UV-CDAT: Exploring and Analyzing MsTMIP dataset

Jorge Poco
Summer Intern at ORNL

Outline

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

MsTMIP Data

available Fall 2012

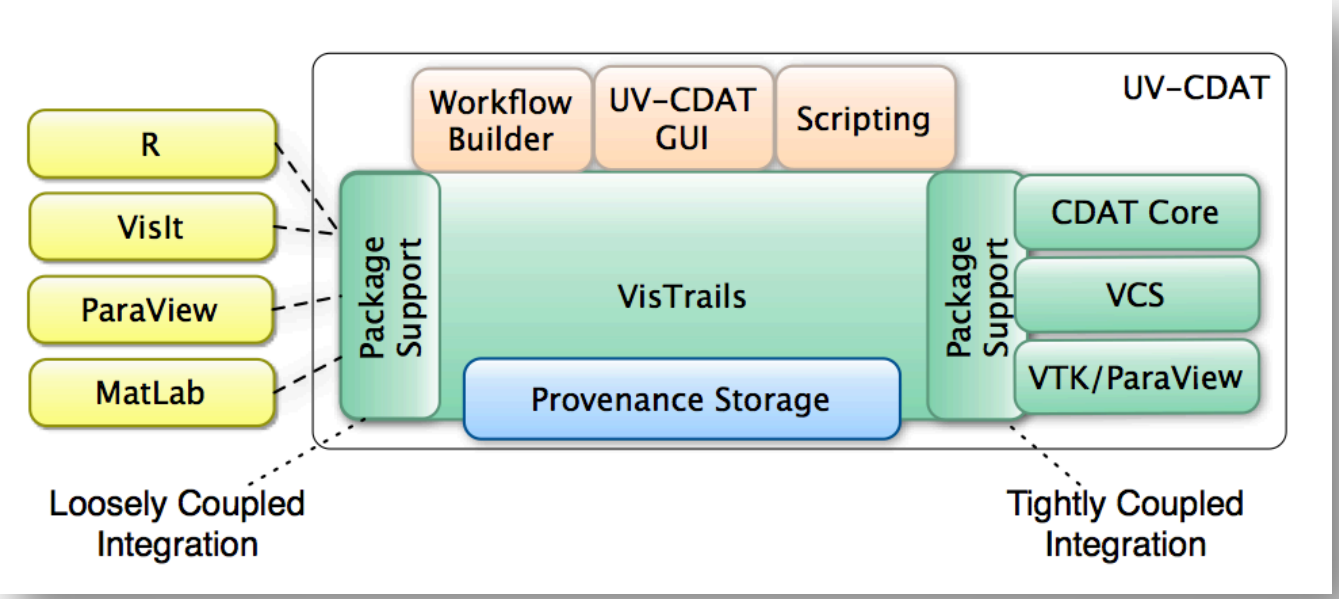
- Model Output Data
  - Variables: carbon fluxes (GPP, NEE, …), carbon pools (TotLivBiom, …), energy fluxes (LW_albedo, …), physical (Evap, …), and other (LAI, …).
  - Global 0.5° and North American 0.25°
  - Monthly, 1901-2010 and 3-hourly, 1980 - 2010
  - CF-compatible NetCDF
  - 27 model teams, 10 simulations
- Benchmark Data
  - Variables: GPP (1982-2008), NEE (1982-2008), Evapotranspiration (1986-1995), …
- Web site: http://nacp.ornl.gov/MsTMIP.shtml

NACP Regional Interim Synthesis Data

- Model Output Data
  - Variables: GPP, NPP, NEE, Rh, Ra, …
  - Global 1°
  - Monthly, in 200X
  - CF-compatible NetCDF
  - 21 Terrestrial Biospheric Models
- Observation Data
  - MODIS GPP/NPP, MODIS Phenology (LAI/EVI/NDVI/fPAR), Forest Biomass, …

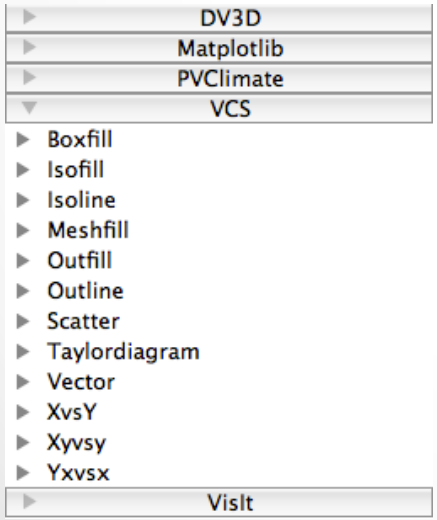
UV-CDAT: overview

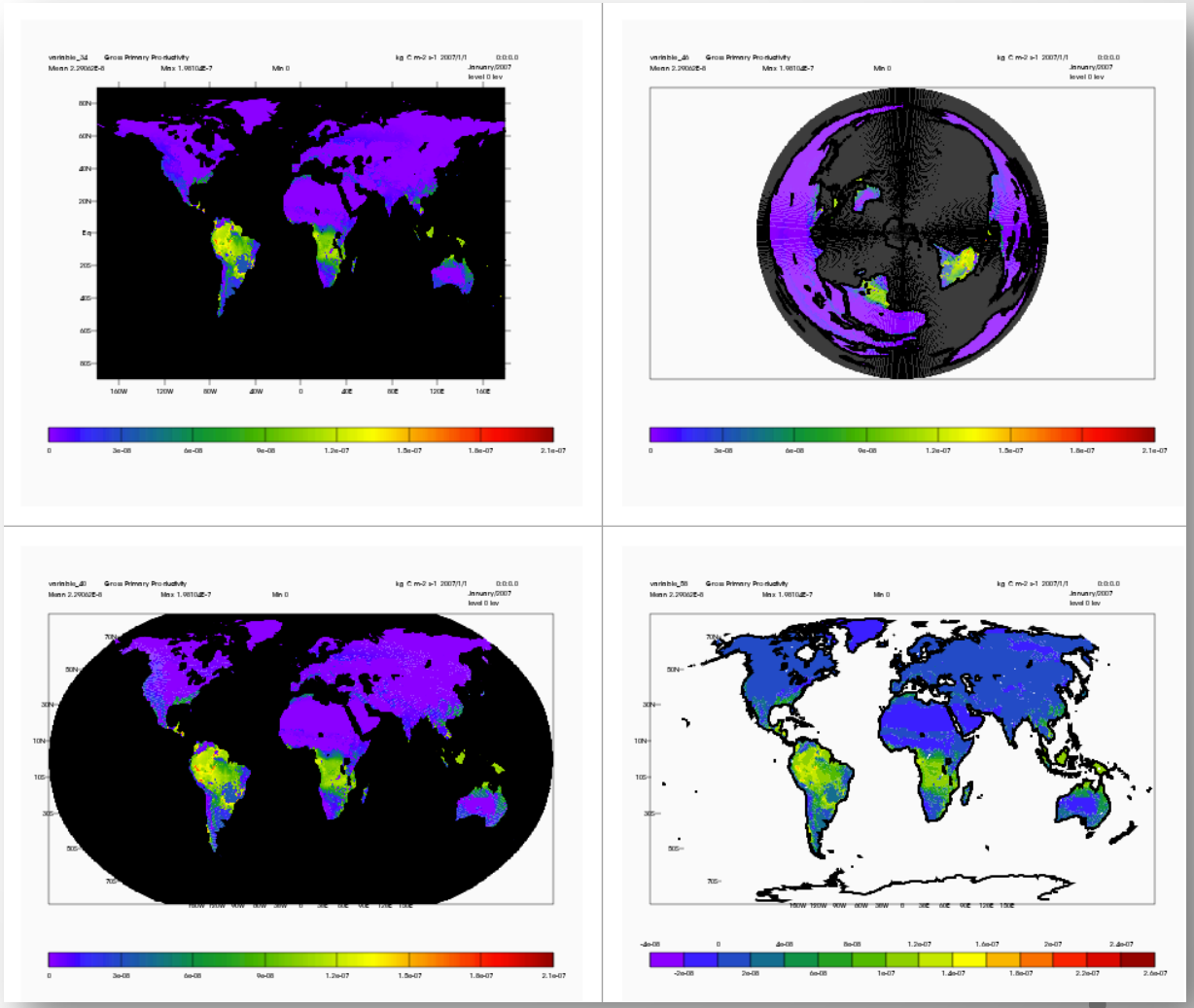
- Ultrascale Visualization: Climate Data Analysis Tools
- Easy to use
- Multiple tools
- Provenance support



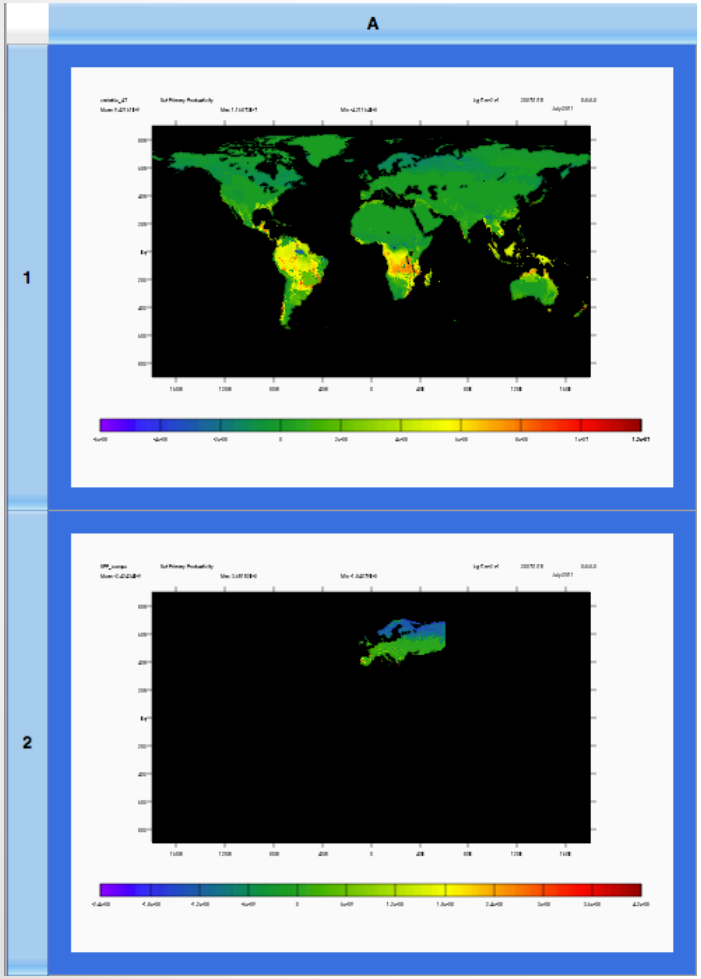
Exploration: Basic Plots

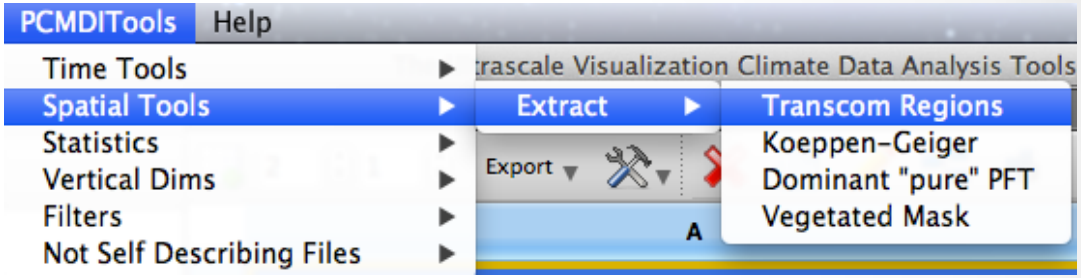
- Many basic plots (boxfill, isofill, isoline, meshfill, etc.)
- Many projections (miller, polar, robinson, etc)

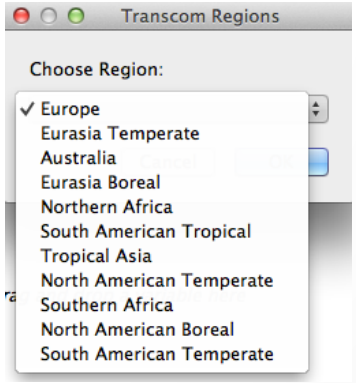




Exploration: Eco-regions extraction



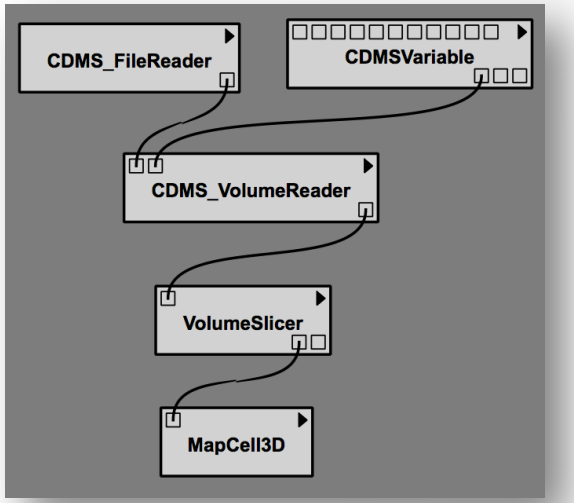




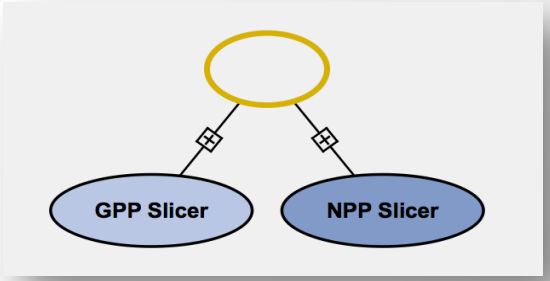
Exploration: Workflows

- 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.

Workflow

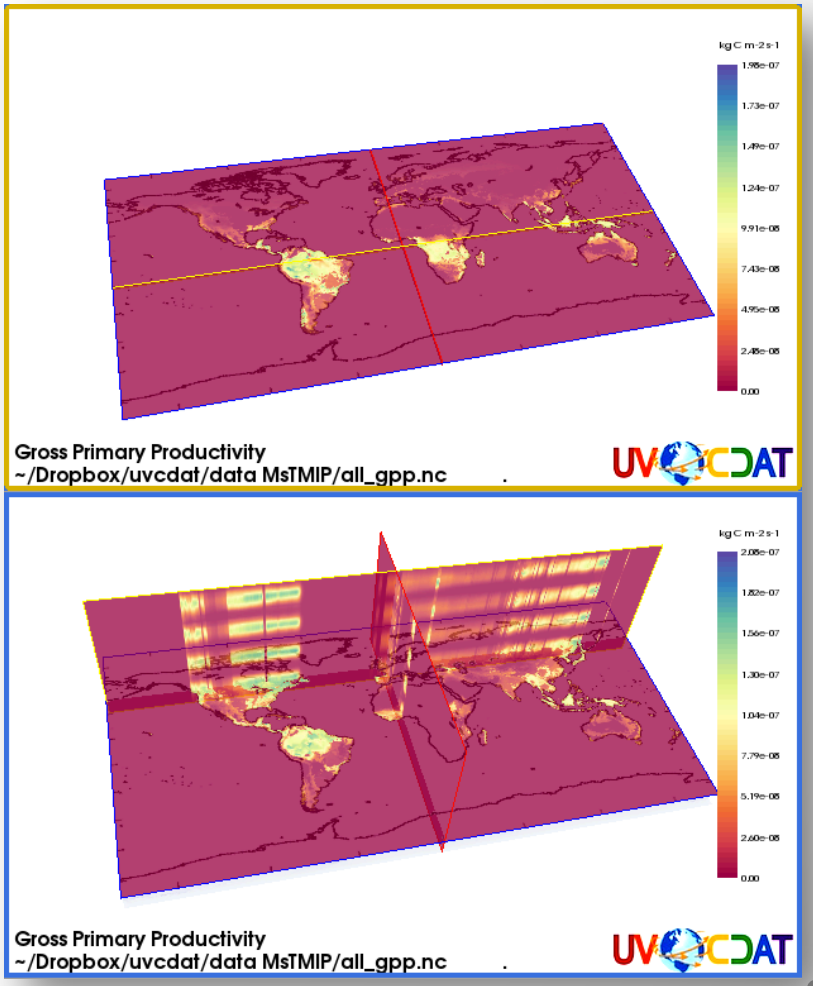


Version Tree



Exploration: Advance Plots

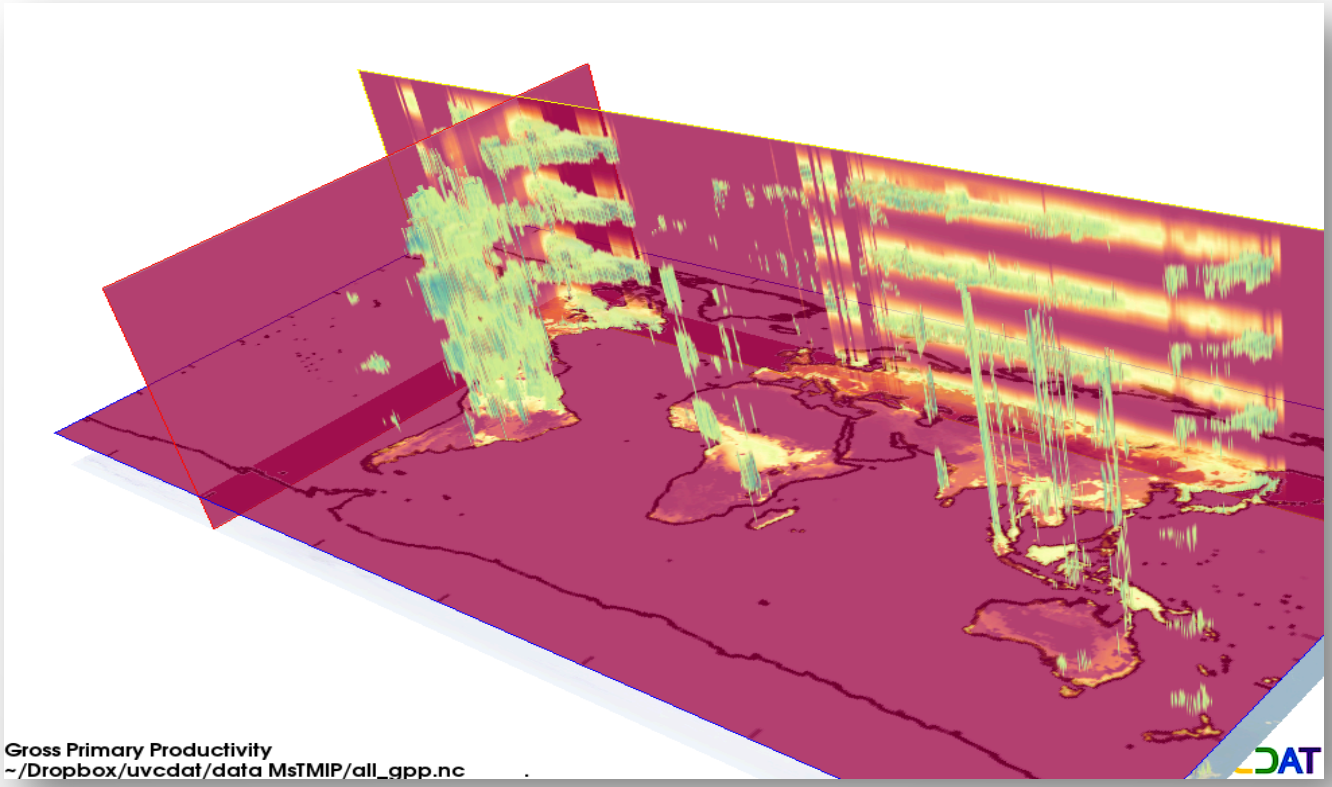
- 3D plots (DV3D, Paraview, Visit)
- Interactivity
- Camera synchronization
- Interactive time series lenses



Exploration: Advance Plots

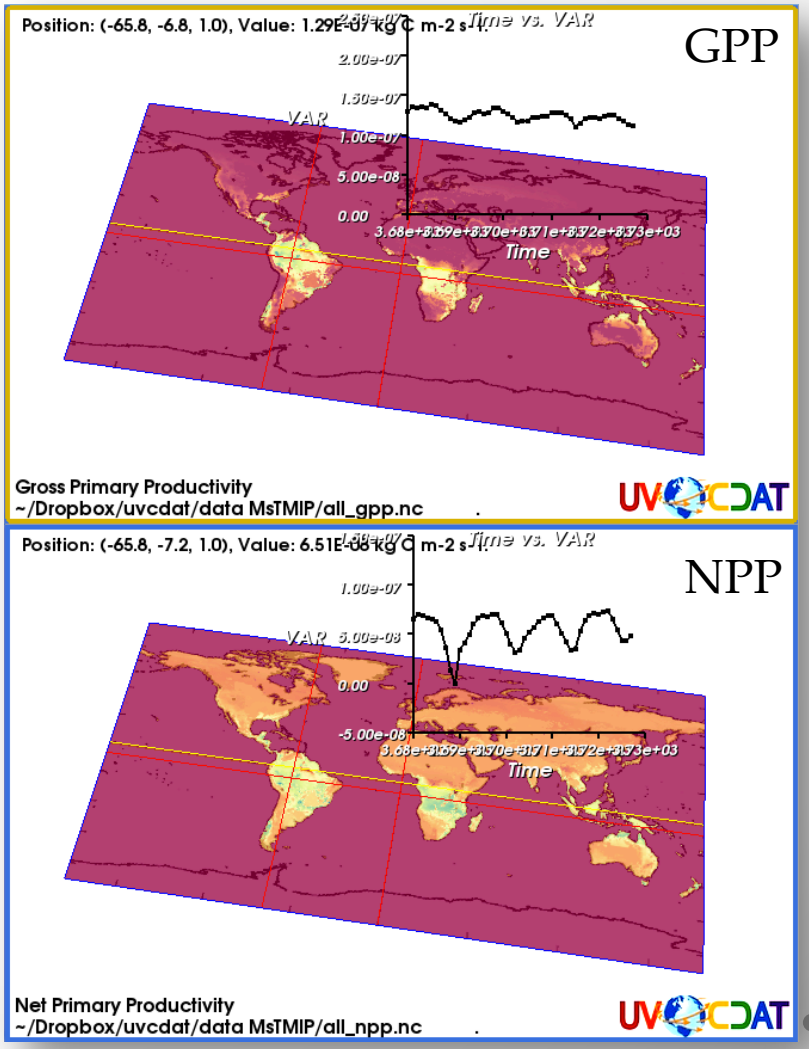
Overlay

- Slicer
- Volume Rendering



Exploration: Time Series Lenses

- Interactive time series lens.
- Compare multiple time series.
- Export and save



Analysis: Monthly long term mean

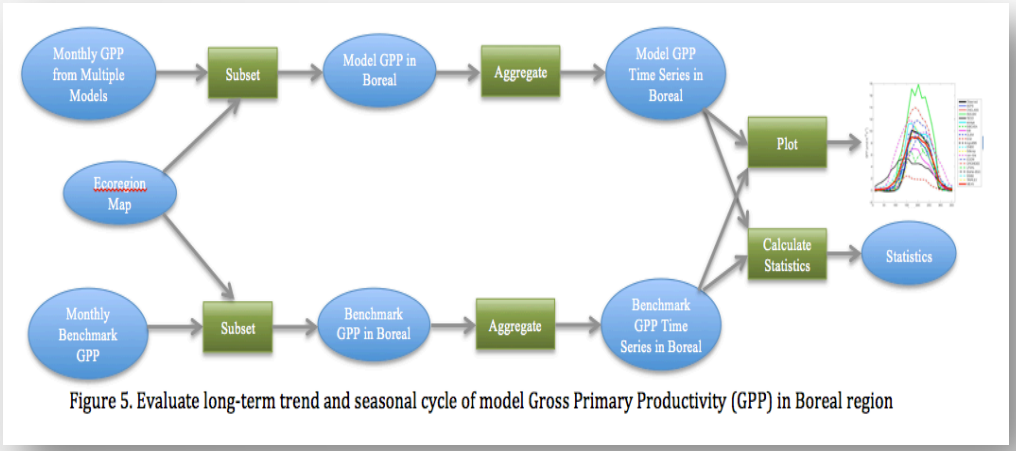
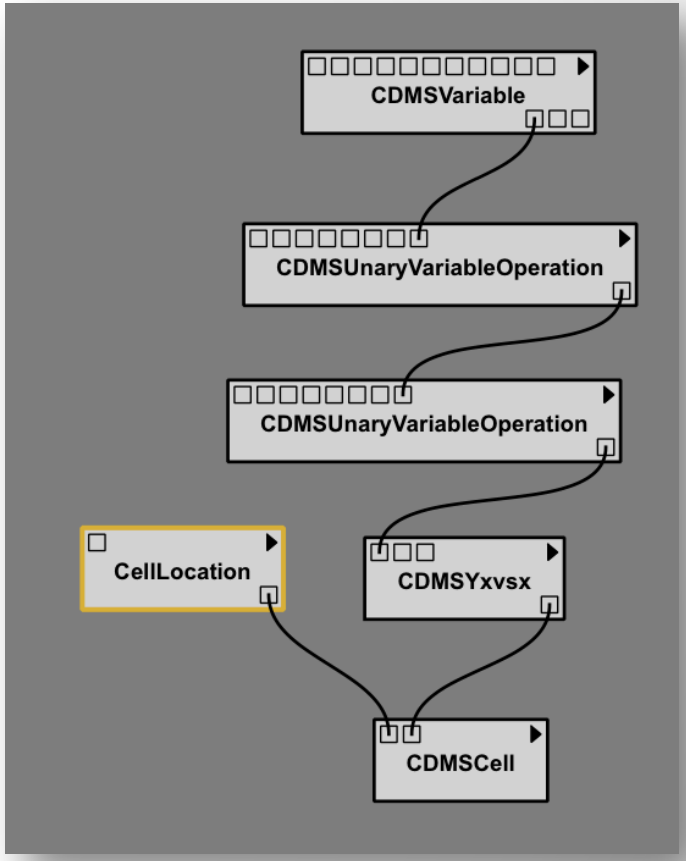
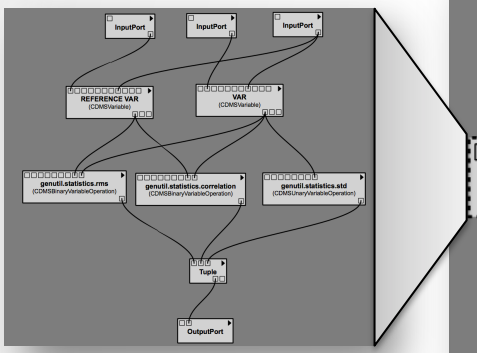
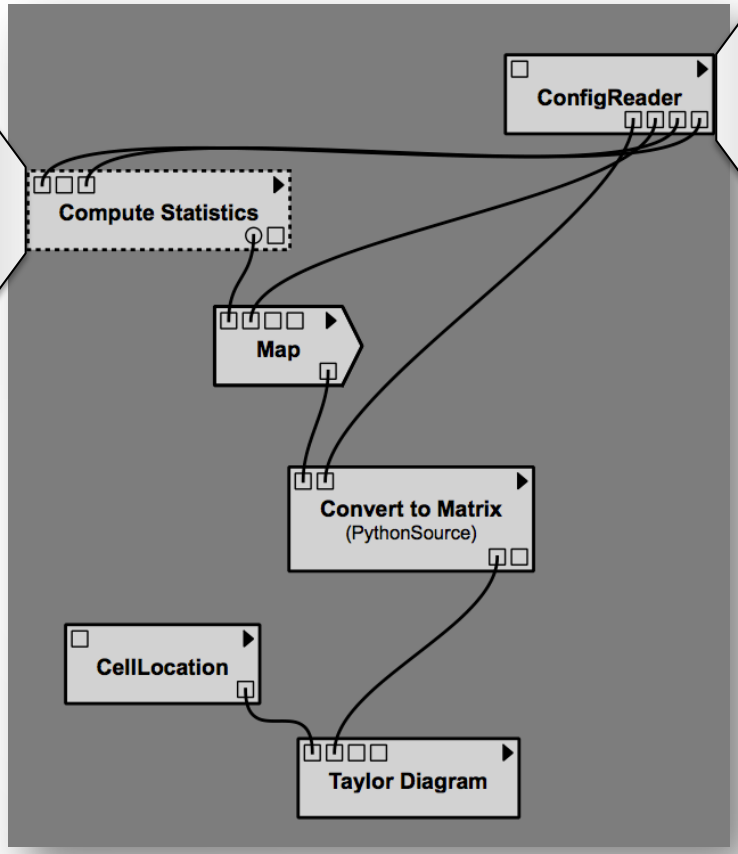


Figure 5. Evaluate long-term trend and seasonal cycle of model Gross Primary Productivity (GPP) in Boreal region

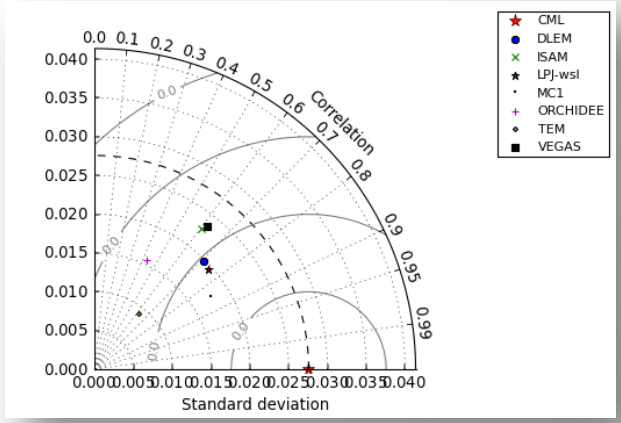


Analysis: Taylor Diagrams





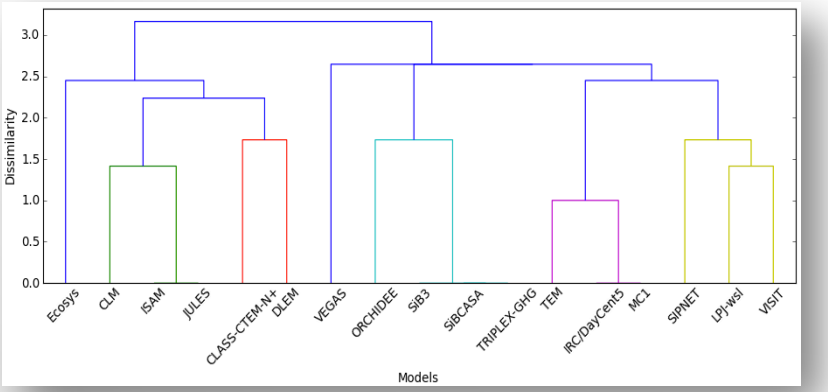

```
[basic]
variable_name: NPP
ref_filename: /Path/CLM-CN_NPP_onedeg_i01.56cn.nc

[file_names]
CML:      /Path/CLM-CN_NPP_onedeg_i01.56cn.nc
DLEM:     /Path/DLEM_NPP_onedeg_v1.0.nc
ISAM:     /Path/ISAM_NPP_onedeg_v1.0.nc
LPJ-wsl:  /Path/LPJmL_NPP_onedeg_v1.0.nc
MC1:      /Path/MC1_NPP_onedeg_v1.0.nc
ORCHIDEE:/Path/ORCHIDEE_NPP_onedeg_v2.1.nc
TEM:      /Path/TEM6_NACP_NPP_onedeg_v1.0.nc
VEGAS:    /Path/VEGAS2_NPP_onedeg_v1.0.nc
```



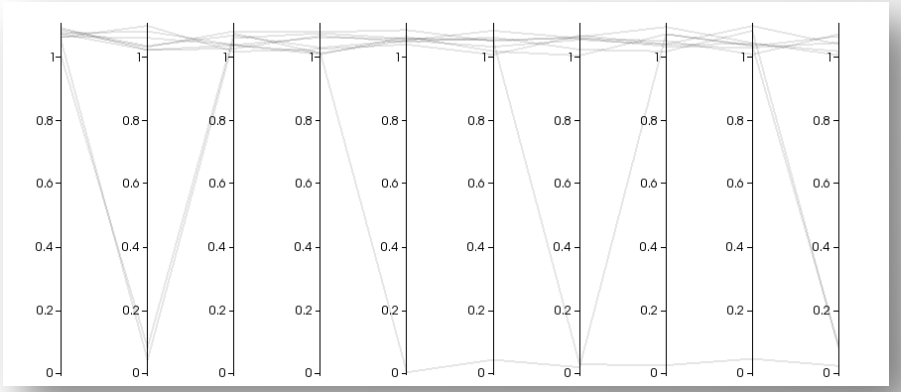
Analysis: Survey data

	CLASS-CTEM-N+	CLM	DLEM	Ecosys	IRC/DayCent 5	ISAM	JULES	LPJ-wsl	MC1	ORCHIDEE
Reflectance/Transmittance/Absorptance computed by model (yes =1, no = 0)	1	1	1	1	0	1	1	0	0	1
RTS = 3-D (yes = 1; no = 0)	0	0	0	1	0	0	0	0	0	0
RTS = 2-stream (yes = 1; no = 0)	0	1	0	0	0	1	1	0	0	0
RTS = Beer's law (yes = 1; no = 0)	1	1	1	0	0	0	0	1	0	0
RTS = Albedo (yes = 1; no = 0)	1	1	1	0	0	0	0	0	0	1
model partitions net radiation into latent & sensible heat (yes = 1; no = 0)	1	1	0	1	0	1	1	0	0	1
Model simlates ground heat flux (yes = 1; no = 0)	1	1	0	1	0	1	1	0	0	1
Canonpy stomatal conductance - shaded leaves (yes = 1; no = 0)	1	1	1	1	0	1	1	0	0	0
Canonpy stomatal conductance - sun leaves (yes = 1; no = 0)	1	1	1	1	0	1	1	0	0	0
Canonpy stomatal conductance - whole canopy (yes = 1; no = 0)	1	0	0	0	0	0	0	1	0	0
Stomatal conductance scheme = Jarvis-type (yes = 1; no = 0)	0	0	0	0	0	0	0	0	0	0
Stomatal conductance scheme = Ball Berry (yes = 1; no = 0)	1	1	1	0	0	1	1	0	0	1
Stomatal conductance connected to photosyn (yes = 1; no = 0)	1	1	1	1	0	1	1	1	0	1

Linked Views



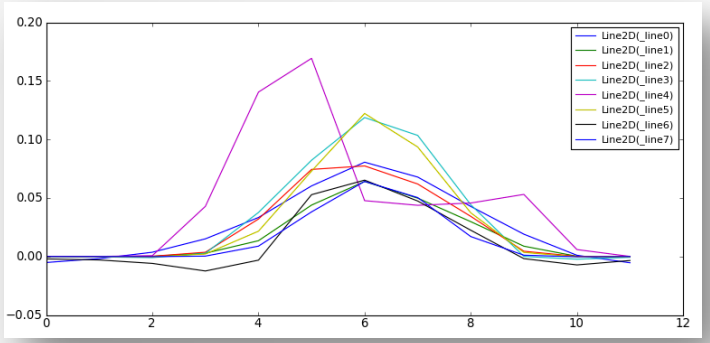
Dendrograms



Parallel Coordinates

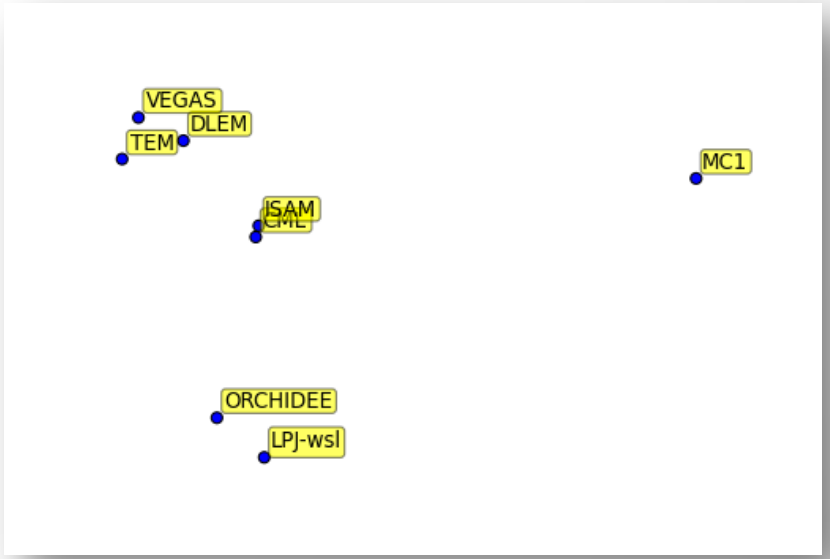
Analysis: Model Output

Extract monthly long term mean



	X/Y	1	2	3	4	5	6	7	8	9	10	total
1	1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3	3	0.00	10.61	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	10.61
4	4	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
5	5	0.00	0.00	79.43	0.00	0.00	0.00	0.00	0.00	0.00	0.00	79.43
6	6	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
7	7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8	8	0.00	0.00	0.00	0.00	0.00	0.00	73.41	0.00	0.00	0.00	73.41
9	9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
10	10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
11	Sum	0.00	10.61	79.43	0.00	0.00	0.00	73.41	0.00	0.00	0.00	163.45

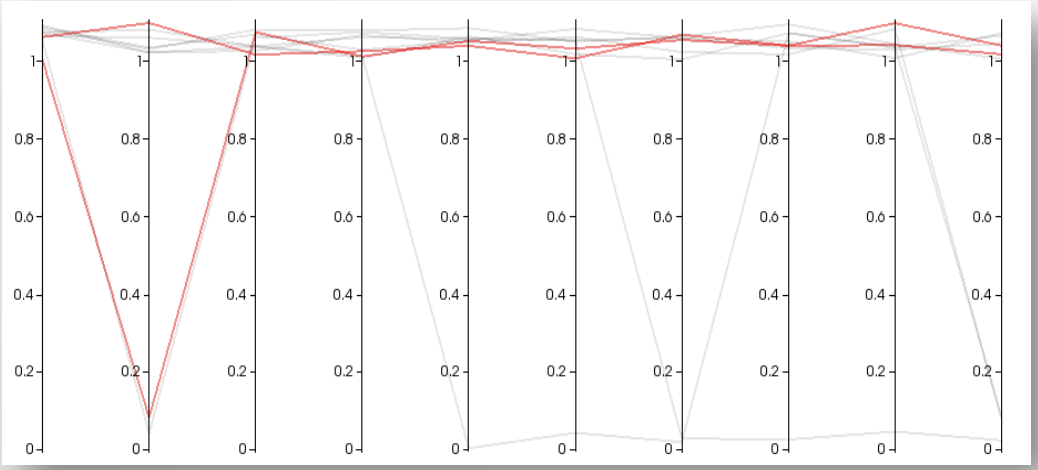
Dimensionality Reduction
(Isomap, PCA, etc.)



- High Dimensional Vectors

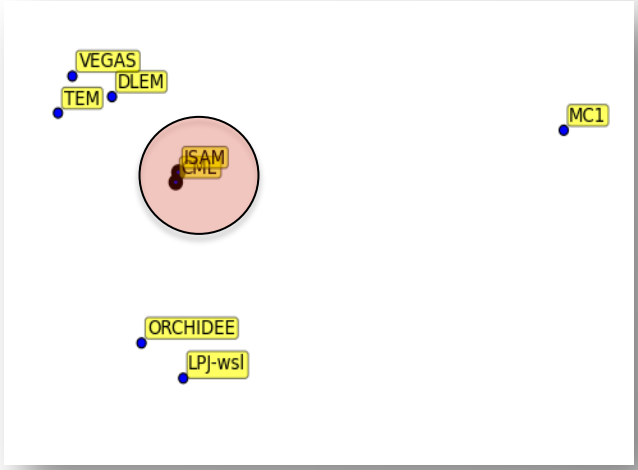
Survey Data – Model Output

Survey Data



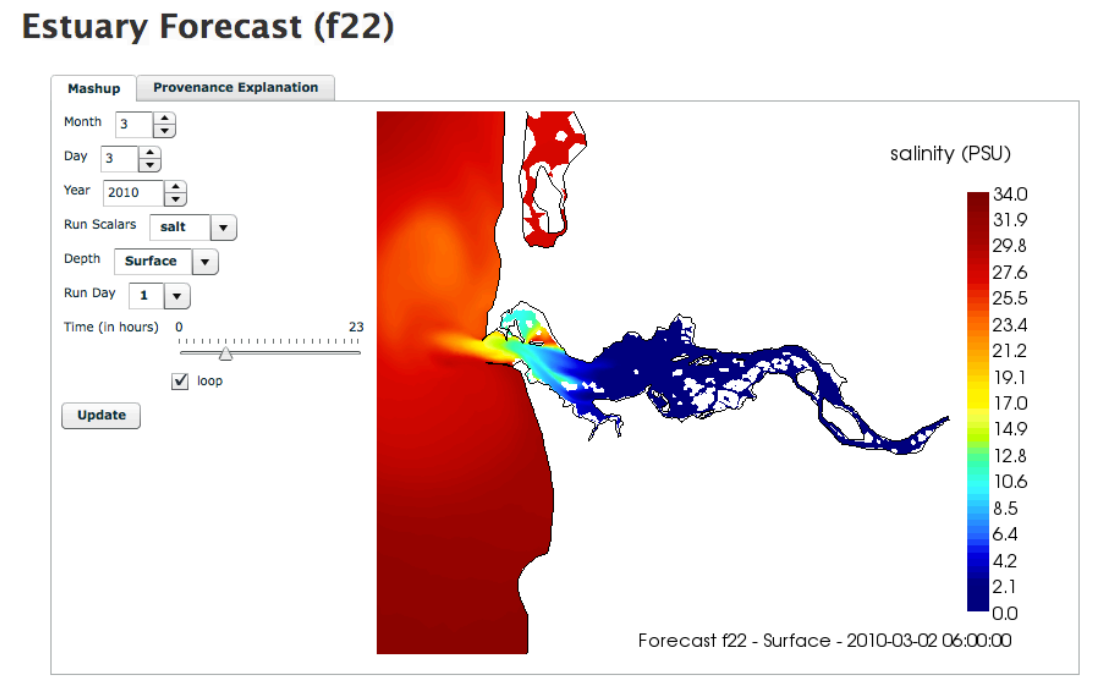
Linked Views

Model Output



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.