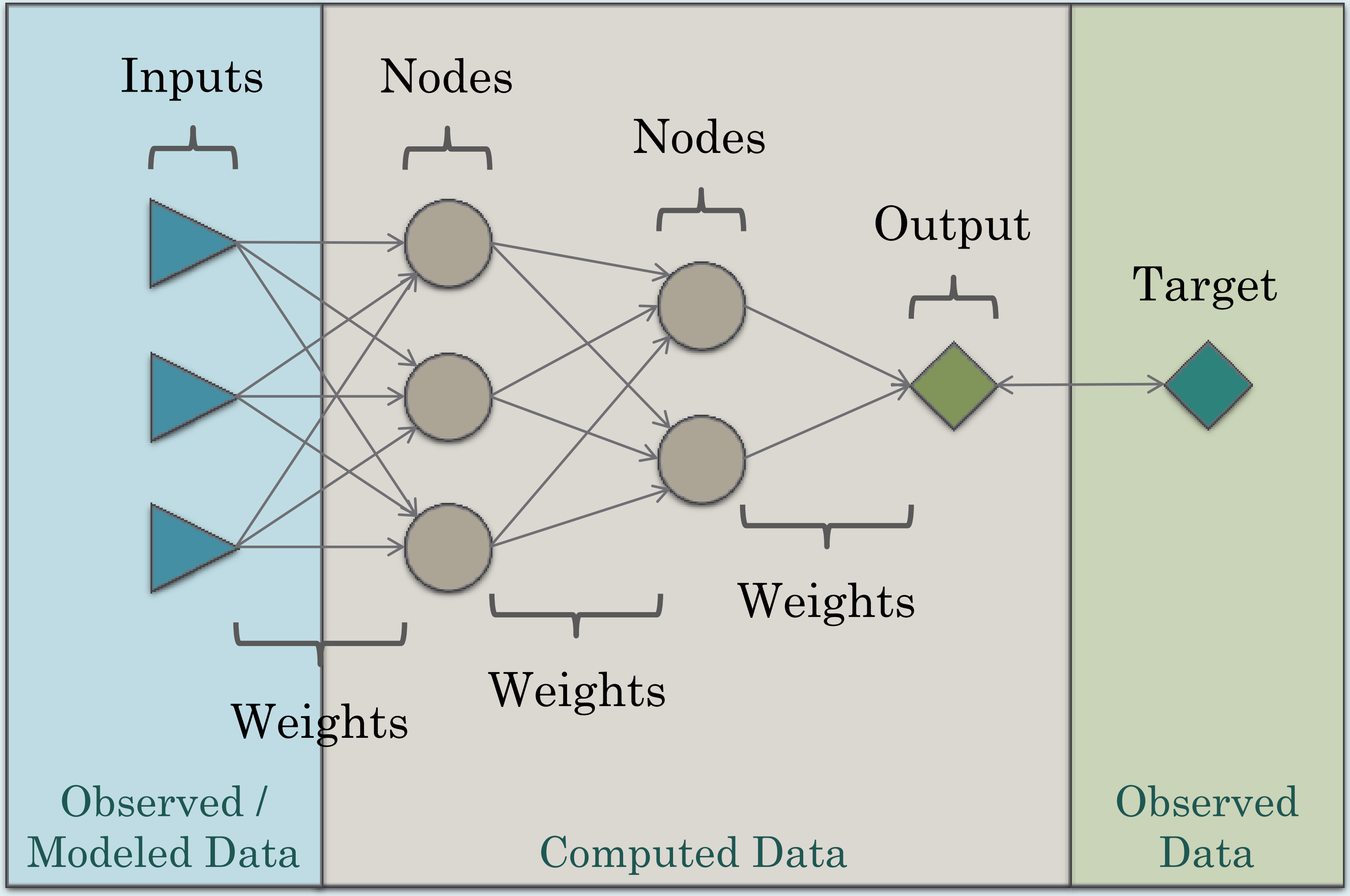
The Biological Neuron

- Neurons are the building blocks of ANNs



http://www.clker.com/cliparts/1/5/b/3/12067397712024942328neuron.svg.hi.png

Artificial Neural Network (ANN)



ANN Development

- **Step 1** — Training Stage
- **Step 2** — Optimization Stage
- **Step 3** — Execution Stage

Requirements of the ANN

- **For Successful Training:**
  - Data with Recorded Changes of Intensity
  - Inputs for Intensity Data
  - Inputs Cause the Change in Intensity
- **For Successful Optimization:**
  - Good Measure of Generalization Error
  - Balance of Overfitting and Underfitting
- **For Successful Execution:**
  - Prediction Data Formatted Identically to the Training Input Data

Methodology

Developing an ANN Configuration

ANN Development

Step 1
- Selection of Targets
- Selection of Inputs

Step 2
- Test Training Configuration(s)
- Select Best Configuration

Step 3
- Selection of Inputs
- Run ANN Configuration

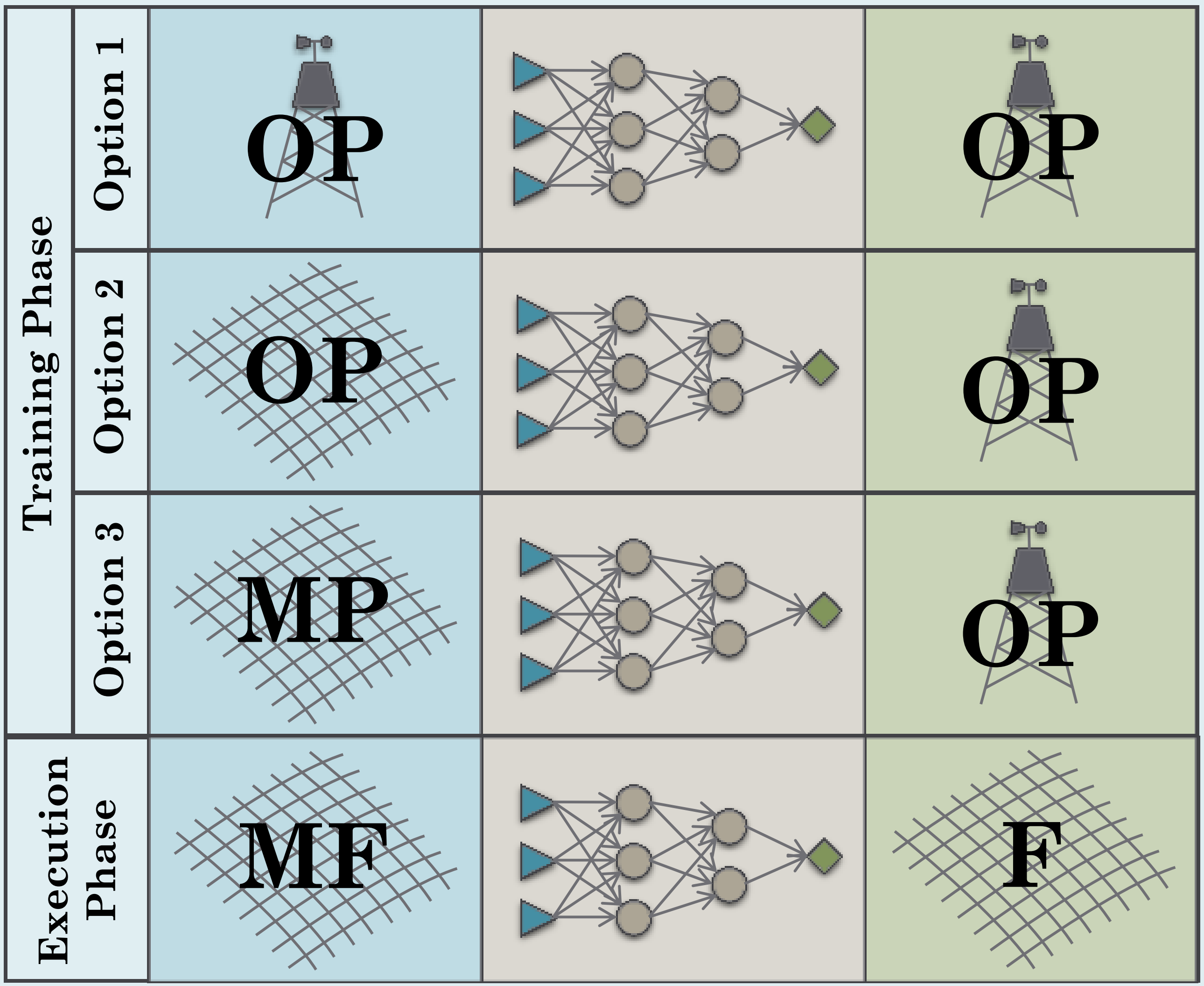
Target Selection

- **Determine Future Rainfall Intensity**
  - NOAA NCDC Station Data
  - Observed 15-Minute Intensity Data
  - Calculate Maximum 30-Minute Intensity
- **Approach to Training**
  - Highly Affected by Research Goal
  - Need the Intensity (I_{30}) to Find EI_{30}
  - EI_{30} is Used for Determining Erosivity
  - Thesis Work in Soil Loss as CC Outcome
- **I_{30} (from Station Data) is the Target**

Input Selection

- **Multiple Training Input Options**
  - Observed Data (Historical Period)
    - Station Data – Point Format
    - Gridded Data – Raster Format
  - Modeled Data (Historical Period)
- **Approach to Training**
  - Observed Data Types Probably Better
  - Station Data has Less Uncertainty
  - Gridded Data has Better Coverage
  - Expensive Computation of Gridded Data
- **Use All Three Methods Beginning with the Simplest and Most Accurate**

ANN Development Overview



P – Present
F – Future
O – Observed
M – Modeled
– Inputs
– Outputs

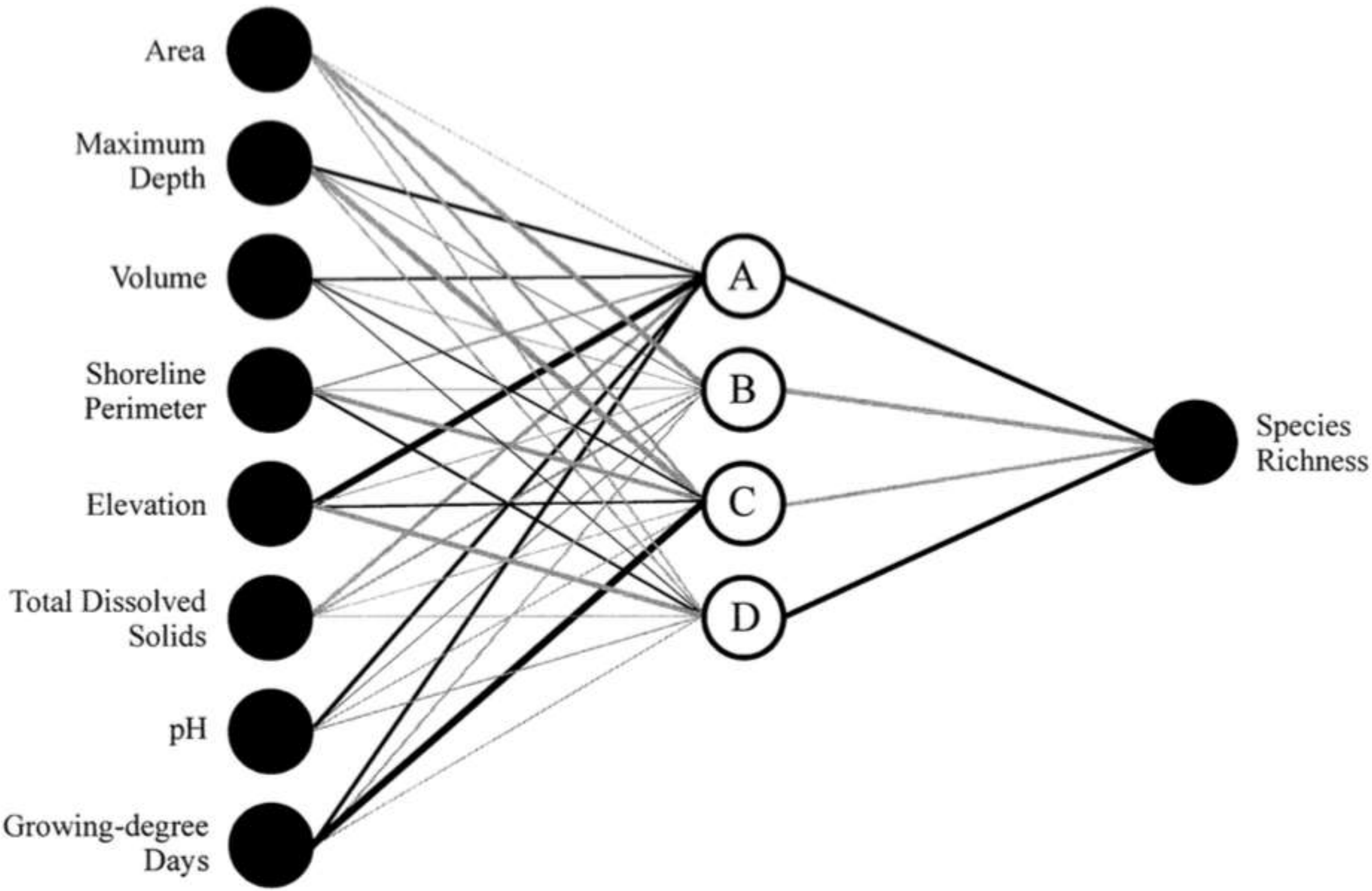
Results

Visualizing and Communicating Outputs

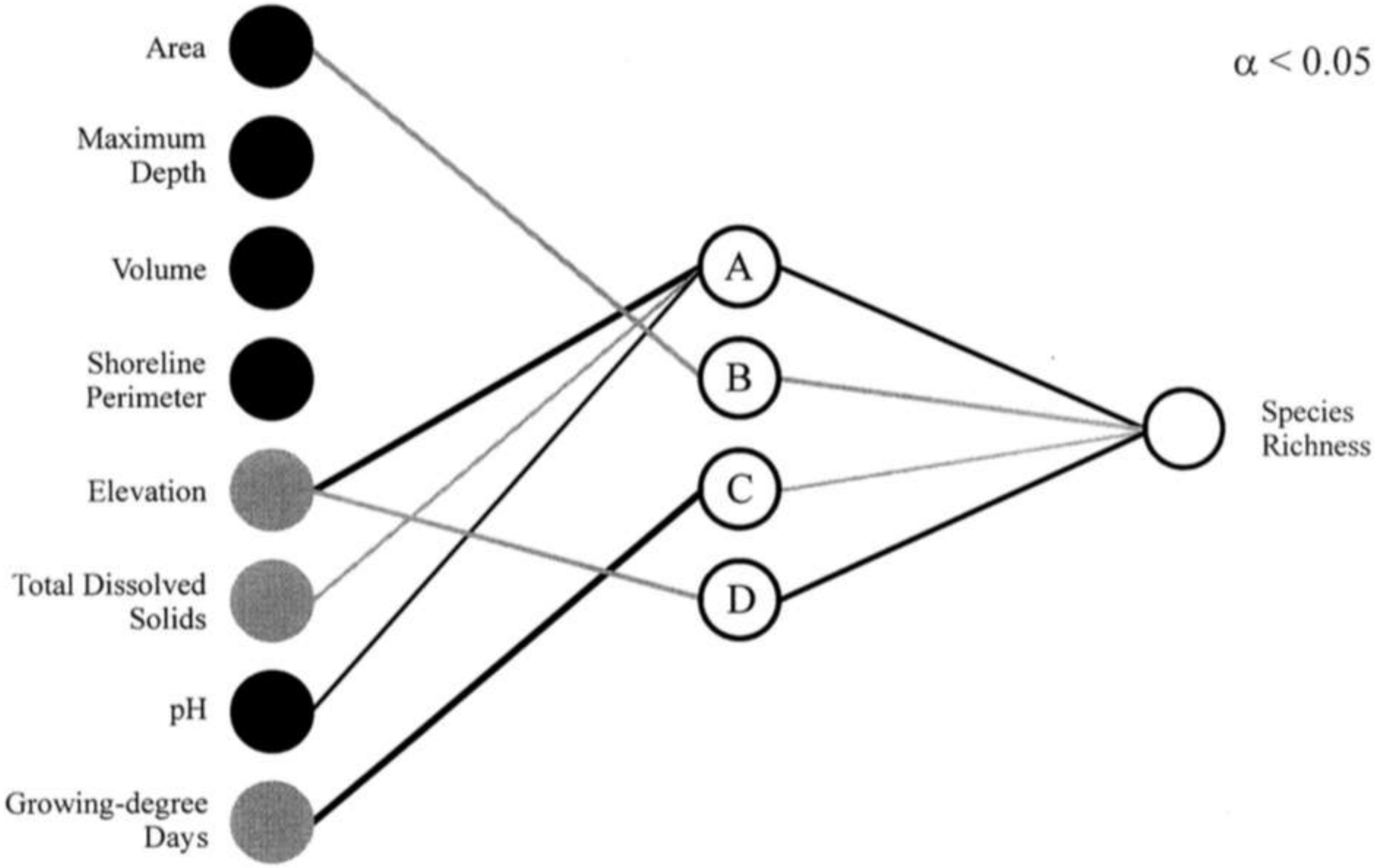
“Illuminating the Black Box”

- **Qualitative Reporting**
  - Neural Interpretation Diagrams (NIDs)
  - Randomization Test for ANNs
- **Quantitative Reporting**
  - Garson’s Algorithm
  - Sensitivity Analysis
- **J.D. Olden, D.A. Jackson / Ecological Modeling 154 (2002) 135-150**
- **Solve a Specific Problem with ANN Random Weighting (No Convergence)**

Neural Interpretation Diagram



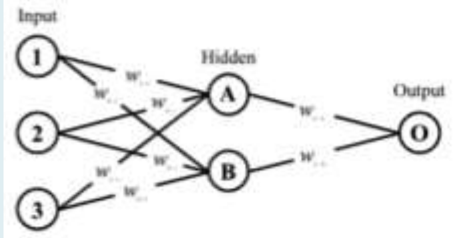
Randomization Test (95%)



Garson's Algorithm

1. Matrix containing input-hidden-output neuron connection weights

	Hidden A	Hidden B
Input 1	w = -2.61	w = -1.23
Input 2	w = 0.13	w = -0.91
Input 3	w = -0.69	w = -2.09
Output	w = 1.11	w = 0.39



2. Contribution of each input neuron to the output via each hidden neuron calculated as the product of the input-hidden connection and the hidden-output connection:
e.g., c = w x w = -2.61 x 1.11 = -2.90

	Hidden A	Hidden B
Input 1	c = -2.90	c = -0.48
Input 2	c = 0.14	c = -0.35
Input 3	c = -0.77	c = -0.82

3. Relative contribution of each input neuron to the outgoing signal of each hidden neuron: e.g., r = $|c|/(|c| + |c| + |c|)$= 2.90 / (2.90 + 0.14 + 0.77) = 0.76;
and sum of input neuron contributions: e.g., S = r + r = 0.76 + 0.29 = 1.05

	Hidden A	Hidden B	Sum
Input 1	r = 0.76	r = 0.29	S = 1.05
Input 2	r = 0.04	r = 0.21	S = 0.25
Input 3	r = 0.20	r = 0.50	S = 0.70

4. Relative importance of each input variable:
e.g., RI = S / (S + S + S) x 100
= 1.05 / (1.05 + 0.25 + 0.70) x 100 = 52.5 %

	Relative importance
Input 1	52.5 %
Input 2	12.5 %
Input 3	35.0 %

Conclusions

Reporting Practical Findings

Artificial Neural Networks

- **They Are Not Black Boxes!**
- **Qualitative and Quantitative Results**
- **Ideal in Many Ecological Scenarios**
  - Numerous Inputs
  - Complex Causal Relationships
  - Easily Measured Outputs / Targets
  - Strong Predictive Power (Climate)
  - Small Data is Best
- **Training and Optimization is Critical**

Rainfall Intensity

- **Easy to Measure and Calculate**
- **Causal Relationships**
  - Perhaps Complex
  - Not Modeled Directly by AOGCMs
- **Requires Huge Data**
  - Small Temporal Scale
  - Long Observation or Modeling Periods
  - Perhaps Multiple Variables
- **Cannot Expect Rain to Follow Same Pattern Under a New Climate Regime**

AUBURN UNIVERSITY

Change in the Southeast

- **Precipitation Change**
  - Slightly More Rain in Most States
  - More Frequent Fall / Winter Rain
  - Less Frequent Spring / Summer Rain
  - Increasing Consecutive Dry Days
  - Larger Maximum Rainfall Events
- **Change Outcomes**
  - Less Soil Moisture (Increasing Withering)
  - Highly Erosive Conditions (Soil Loss)
  - Increased Runoff, Flooding, Sedimentation
  - Increased Drought Vulnerability & Irrigation
  - Decreased Surface and Groundwater

Discussion

Additional Content on Following Slides

Determining Soil Loss

$$A = \left(\sum_{i=a}^{365m} a_i \right) \Big/ m$$

$$a_i = r_i k_i l_i S c_i p_i$$

$$a_u = rK$$

$$r_k = f_k R$$

$$R = \sum R_m$$

$$R_{m(j)} = \alpha_{(j)} P_{md(j)}$$

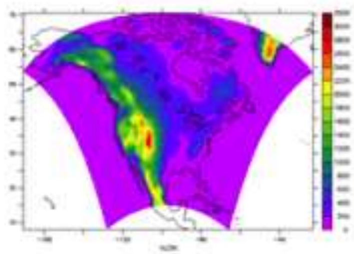
$$\alpha = \frac{\sum_{i=1}^{n} E_{(i)} I_{30(i)}}{\sum P_{15}}$$

$$E = \sum_{k-1}^{m} e_k \Delta V_k$$

$$e_k = 0.29[1 - 0.72 \exp(-0.082 i_k)]$$

NARCCAP Data

North American Regional Climate Change Assessment Program (NARCCAP)



The NARCCAP dataset contains high-resolution climate change scenario simulation output from multiple RCMs (regional climate models) nested within multiple AOGCMs (atmosphere-ocean general circulation models) for 30-year current and future periods.

The RCMs are run at 50-km spatial resolution over a domain convering the conterminous United States and most of Canada; results are recorded at 3-hourly intervals. The driving AOGCMs are forced with the A2 SRES emissions scenario in the future period. This dataset also include output from two timeslice experiments and a set of 25-year RCM simulations driven with NCEP-2 reanalysis data. These simulation results are useful for impacts analysis, further downscaling experiments, and analysis of model performance and uncertainty in regional scale projections of future climate.

When publishing research based on NARCCAP data, please include a citation for the dataset itself, such as the following:

> Mearns, L.O., et al., 2007, updated 2014. *The North American Regional Climate Change Assessment Program dataset*, National Center for Atmospheric Research Earth System Grid data portal, Boulder, CO. Data downloaded 2015-06-17. [doi:10.5065/D6RN35ST]

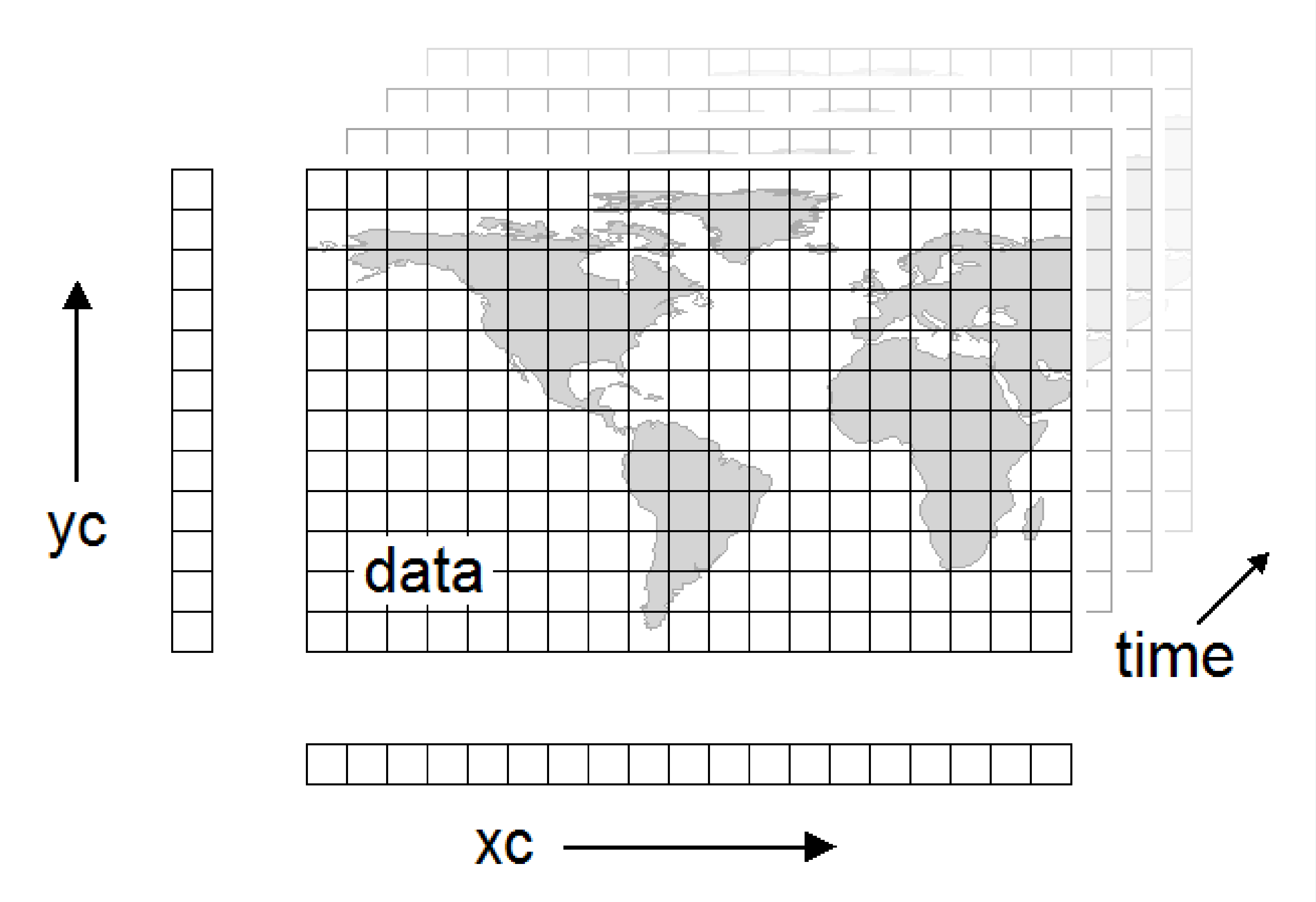
NARCCAP Homepage
Model Information

This dataset is open access. Registration is not required, but we encourage NARCCAP data users to share their research interests at the NARCCAP User Directory.

RCM	Driving Model				
	NCEP	CCSM	CGCM3	GFDL	HadCM3
CRCM	data	data	data		
ECP2	data			data	data
HRM3	data			data	data
MM5I	data	data			data
RCM3	data		data	data	
WRFG	data	data	data		
Timeslice		data		data	
ECPC	data				
WRFP	data				

Download HadCM3 Boundary Condition Data

Data Structure



Study Data Sources

- **AOGCMs**
  - GFDL
  - CCSM
  - HADCM3
- **RCMs**
  - HRM3
  - MM5I
- **Observed Data**
  - NCEP (II) – Gridded Reanalysis Data
  - NOAA NCDC – Observed Station Data

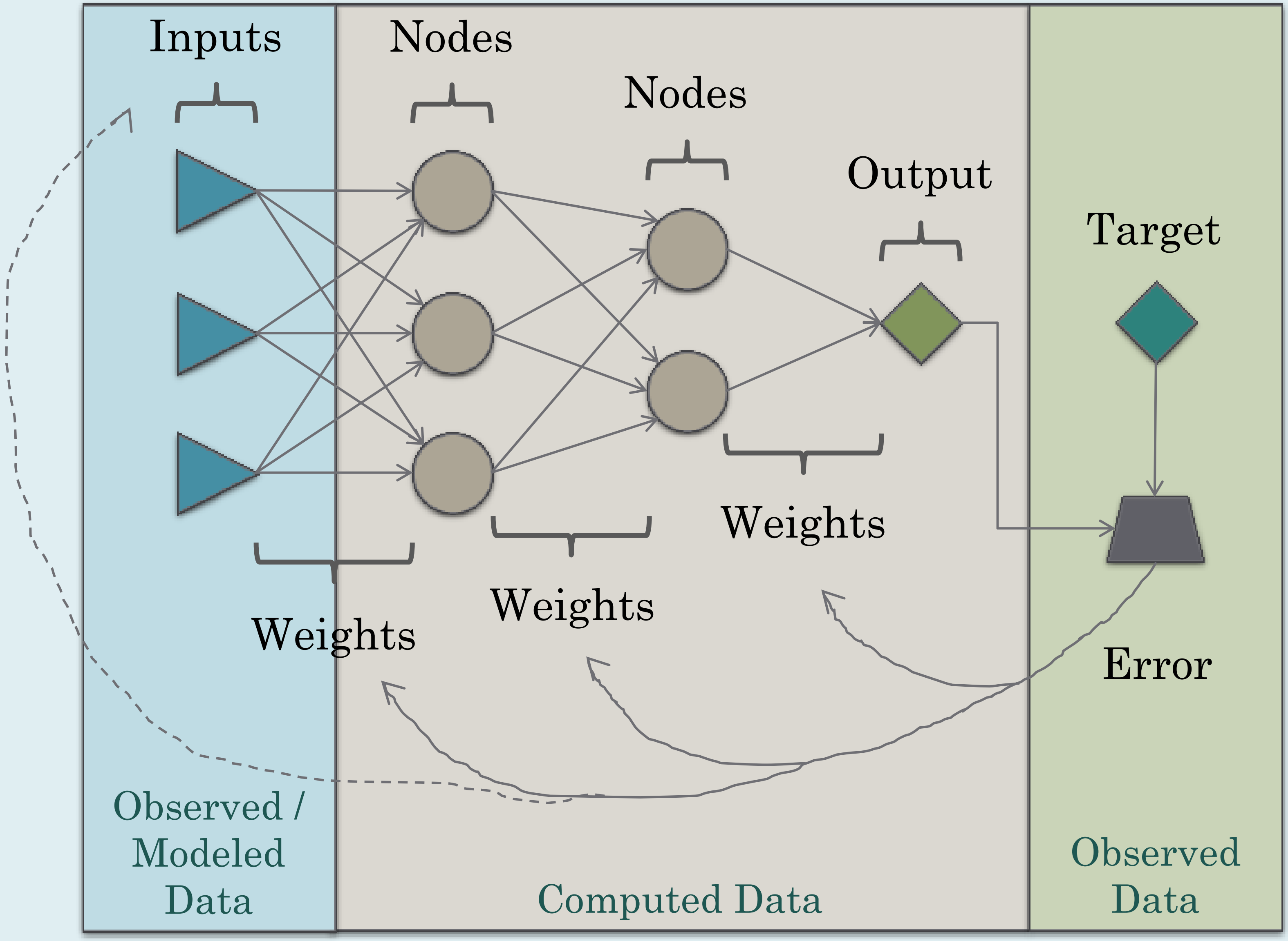
AUBURN UNIVERSITY

Review / Notes

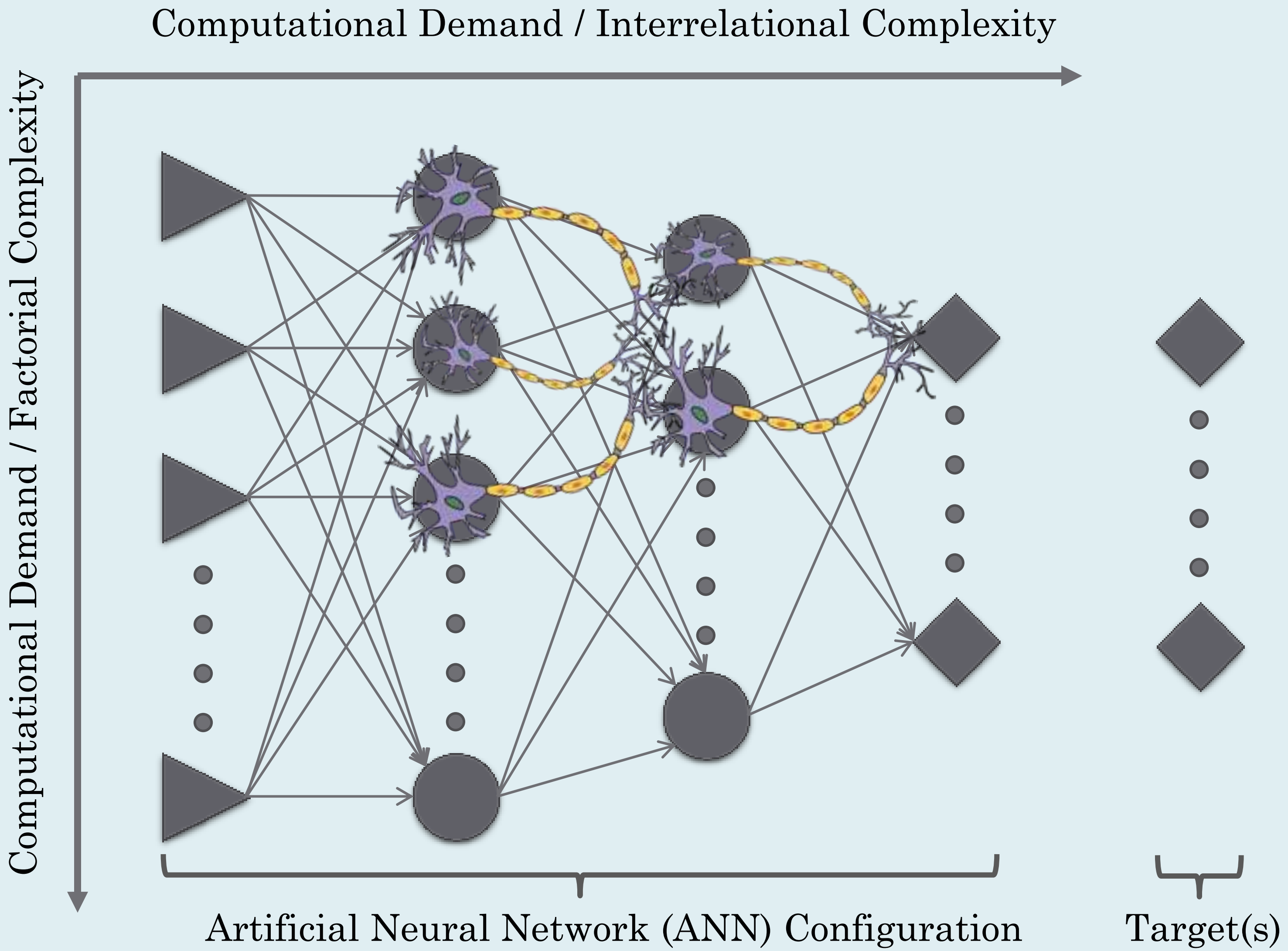
- Models run on various forcing conditions
- Results reporting usually looks at worst case and best case (i.e. RCP 8.5 and 2.6)
- Regardless of forcing conditions, trends still can be seen in the direction of climate change even when magnitude is uncertain
- In addition to forcing conditions, there are uncertainties in physical relationships (due to a lack of knowledge) as well as model limitations such as scale (due to a lack of computing power)

AUBURN UNIVERSITY

Artificial Neural Network (ANN)



ANN Complexity



Validation Techniques

- **Holdout Cross Validation**
  - Simple and Most Common
  - Training Set and Validation Set
- **K-Fold Cross Validation**
  - More Involved and Not Common
  - Rotates Data Partitioning
  - All Data is Used for Validation
  - Rotations / Partitions Determined by K

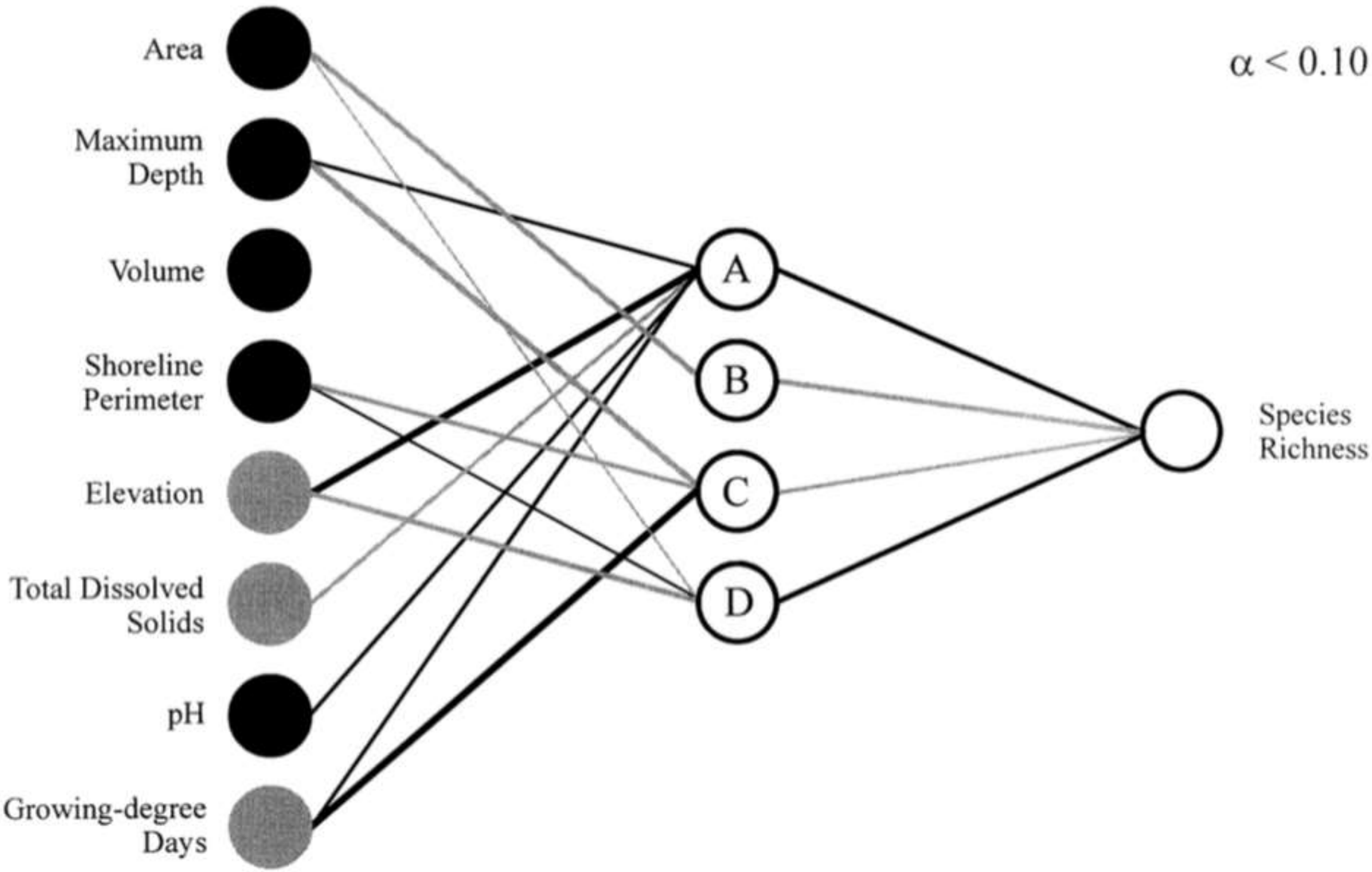
AUBURN UNIVERSITY

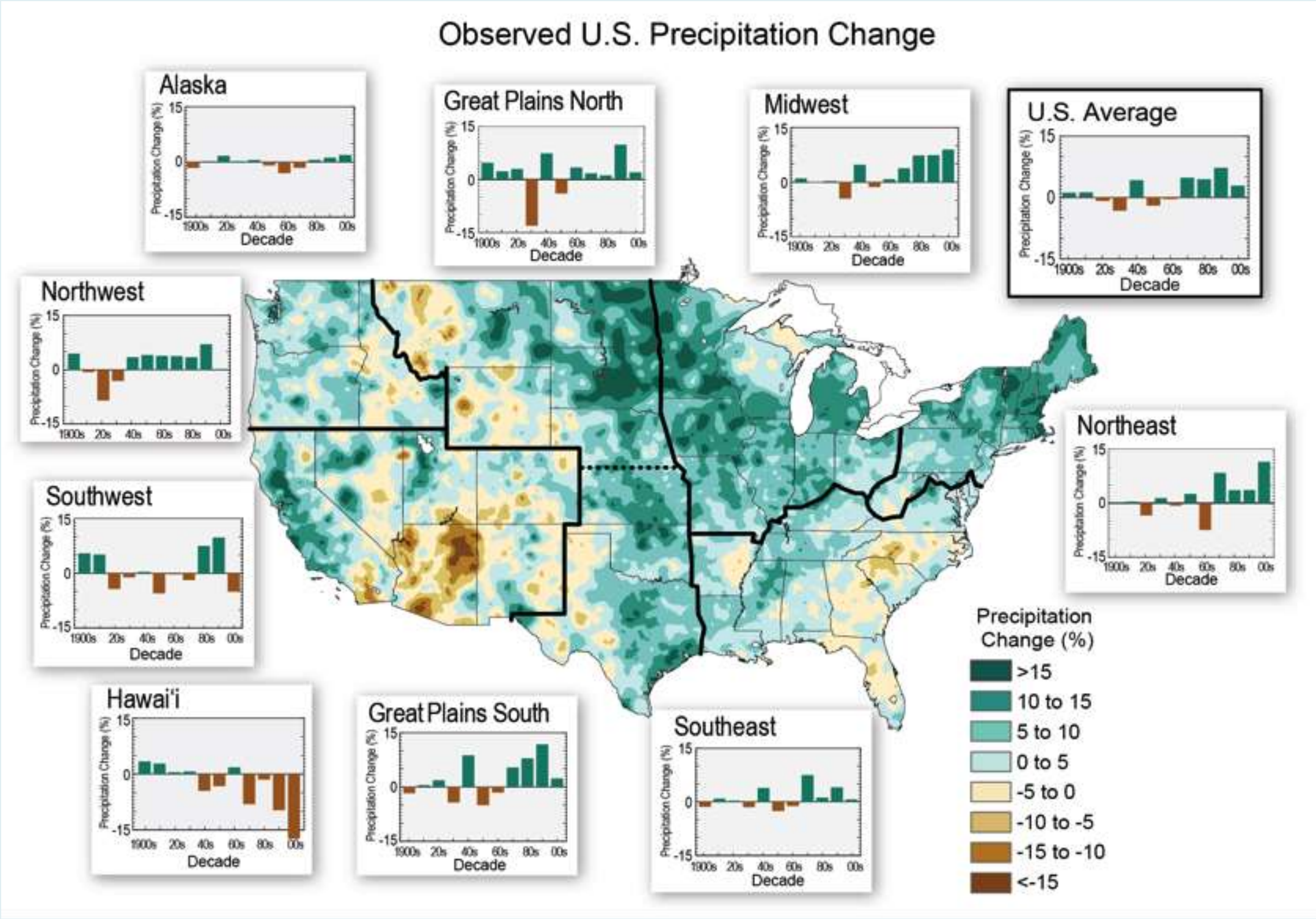
Execution Stage

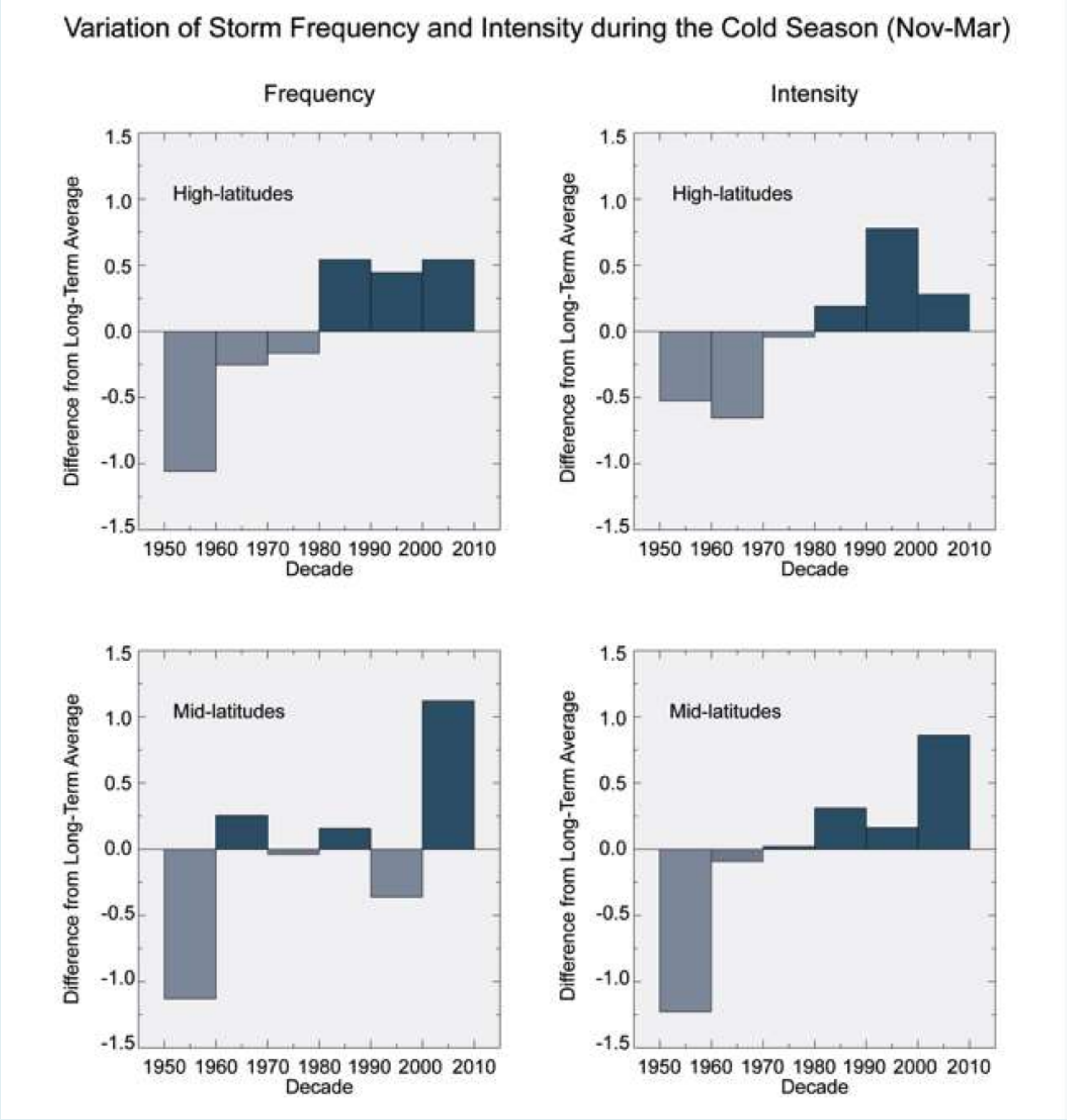
- **After Training and Optimization the ANN Configuration is Executed for the Following Model Combinations:**
  - Observed Gridded Data for Each RCM
  - Modeled Data for Historical Periods
  - Modeled Data for Future Periods
- **10 Total Gridded Periods**
  - 4 Historical, 4 Future, 2 Reanalysis

GCM (Right) RCM (Down)	NCEP (Reanalysis)	CCSM	GFDL	HADCM3
HRM3	X		X	X
MM5I	X	X		X

Randomization Test (90%)

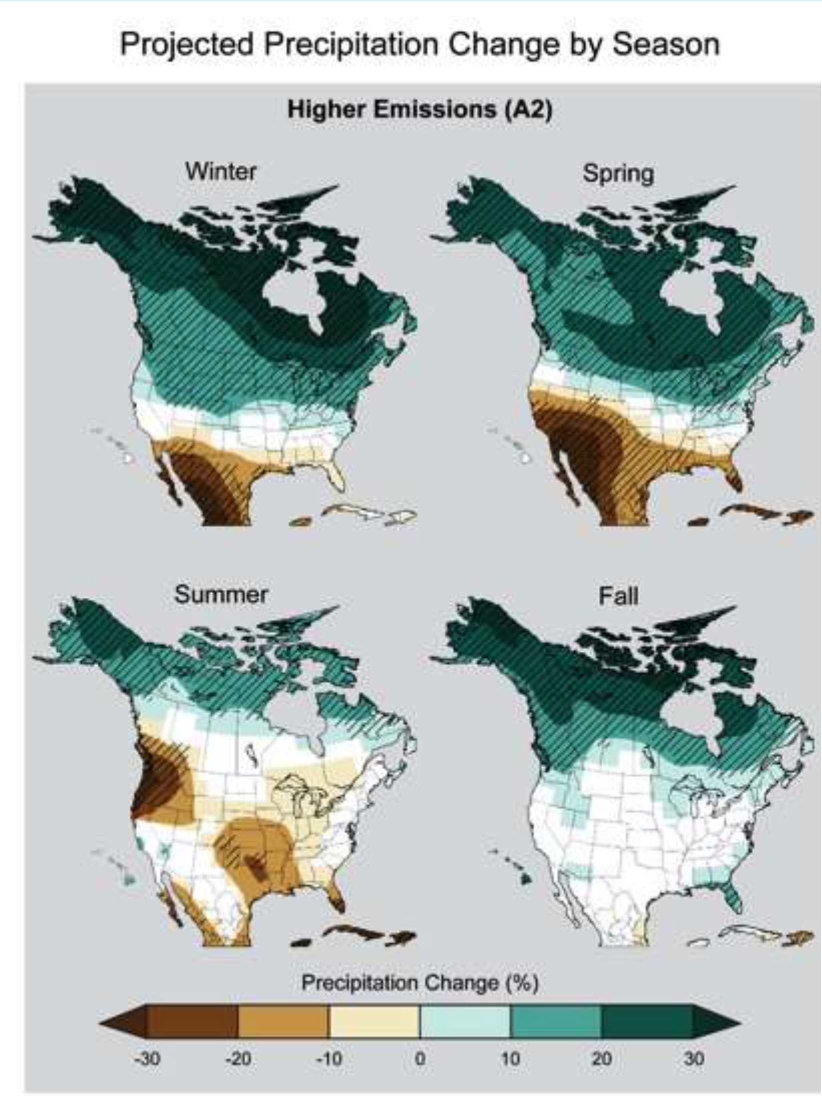


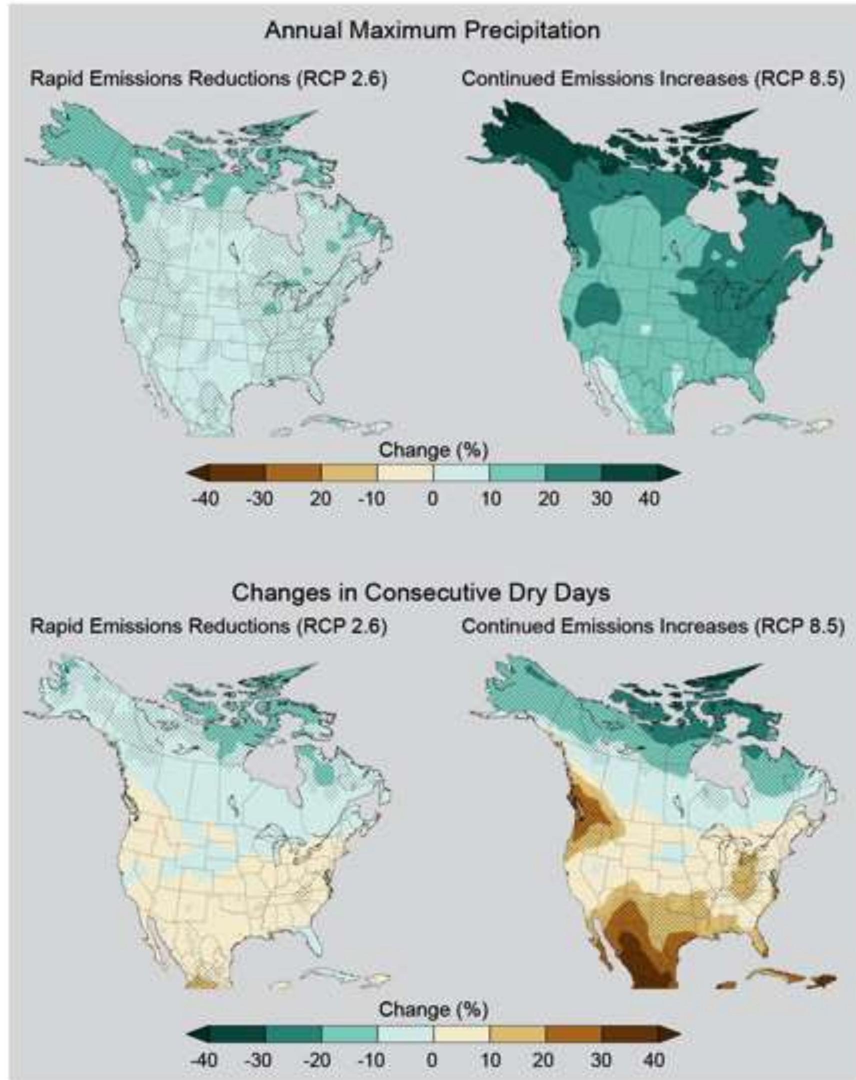

Observed U.S. Precipitation Change
Alaska
Great Plains North
Midwest
U.S. Average
Northwest
Southwest
Northeast
Hawai'i
Great Plains South
Southeast
Precipitation Change (%)
Decade
1900s 20s 40s 60s 80s 00s
15
0
-15
Precipitation Change (%)
>15
10 to 15
5 to 10
0 to 5
-5 to 0
-10 to -5
-15 to -10
<-15

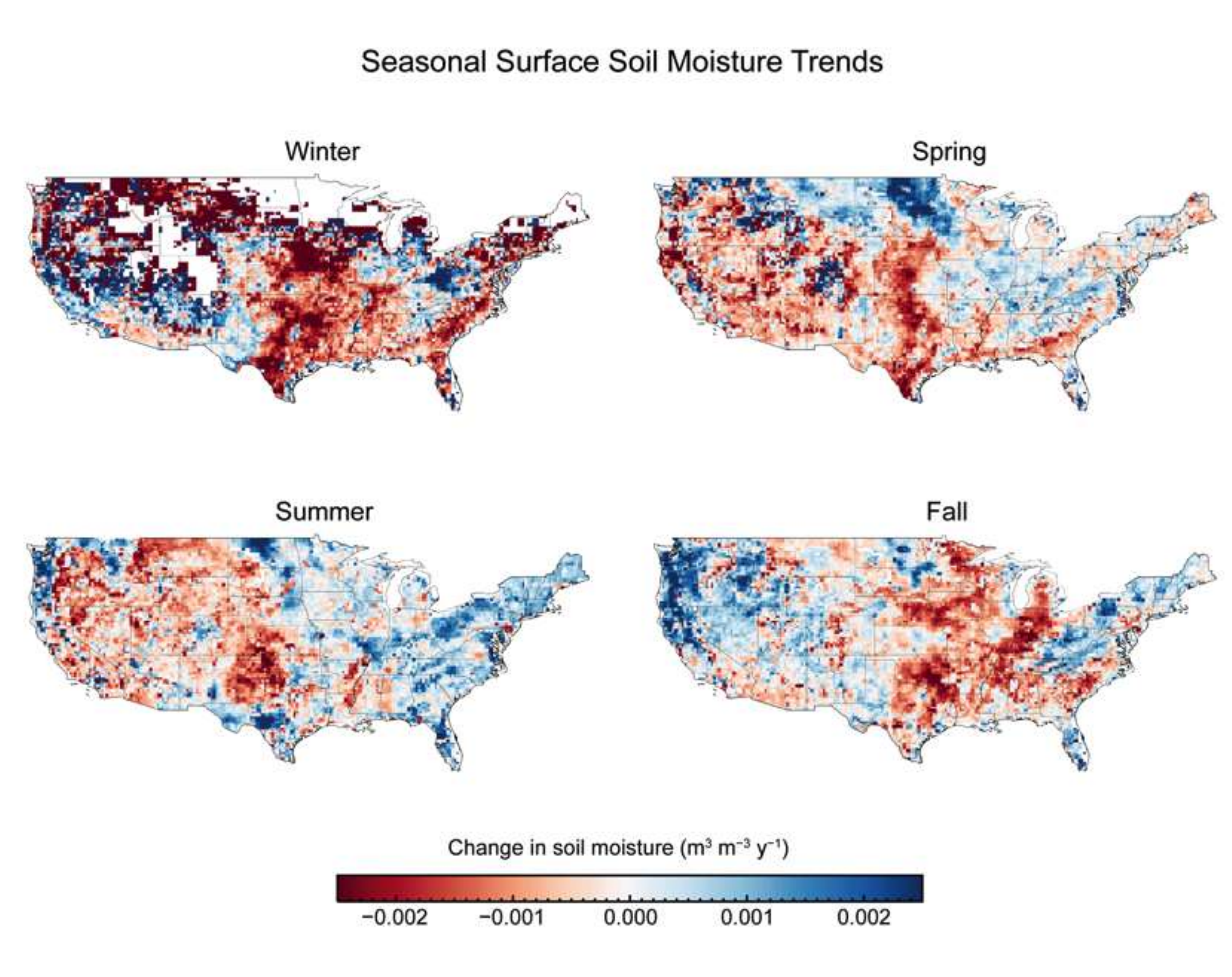


Variation of Storm Frequency and Intensity during the Cold Season (Nov-Mar)
Frequency
Intensity
High-latitudes
Mid-latitudes
Difference from Long-Term Average
Decade
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
1950 1960 1970 1980 1990 2000 2010

AUBURN UNIVERSITY






Seasonal Surface Soil Moisture Trends
Winter
Spring
Summer
Fall
Change in soil moisture ($m^3 m^{-3} y^{-1}$)
−0.002
−0.001
0.000
0.001
0.002