Biome Classification in the United States

Jeffrey Blodgett, Nathan Madsen

ECEn 670
Brigham Young University

December 3, 2013
Overview

1. Introduction
2. Theory
3. Experiment
4. Results
5. Analysis
6. Conclusions
Biomes

- We will create and analyze a classification system for biomes in the United States using scatterometer and radiometer measurements.
- Four biome categories: Temperate Forests, Plains, Deserts, Northwest Forested Mountains

Sensors

- Advanced Scatterometer (ASCAT)
- Advanced Microwave Scanning Radiometer - Earth Observing System (AMSR-E)
Data Types

- Each data type separated: Ascending/Descending, Summer/Winter
- 16-dimensional vector

<table>
<thead>
<tr>
<th>ASCAT</th>
<th>AMSR-E</th>
</tr>
</thead>
<tbody>
<tr>
<td>A image ($\sigma_0$ or backscatter)</td>
<td>$T_B$ H-pol</td>
</tr>
<tr>
<td>B image (incidence angle dependence)</td>
<td>$T_B$ V-pol</td>
</tr>
</tbody>
</table>
Overview

1. Introduction

2. Theory

3. Experiment

4. Results

5. Analysis

6. Conclusions
Model

- Treat data for each biome as a 16-D random vector
- Assume multivariate gaussian distribution

\[ \tilde{X}_{biome} = \tilde{X}_{true} + n_{terrain} + n_{weather} + n_{sensor} \]

\[
f_{\tilde{X}_{biome}} (\tilde{x}) = \frac{1}{\sqrt{(2\pi)^k \det(\Lambda)}} e^{-\frac{(\tilde{x} - \tilde{\mu})^T \Lambda^{-1} (\tilde{x} - \tilde{\mu})}{2}}
\]
Use a maximum likelihood estimator to determine the biome given a measured vector $\vec{x}_{\text{meas}}$ and the pdf $f_{\vec{X}_{\text{biome}}} (\vec{x})$ for each biome

$$\hat{X}_{\text{ML}} (\vec{y}) = \arg\max f_{\vec{X}_{\text{biome}} | \vec{x}_{\text{meas}}} (\vec{y} | \vec{y} = \vec{x}_{\text{meas}})$$
Before applying the estimator, principle component analysis (PCA) is used to reduce the complexity of the problem.
Overview

1. Introduction
2. Theory
3. Experiment
4. Results
5. Analysis
6. Conclusions
Training Sets

- Areas representing each biome selected to estimate each biome’s distribution
Simulation

- Distributions determined for each biome from training sets
- Use these distributions to generate random vectors then classify

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Temp. For.</td>
<td>0.9991</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0008</td>
</tr>
<tr>
<td>Gr. Plains</td>
<td>0.0001</td>
<td>0.9999</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>Desert</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.9952</td>
<td>0.0048</td>
</tr>
<tr>
<td>NW Mount.</td>
<td>0.0022</td>
<td>0.0000</td>
<td>0.0044</td>
<td>0.9934</td>
</tr>
</tbody>
</table>
Iterative Process

- Training areas may not give an accurate distribution for an entire biome
- Increase the sample size by classifying the entire United States
- Recalculate the distribution
- Repeat until it stabilized (under 3% of pixels changed)
Iterative Process

Initial Classification Using Training Sets
Iterative Process

Iteration 1

[Map of the United States showing different biomes with legends: Temperate Forests, Great Plains, Desert, Northwest Forests]
Iterative Process

Iteration 2

- Temperate Forests
- Great Plains
- Desert
- Northwest Forests
Iterative Process

Iteration 3

- Temperate Forests
- Great Plains
- Desert
- Northwest Forests
Iterative Process

Iteration 4

- Temperate Forests
- Great Plains
- Desert
- Northwest Forests
Iterative Process

Iteration 5

![Map of the United States showing different biomes: Temperate Forests, Great Plains, Desert, Northwest Forests]
Iterative Process

Iteration 6

![Map of the United States showing biomes: Temperate Forests, Great Plains, Desert, and Northwest Forests.](image)
Iterative Process

Iteration 7

![Map of the United States with different biomes shaded in various colors. The legend shows different biomes, including Temperate Forests, Great Plains, Desert, and Northwest Forests.]
Iterative Process

Iteration 8

![Map of the United States with biomes classified]

- Blue: Temperate Forests
- Cyan: Great Plains
- Yellow: Desert
- Red: Northwest Forests
Iterative Process

Convergence of Iterations
Overview

1. Introduction
2. Theory
3. Experiment
4. Results
5. Analysis
6. Conclusions
Results

- Gaussian distributions determined using iterated classification
- Vectors for each biome generated and then classified

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Temp. For.</td>
<td>0.9845</td>
<td>0.0028</td>
<td>0.0055</td>
<td>0.0072</td>
</tr>
<tr>
<td>Gr. Plains</td>
<td>0.0087</td>
<td>0.9390</td>
<td>0.0359</td>
<td>0.0164</td>
</tr>
<tr>
<td>Desert</td>
<td>0.0115</td>
<td>0.0298</td>
<td>0.9212</td>
<td>0.0376</td>
</tr>
<tr>
<td>NW Mount.</td>
<td>0.0225</td>
<td>0.0129</td>
<td>0.0343</td>
<td>0.9303</td>
</tr>
</tbody>
</table>
Results

Biome characterization system applied to 2011 data
## Results

<table>
<thead>
<tr>
<th></th>
<th>2010</th>
<th>2011</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tem. For.</td>
<td>Gr. Plains</td>
<td>Desert</td>
<td>NW Moun.</td>
</tr>
<tr>
<td>Temp. For.</td>
<td>0.4193</td>
<td>0.1159</td>
<td>0.0973</td>
<td>0.3675</td>
</tr>
<tr>
<td>Gr. Plains</td>
<td>0.0082</td>
<td>0.8665</td>
<td>0.0622</td>
<td>0.0631</td>
</tr>
<tr>
<td>Desert</td>
<td>0.0124</td>
<td>0.3043</td>
<td>0.5843</td>
<td>0.0990</td>
</tr>
<tr>
<td>NW Mount.</td>
<td>0.0158</td>
<td>0.0515</td>
<td>0.0843</td>
<td>0.8484</td>
</tr>
</tbody>
</table>
Overview

1. Introduction
2. Theory
3. Experiment
4. Results
5. Analysis
6. Conclusions
Preliminary analysis suggests biomes are statistically distinct
When compared with 2011 data not as accurate as in simulation
Unaccounted for interannual variation in distributions
Greater diversity in biome types
Some 'noise' signals may not be Gaussian
Lack of prior distribution of biomes
Overview

1 Introduction

2 Theory

3 Experiment

4 Results

5 Analysis

6 Conclusions
Conclusions

- Gaussian distribution can describe means and linear correlations
- Not all distributions are Gaussian
- The classification did not fall always along biome lines
- May be detecting signals other than simply inter-biome variation