Evaluating factors influencing forest growth across climatic and silvicultural gradients in the northern forests of the United States

Sarah Johnson
M.S. Defense
University of Maine
August 2 2012
Forest Growth & Yield

• Foundation of forest management
  – Many assumptions
  – Confounding factors
  – Multiple metrics

• Complex systems
  – Highly dynamic
  – Limited scope of inference
  – Differing spatial and temporal gradients

• Long-term growth and yield data is invaluable, but relatively rare

Figure 1.1. Number of publications on growth and yield, by publication year, based on a keyword search the CAB Direct Database (www.cabdirect.org, accessed December 21, 2010).
Silvicultural Research

• Long-term and large-scale

• Open-ended questions

• Unique opportunities

• Diverse challenges
MS Context

Can we use historical silvicultural experiments to better understand variability of growth across regional scales and multiple gradients?

How?

Challenges?

Opportunities?

Importance?
Project Significance

- Multiple gradients present & influential in NE
  - First effort to compare net growth of EFRs across multiple gradients
  - 80 years of ind. tree-level net growth to study
  - Increase accessibility and highlight importance of these LT studies
  - Explain local variation in silvicultural outcome based on multiple factors
Thesis Goals

• Increase scientific collaboration within NE

• Use historical data for silvicultural synthesis
  – What is needed?
  – What is available?
  – Does silvicultural outcome vary across NE?
Research Questions

• How can we reuse previously collected data from complex historical studies?

• What are minimum data requirements?

• What can we do in future to better store and utilize these data?

• What types of comparison can we make with these data?
Objectives

• Increase data quality and accessibility with creation of relational database
  – Provide simple guidelines for future standardization of silvicultural data

• Contrast stand-level growth rates across long-term silvicultural studies
  – Showcase the importance of cataloguing long-term forest growth records
Forest Data

• 8 long-term experiments
• 28 unique treatments
  - Control
  - EA
  - UEA
• Multiple data formats
  - Aggregated vs. individual
  - txt, xls, dat, accbd
  - Long vs. wide
Long-term Data

• All plot- and tree-level data previously collected
  – 1927 – 2010

• 3 categories and sources
  – Forest growth data:
    • USFS NRS
  – Soils data:
    • NRCS soil surveys
  – Climatic data:
    • PRISM Climate Group (OSU)
Soil Data

Type

Drainage

WHC

- Very Poorly Drained
- Poorly Drained
- Moderately Well Drained
- Well Drained
- Very Well Drained

- Very Low
- Low
- Moderate
- High
- Very High

Background | Introduction | Chapter 1 | Chapter 2 | Summary
Methodology

Original Records

Unaltered Raw

Standardization

Ancillary Site Data

Precipitation

Temperature

Soil

Standardized Raw

Site-Specific Minimum Diameter Limits

Site-Specific Truncated Plot Summaries

Standardized Minimum Diameter Limits

Standardized Truncated Plot Summaries

Across-Site Comparisons

n=361,775 records

n=245,357 records

n=239,362 records

n=184,144 (30,905 trees)
Relational Database

- **Site**
  - Precipitation
  - Temperature
  - Geological attributes

- **Replicate**
  - Soil attributes
  - Time since study initiation

- **Tree**
  - Tree number
  - Species
  - Diameter
  - Status code
  - Additional notes

- **Plot**
  - Treatment type
  - Time since stand entry
  - Summarized stand attributes
  - Plot size
Initial Trends – Relative Density

Regional RD = 0.20 ± 0.001
(using additive SDI and specific gravity)
Initial Trends – Species Richness

Regional Richness = 4.02 ± 0.05
Initial Trends – Growth Rates

Regional PAI = 0.48 ± 0.25 m²ha⁻¹yr⁻¹
(2.09 ± 1.09 ft²ac⁻¹yr⁻¹)
Argonne  Penobscot
Limitations

- **Volume**
  - Not all site had tree height measurements

- **Inherent species differences**
  - 72 species present

- **Understory dynamics**
  - Diameter truncations
Challenges

• Scientist general aversion to data sharing

• Data accessibility

• Clear, effective communication of complex experiments
Looking Forward

**Increasing ease of collaboration**

- Know what types of data are available at your site
- Identify missing information or unclear records that can be immediately fixed
- Prepare documents that are associated with silvicultural experiment
- Begin metadata compilation
- Design data use policies

**LT Data Opportunities**

- Increase record robustness by compiling descriptive data
- Increase the strength of data storage by begging metadata
- Standardize raw data as part of metadata compilation
- Continue to use database for future comparisons
- Flexible multi-scale analysis with future collaboration
Summary

• Data standardization can facilitate large-scale comparisons of silvicultural studies

• Additional site-level data management needed

• Database is a flexible approach to increase future use of standardized silvicultural data records
Research Questions

• What is influencing NE forest growth?

• Which factors are most important?

• Are there differences across multiple gradients?
  – Climatic gradients?
  – Silvicultural gradients?
  – Other gradients?
Goals

• Use standardized data for large scale comparisons

• Better understand influential factor variability associated multiple gradients

Objectives

- Highlight general trends and variability of stand-level basal area periodic annual increment
- Identify the relative ranking of factors influencing growth
- Explore the influence of specific factors on stand growth response
Data Synthesis

- Site data
  - Soils
    - Depth
    - Drainage
  - Parent Material

- Interval data
  - PAI (ba)
  - Climatic
    - Temperature
    - Precipitation
  - Stand characteristics
    - Composition
    - Size
<table>
<thead>
<tr>
<th>Method</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear models</td>
<td>-Simple approach</td>
<td>-Not appropriate for complex ecological relationships</td>
</tr>
<tr>
<td></td>
<td>-Easy to interpret</td>
<td>-Required assumptions</td>
</tr>
<tr>
<td>Additive Models</td>
<td>-Flexible functional forms</td>
<td>-May be difficult to interpret</td>
</tr>
<tr>
<td></td>
<td>-Good for complex relationships</td>
<td>-Complicated to fit</td>
</tr>
<tr>
<td>Mixed-effect models</td>
<td>-Account for data hierarchy</td>
<td>-Assumptions of data importance and relationships across area</td>
</tr>
<tr>
<td></td>
<td>-Linear or nonlinear</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-Repeated measures</td>
<td></td>
</tr>
<tr>
<td>Meta-analysis</td>
<td>-Multiple large scale comparisons</td>
<td>-Multiple biases</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-Low transparency</td>
</tr>
<tr>
<td>Regression Trees</td>
<td>-No assumptions on data structure required</td>
<td>-Low level of prediction strength</td>
</tr>
<tr>
<td></td>
<td>-Multiple data types</td>
<td>-May be difficult to interpret</td>
</tr>
<tr>
<td>Boosted Regression Trees</td>
<td>- 2 ML techniques to increase prediction strength of RT</td>
<td>-Sensitive to data inputs</td>
</tr>
<tr>
<td></td>
<td>-Iterative model fitting (CV)</td>
<td>-Model structure</td>
</tr>
<tr>
<td></td>
<td>-Automatically model interactions</td>
<td>-May be difficult to interpret standard fit statistics</td>
</tr>
</tbody>
</table>
Boosted Regression Trees

- Regression Trees
  - Multiple data formats
  - Partitions data

- Boosting
  - Several iterations
  - Cross-validated

- Output
  - Influential factors
  - Relative ranking
  - Observed relationship
Model Covariates

- Composition
- Silviculture
- Climate
- Soils
- Diameter
- Diversity
- Density*

*maximum RD based on Woodall et al. 2005
Models

- Good fit for regional model – 3-way interactions
- Variable fits for site models – Small sample sizes

<table>
<thead>
<tr>
<th>Data</th>
<th>TC</th>
<th>LR</th>
<th>Optimal trees</th>
<th>No. Obs.</th>
<th>$R^2$</th>
<th>RMSE</th>
<th>Mean Abs. Bias</th>
<th>Mean % Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>4</td>
<td>0.10</td>
<td>2650</td>
<td>2613</td>
<td>0.98</td>
<td>0.04</td>
<td>0.03</td>
<td>6.6</td>
</tr>
<tr>
<td>All</td>
<td>4</td>
<td>0.05</td>
<td>5200</td>
<td>2613</td>
<td>0.97</td>
<td>0.05</td>
<td>0.04</td>
<td>7.7</td>
</tr>
<tr>
<td>All</td>
<td>3</td>
<td>0.10</td>
<td>3900</td>
<td>2613</td>
<td>0.96</td>
<td>0.05</td>
<td>0.04</td>
<td>8.1</td>
</tr>
<tr>
<td>All</td>
<td>3</td>
<td>0.05</td>
<td>5800</td>
<td>2613</td>
<td>0.95</td>
<td>0.06</td>
<td>0.05</td>
<td>7.6</td>
</tr>
<tr>
<td>AEF</td>
<td>4</td>
<td>0.01</td>
<td>3750</td>
<td>1417</td>
<td>0.79</td>
<td>0.17</td>
<td>0.09</td>
<td>18.75</td>
</tr>
<tr>
<td>BLS</td>
<td>4</td>
<td>0.01</td>
<td>350</td>
<td>80</td>
<td>0.79</td>
<td>0.11</td>
<td>0.08</td>
<td>18.1</td>
</tr>
<tr>
<td>DEF</td>
<td>4</td>
<td>0.01</td>
<td>3650</td>
<td>424</td>
<td>0.97</td>
<td>0.03</td>
<td>0.03</td>
<td>6.4</td>
</tr>
<tr>
<td>FEF</td>
<td>4</td>
<td>0.10</td>
<td>500</td>
<td>107</td>
<td>0.78</td>
<td>0.12</td>
<td>0.09</td>
<td>17.43</td>
</tr>
<tr>
<td>KEF</td>
<td>4</td>
<td>0.001</td>
<td>1450</td>
<td>105</td>
<td>0.65</td>
<td>0.14</td>
<td>0.11</td>
<td>29.03</td>
</tr>
<tr>
<td>PEF</td>
<td>4</td>
<td>0.01</td>
<td>5500</td>
<td>334</td>
<td>0.99</td>
<td>0.02</td>
<td>0.02</td>
<td>3.1</td>
</tr>
<tr>
<td>SEF</td>
<td>4</td>
<td>0.001</td>
<td>3100</td>
<td>45</td>
<td>0.53</td>
<td>0.17</td>
<td>0.13</td>
<td>32.6</td>
</tr>
<tr>
<td>VFEF</td>
<td>4</td>
<td>0.0001</td>
<td>3450</td>
<td>101</td>
<td>0.74</td>
<td>0.15</td>
<td>0.11</td>
<td>19.6</td>
</tr>
</tbody>
</table>

Comprehensive model
- AEF
- BLS
- DEF
- FEF
- KEF
- PEF
- SEF
- VFEF
Influential Factors

• Comprehensive model
  – Density and diameter most influential

• Site models
  – Mixed species RD
  – Climate
    • Variable influential factors
Influential Factors

- Density
- Diameter
- Silviculture
- Soil
- Composition
- Diversity
## Influential Factors

### Regional Model

<table>
<thead>
<tr>
<th>Rank</th>
<th>Relative Influence</th>
<th>Variable Name</th>
<th>AEF</th>
<th>BLS</th>
<th>DEF</th>
<th>FEF</th>
<th>KEF</th>
<th>PEF</th>
<th>SEF</th>
<th>VFSEF</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8.46</td>
<td>WD05_RD*</td>
<td>6.07 (4)</td>
<td>-</td>
<td>9.73 (3)</td>
<td>1.57 (15)</td>
<td>3.14 (10)</td>
<td>16.96 (1)</td>
<td>18.11 (1)</td>
<td>0.50 (39)</td>
</tr>
<tr>
<td>2</td>
<td>7.84</td>
<td>allba1</td>
<td>6.4  (2)</td>
<td>0.74 (29)</td>
<td>17.74 (1)</td>
<td>4.41 (5)</td>
<td>1.53 (23)</td>
<td>12.17 (2)</td>
<td>0.10 (36)</td>
<td>6.33 (5)</td>
</tr>
<tr>
<td>3</td>
<td>7.27</td>
<td>addSDI</td>
<td>6.53 (1)</td>
<td>0.75 (28)</td>
<td>12.77 (2)</td>
<td>1.01 (24)</td>
<td>2.19 (14)</td>
<td>10.78 (3)</td>
<td>6.98 (2)</td>
<td>2.72 (10)</td>
</tr>
<tr>
<td>4</td>
<td>4.31</td>
<td>m_d</td>
<td>4.92 (7)</td>
<td>3.68 (8)</td>
<td>2.77 (12)</td>
<td>1.62 (14)</td>
<td>2.92 (11)</td>
<td>1.75 (17)</td>
<td>0.30 (28)</td>
<td>0.51 (38)</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>sd_d</td>
<td>5.86 (5)</td>
<td>0.91 (22)</td>
<td>2.92 (8)</td>
<td>3.21 (7)</td>
<td>3.22 (9)</td>
<td>2.93 (5)</td>
<td>2.69 (6)</td>
<td>13.44 (2)</td>
</tr>
</tbody>
</table>

### Site Models
Climate x Stand x Growth

Three-way interactions across multiple gradients present
Stand x Stand x Growth

TR=0 (n=1996)  TR=1 (n=586)  TR=2 (n=31)

Not all factors exhibit complex interactions at regional level
Future Implications?

- Climate predictions
  - More precipitation in NE
  - Longer growing seasons
  - Changes in productivity

1Rustad et al. 2012 Changing Climate, Changing Forests: The impact of climate change on forest of the northeastern United States and Eastern Canada
Climate Influencing Growth

• Precipitation
  – DEF, FEF

• Temperature
  – BLS, VFSEF

• Both
  – KEF, SEF

• Neither*
  – Region
  – AEF, PEF

• Importance of short-term weather events?

* In ten-most influential factors

Huntington et al. 2009. Climate and hydrological changes in the northeastern United States: recent trends and implications for forested and aquatic ecosystems. CJRF
Limitations

• Nonparametric results are dependent on the size, quality, and scope

• Measures of time used in the analysis may not be ecologically important
  – Unknown stand histories

• Confounding of silviculture and stand structural attributes
  – Direct vs. indirect results of silviculture
Influential Factors Summary

• Local and regional trends vary
  – Climatic factors
  – Interactions

• Multiple gradients influential in NE

• Additional study is required
Synthesis Conclusions

• BRT are one possible framework for large scale synthesis
  – Provide influential factor tabulations
  – Dependent on data inputs

• Future additions possible
  – Treatment comparisons
  – Temporal comparisons
Overall Conclusions

Current needs for data management and record archival are high

Large scale conclusions can be drawn from raw data using independent analyses

LT data can be flexible tool for future analyses
Future Efforts

• Additional sites
• Additional data
• Additional responses
  – DWD
  – Tree recruitment
  – Mortality
  – Volume accretion
  – Carbon storage
Acknowledgements

Funding provided by NRS internal funding and additional USFS funding sources