SimilarityExplorer: A Visual Inter-Comparison Tool for Multifaceted Climate Data

Jorge Poco¹, Aritra Dasgupta¹, Yaxing Wei², William Hargrove³, Robert Cook², Christopher Schwalm⁴, Enrico Bertini¹ and Claudio Silva¹

¹. New York University, ². Oak Ridge National Laboratory, ³. United States Forest Service, ⁴. Northern Arizona University
Model Simulation

ecosystem processes  carbon exchange

Input/Drivers  Structure  Output

e.g., temperature, pressure, humidity, etc.

e.g., vegetation type, moisture influence

eg., productivity, soil characteristics, etc.

Holy grail: Analyze similarity among model input, model structure and model output
Model Simulation
ecosystem processes carbon exchange

Explore spatial similarity

Explore temporal similarity

*Images from Huntzinger et al., '11
Challenges

- Multiple models, 100 years of data, multiple variables, multiple regions.
- State-of-the-art: Static scripts
- Need interactive visualization

"Without this tool scientists would literally print hundreds of plots and pin them on the wall, this tool solves this problem".
Inter-comparison of multiple facets (space, time, variables, model similarity)

Where: global, regional

When: annual, seasonal, and monthly

What: multiple variables
Design Study Process

• Collaboration with climate scientists from ORNL, NAU and USFS

• Year-long association: workshops, phone interviews

• Iterative refinement of prototypes through discussions
Contributions

• Domain Characterization for Model Inter-Comparison

• Characterization of the Visualization Design Space

• SimilarityExplorer tool: Output of the Design Study
Related Work

Inspired by previous work on integration of space and time
Peaquet’s triad framework (’94), Andrienko’s visual analytics model (2010)
Related Work

Multifaceted scientific data analysis

Conceptualized by Kehrer’11, 13. Existing work deals with at most two facets

[Kehrer13]

[Kehrer11]
Related Work

Climate data visualization tools

Tools like UVCDAT do not allow multi-model analysis

[WBD*13]
Domain questions

Q₁: How much similar are models, with respect to any variable?

Q₂: How does multi-model similarity vary across space and time?

Q₃: Does global similarity agree with sub-region similarity?

Q₄: How can we associate similarity with data distribution?
Similarity Abstraction

Inter-comparison among model simulations

Compress time

Spatial Correlation between $M_1$ and $M_2$ for output variable $V_1$

Temporal Correlation between $M_1$ and $M_2$ for the time series of output variable $V_1$

Compress space
Design Space

- Multifaceted Data
- Domain Questions
- Visualization Tasks
- Views and Interaction

Questions:

- Q1
- Q2
- Q3
- Q4
Classification Scheme

- **Domain Questions**
  - Identify
  - Compare
  - Associate
  - Distribution

- **Visualization Tasks**
  - Space
  - Time
  - Variables
  - Similarity

- **Facets**
  - Views
  - Comparison
  - Method

- **Visualization Design**
## Classification Scheme

<table>
<thead>
<tr>
<th>Questions</th>
<th>Tasks</th>
<th>Facets</th>
<th>Visualization Design</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Space</td>
<td>Time</td>
</tr>
<tr>
<td><strong>Q1</strong></td>
<td>identify(p)</td>
<td>g/r</td>
<td>a/s/m</td>
</tr>
<tr>
<td></td>
<td>identify(t)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>identify(p,t)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Q2</strong></td>
<td>compare(p,v)</td>
<td>g,r</td>
<td>a,s,m</td>
</tr>
<tr>
<td></td>
<td>compare(t,v)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Q3</strong></td>
<td>associate(p)</td>
<td>r</td>
<td>s,m</td>
</tr>
<tr>
<td></td>
<td>associate(t)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Q4</strong></td>
<td>distribution(p,v)</td>
<td>r</td>
<td>s,m</td>
</tr>
<tr>
<td></td>
<td>distribution(t,v)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p*: Space  
*t*: Time  
*v*: Variables  
*g*: global  
*r*: regional  
*a*: annual  
*s*: seasonal  
*m*: monthly  
*smlt*: small multiples
Similarity Explorer

Filters  Pairwise Similarity  Multiway Similarity

Detailed Exploration: Regions

NYU  Polytechnic School of Engineering
Pairwise Similarity

Spatial Correlation

Temporal Correlation
Multiway Similarity
Interaction with the Matrix

Compare regional and seasonal similarity
Data View
Compare seasonal similarity with data distribution
• Demo
Feedback

Efficiency
“Without this tool scientists would literally print hundreds of plots and pin them on the wall, this tool solves this problem”.

Flexibility
“.the free-style nature of the exploration lends well to shift from one variable to another and support root-cause analysis”.

Prognostic value
“.this would allow them to develop hypotheses on performing additional experiments”
Conclusion

• SimilarityExplorer: A tool that support multi-model comparison for climate data

• Tool currently deployed, used, and developed further

• Build upon the multifaceted analysis paradigm and apply to other domains
Acknowledgements

- DataONE project sponsored by the NSF Grant number OCI-0830944, NSF CNS-1229185, NASA ROSES 10-BIOCLIM10-0067, and DOE Office of Science Biological and Environmental Research (BER).

- Data acquired through the MASTDC (NASA Grant NNH10AN68I) and MsTMIP (NASAGrant NNH10AN68I) projects funded by NASA’s Terrestrial Ecology Program.

- Collaborators: Members of the Exploration, Visualization, and Analysis Working Group under DataONE.
SimilarityExplorer: A Visual Inter-Comparison Tool for Multifaceted Climate Data

Thank You!

jpocom@nyu.edu
adasgupt@nyu.edu