UV-CDAT: Exploring and Analyzing MsTMIP dataset

Jorge Poco
Summer Intern at ORNL
Outline

• Data explanation
• UV-CDAT overview
• Exploration (UV-CDAT)
  o Basic plots
  o Advance plots
  o New interactive features
• Analysis (Vistrails)
  o Basic Plots (dendrogram, taylor diagram, etc.)
  o Real case: Survey data and model output
    • Multidimensional projections
    • Parallel coordinates
    • Linked views
MsTMIP Data

available Fall 2012

• Model Output Data
  o Variables: carbon fluxes (GPP, NEE, ...), carbon pools (TotLivBiom, ...), energy fluxes (LW_albedo, ...), physical (Evap, ...), and other (LAI, ...).
  o Global 0.5° and North American 0.25°
  o Monthly, 1901-2010 and 3-hourly, 1980 - 2010
  o CF-compatible NetCDF
  o 27 model teams, 10 simulations

• Benchmark Data

• Web site:  http://nacp.ornl.gov/MsTMIP.shtml
NACP Regional Interim Synthesis Data

• Model Output Data
  o Variables: GPP, NPP, NEE, Rh, Ra, …
  o Global 1°
  o Monthly, in 200X
  o CF-compatible NetCDF
  o 21 Terrestrial Biospheric Models

• Observation Data
  o MODIS GPP/NPP, MODIS Phenology (LAI/EVI/NDVI/fPAR), Forest Biomass, …
UV-CDAT: overview

• Ultrascale Visualization: Climate Data Analysis Tools
• Easy to use
• Multiple tools
• Provenance support
Exploration: Basic Plots

- Many basic plots (boxfill, isofill, isoline, meshfill, etc.)
- Many projections (miller, polar, robinson, etc)
Exploration: Eco-regions extraction
Exploration: Workflows

- A pipeline is created automatically for each plot.
- A history of all the changes are saved in Vistrails.
- We can reuse the pipelines to create more advances plots.
Exploration: Advance Plots

- 3D plots (DV3D, Paraview, Visit)
- Interactivity
- Camera synchronization
- Interactive time series lenses
Exploration: Advance Plots

Overlay
• Slicer
• Volume Rendering
Exploration: Time Series Lenses

- Interactive time series lens.
- Compare multiple time series.
- Export and save
Analysis: Monthly long term mean

Figure 5. Evaluate long-term trend and seasonal cycle of model Gross Primary Productivity (GPP) in Boreal region
Analysis: Taylor Diagrams
Analysis: Survey data

<table>
<thead>
<tr>
<th>Model Name</th>
<th>CLM</th>
<th>DLEM</th>
<th>Ecoys</th>
<th>IRC/DayCent5</th>
<th>ISAM</th>
<th>JULES</th>
<th>LPJ-wal</th>
<th>MC1</th>
<th>ORCHIDEE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reflectance/Transmittance/Absorptance computed by model (yes = 1, no = 0)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>RTS = 3-D (yes = 1; no = 0)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>RTS = 2-stream (yes = 1; no = 0)</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>RTS = Beer’s law (yes = 1; no = 0)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>RTS = Albedo (yes = 1; no = 0)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Model partitions net radiation into latent &amp; sensible heat (yes = 1; no = 0)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Model simulates ground heat flux (yes = 1; no = 0)</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Canopy stomatal conductance - shaded leaves (yes = 1; no = 0)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Canopy stomatal conductance - sun leaves (yes = 1; no = 0)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Canopy stomatal conductance - whole canopy (yes = 1; no = 0)</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Stomatal conductance scheme = Jarvis-type (yes = 1; no = 0)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Stomatal conductance scheme = Ball Berry (yes = 1; no = 0)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Stomatal conductance connected to photosyn (yes = 1; no = 0)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Linked Views

- Dendrograms
- Parallel Coordinates
Analysis: Model Output

Extract monthly long term mean

High Dimensional Vectors

Dimensionality Reduction (Isomap, PCA, etc.)
Survey Data – Model Output

Survey Data

Model Output

Linked Views
Future work:

• Web version using crowdlabs

http://www.crowdlabs.org/vistrails/medleys/details/13/
Future work

• Use parallel implementation of some components (e.g. standard deviation, correlation, mean, etc.)
• Increase complexity in the configuration file to support more complex pipelines.
• Apply the workflows implemented to the complete MsTMIP data collection, which will be available in Fall 2012.