Climatology and Cluster Analysis:
Self-Organizing Maps (SOMs)

Ayzha D. Ward
Faculty Mentor: Dr. Benjamin Lintner
Graduate Mentor: Maxwell Pike
Department of Environmental Science
Background

- Synoptic Climatology
  - The study of atmospheric behavior and characteristics
    - Local or regional climate in specific area
  - Visual Representations
    - Global Climate Models: Coupled Model Intercomparison Project “Phase 5” (CMIP5) ensemble
      - GCMs have biases, inaccuracies that differ from reality

- Question
  - How can we obtain more accurate models?
Objectives

**Broad Objectives**

- Improve our understanding of climate: how and why it works the way that it does
  - Precipitation

- Test efficiency of SOMs
  - Ability to organize large climate data sets

**Specific Objectives**

- To identify the causes of biases in global climate models

- To learn more about SPCZ and ITCZ precipitation patterns
  - Possible relations the two large rainbands may have
Where are they located?

How SOM Analysis Works

Data Set \rightarrow \text{MATLAB} \rightarrow \text{“SOM Toolbox”} \rightarrow \text{SOM Map}
Self-Organizing Maps

- “Train” SOM neurons to match data points
  - How neurons learn data
- Colors have 3 parameter values: (Red, Blue, Green)
  - Red = (6,0,0)
  - Green = (0,0,6)
  - Neuron = (1,1,6)
SOM Analysis Methodology

**TRMM Data Set**
- NASA’s Tropical Rainfall Measurement Mission (TRMM)
- Observational Data
  - Observational benchmark
  - Help identify the biases

**CMCC Data Set**
- CMIP5 ensemble model
- Known to have biases, expect to see such
- Will be compared to TRMM SOM analysis

VS

SOM Map

SOM Map
TRMM SOM Analysis

ITCZ

SPCZ

Precip in mm
CMIP5/CMCC SOM Analysis

Precip SOM, 2 Patterns, 1999-2005 20 - 35 120 270

ITCZ

SPCZ
TRMM SOM Map Node Differences

1

2

3

4
TRMM SOM Map Nodes 1 vs. 2
TRMM SOM Map Nodes 3 vs. 4
Conclusion/Future Work

• Research strongly suggests SOMs to be an efficient data analysis tool
  – Reduces high-res data to lower, visible dimensions
  – Ability to “learn” the data set that is being used
• Results suggest anomalies are not produced by SOM analysis
  – Most likely due to CMIP5 model parameterization
• Future Work
  – Conducting SOM analysis on:
    • Other CMIP5 models
  – Plotting difference between TRMM and CMIP5 SOM nodes
Acknowledgements

• Department of Environmental Science
  – Dr. Benjamin Lintner
  – Mr. Maxwell Pike

• RiSE REU
  – Dr. Evelyn Erenrich
  – Ms. Ana Rodriguez

• DIMACS REU
  – Dr. Eugene Fiorini

• Texas Southern University
  – Dr. Lila Ghemri

• CCICADA
Climatology and Cluster Analysis: Self-Organizing Maps (SOMs)

Thank you for your time!

Questions?