Scaling Information on 'Biosphere Breathing' from Chloroplast to the Globe

Dennis Baldocchi, Youngryel Ryu & Hideki Kobayashi
University of California, Berkeley

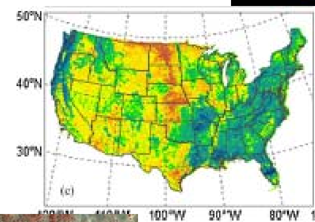
Catharine Van Ingen
Microsoft

The Breathing of an Ecosystem is Defined by the Sum of an Array of Coupled, Non-Linear, Biophysical Processes that Operate across a Hierarchy/Spectrum of Time and Space Scales

Spatial Scales of Inquiry:

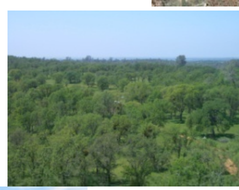
Span 13-14 orders of Magnitude

Globe: 10,000 km (10^7 m)



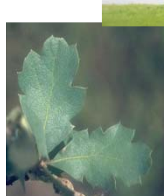
Continent: 1000 km (10^6 m)

Landscape: 1-100 km



Canopy: 100-1000 m

Plant: 1-10 m



Leaf: 0.01-0.1 m

Stomata: 10^{-5} m

Bacteria/Chloroplast: 10^{-6} m

Critical Time Scales

- **Seconds/Hours/Day, 10^{-6} to 10^{-3} yr**
  - Physiology
    - Photosynthesis, Respiration, Transpiration, Stomatal Conductance
- **Seasonal & Annual, 10^{0} yr**
  - Net and Gross Primary Productivity
  - Autotrophic and Heterotrophic Respiration and Decomposition
  - Plant Acclimation
  - Mineralization and Immobilization

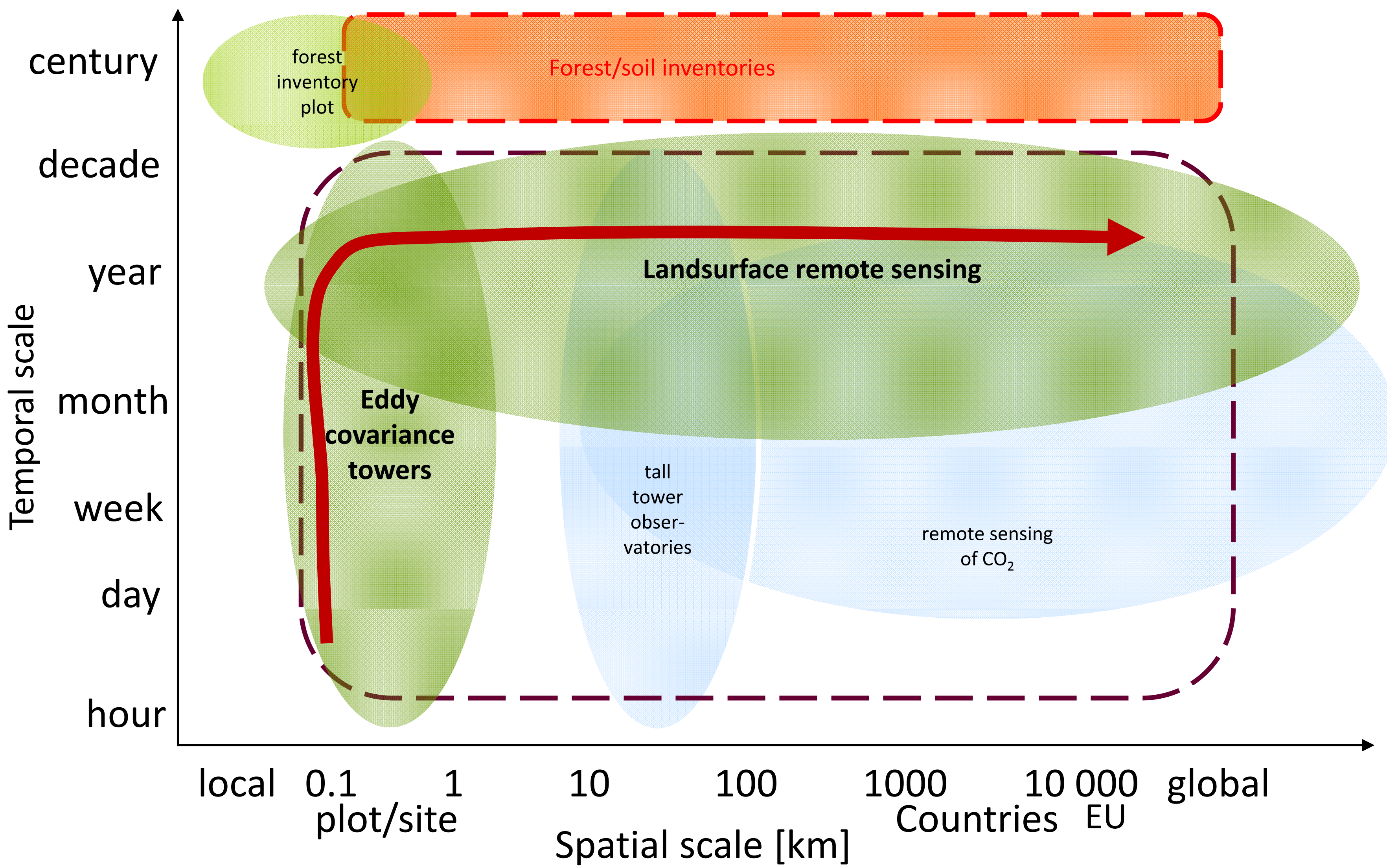
- **Decadal, 10^{1} yr**
  - Competition, Gap-Replacement, Stand Dynamics
  - Changes in Soil Organic Matter
- **Century, 10^{2} yr**
  - Succession, Mortality
- **Millennia, 10^{3} yr**
  - Species migration
  - Soil Formation

- **Geological Periods, 10^{6} yr**
  - Evolution, Speciation, Extinction, Climate Regimes
- **Eons, 10^{9} yr**
  - Evolution of Life and the Formation of our Atmosphere

Big Picture Question Regarding Predicting and Quantifying the 'Breathing of the Biosphere':

- **How can We Be 'Everywhere All the Time?'**

Components of an Integrated Earth System EXIST, but are Multi-Faceted



From: Markus Reichstein, MPI

Challenges and Points of Collaboration between the Computational and Scientific Communities

•**Part 1, Upscaling Gas Fluxes from Leaves to Landscapes**

➢Requires Complex, Monte-Carlo Light-Rendering, like that used in 'A Bug's Life'

❖A potential Microsoft Cloud Computing - UCB Partnership

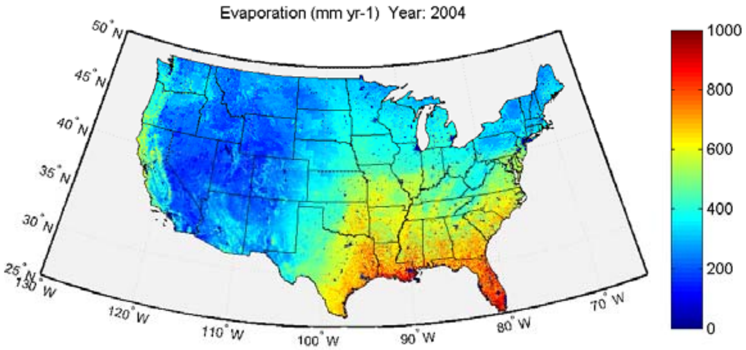
•**Part 2, Upscaling Fluxes from Landscapes to the Globe**

➢Download, Transform and Integrate High-Resolution Remote Sensing Data to produce Next Generation Carbon-Water Flux Products

❖MODIS-Azure-UCB Partnership

➢ Acquire, Process, Distill and Disseminate Vast Quantities of Flux Data for Model Ground-Truthing, Validation and Parameterization

❖FLUXNET-Microsoft Partnership



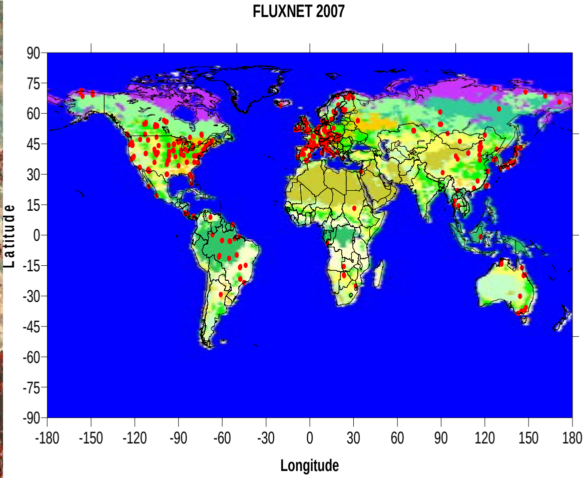
A Challenge for Leaf to Landscape Upscaling:

Transform Weather Conditions from
a Weather Station to that of the **Leaves** in a Canopy with Their Assortment of Angles and Layers Relative to the Sun and Sky

Challenge for Landscape to Global Upscaling

Converting Virtual ‘Cubism’ back to Virtual ‘Reality’

Realistic Spatialization of Flux Data
Requires the Merging Numerous Data Layers with varying Time Stamps (hourly, daily, weekly), Spatial Resolution (1 km to 0.5 degree) and Data Sources (Satellites, Flux Networks, Climate Stations)

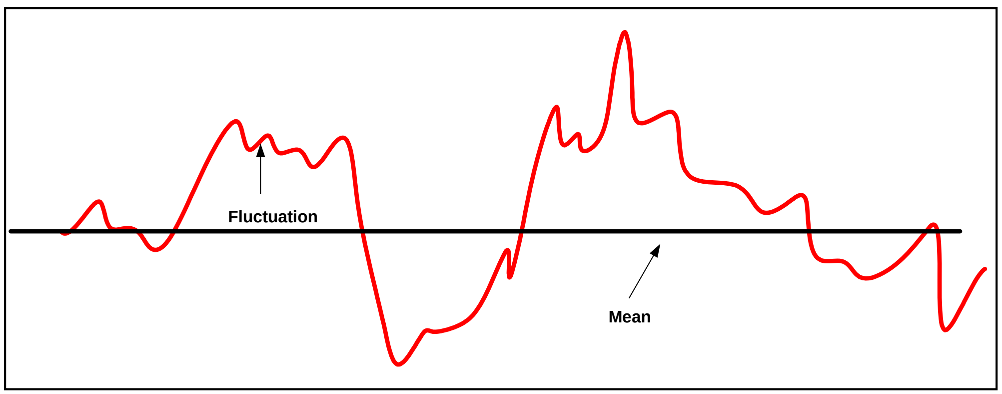


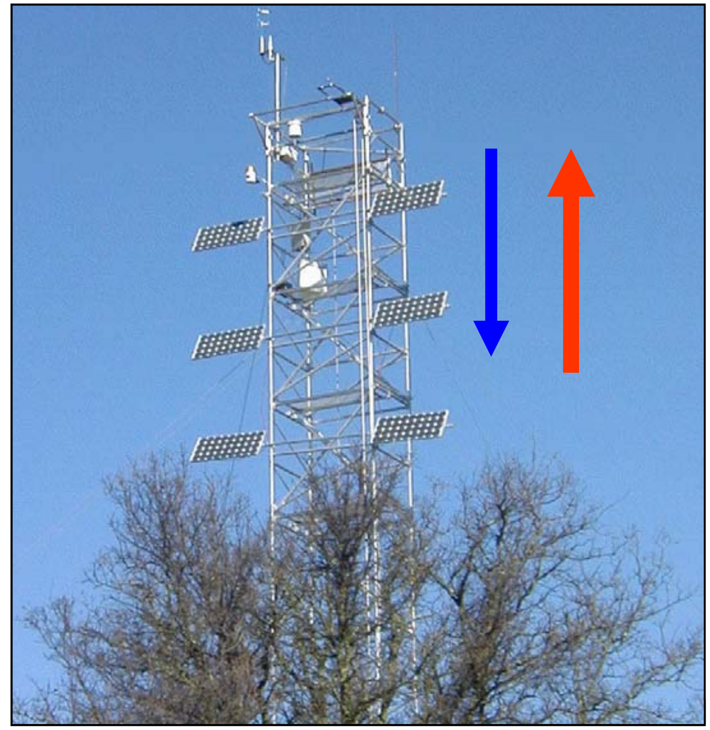
To Develop a Scientifically Defensible Virtual World
'You Must get your boots dirty', too

Collecting Real Data Gives you Insights on What is Important & Data to Parameterize and Validate Models

Eddy Covariance Fluxes

$$F = \overline{w'c'}$$

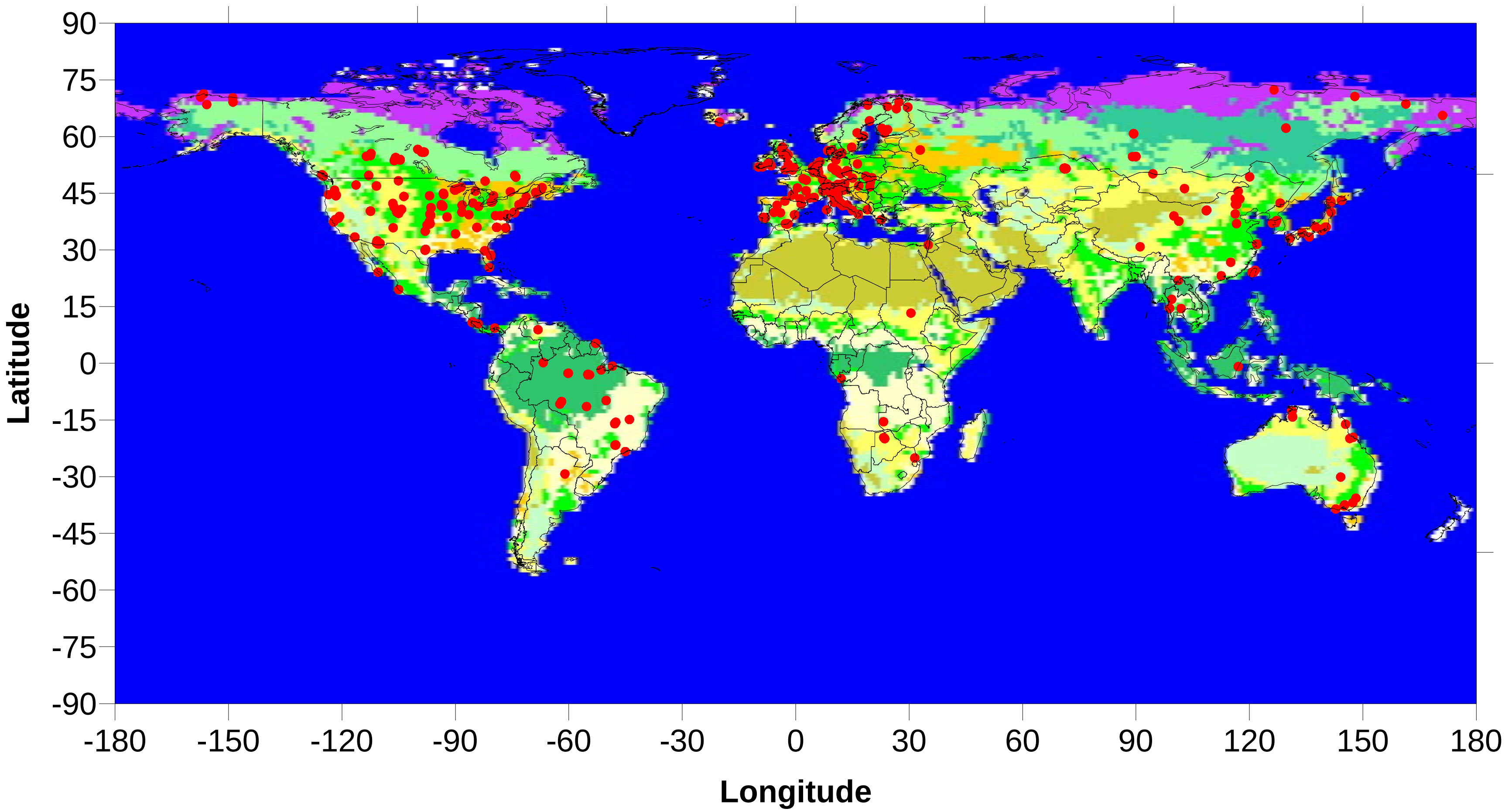




Net Flux is the sum of the mass flux moving across a plane (w C) in the up and down-drafts of Air and Wind Sampled 10 times per second

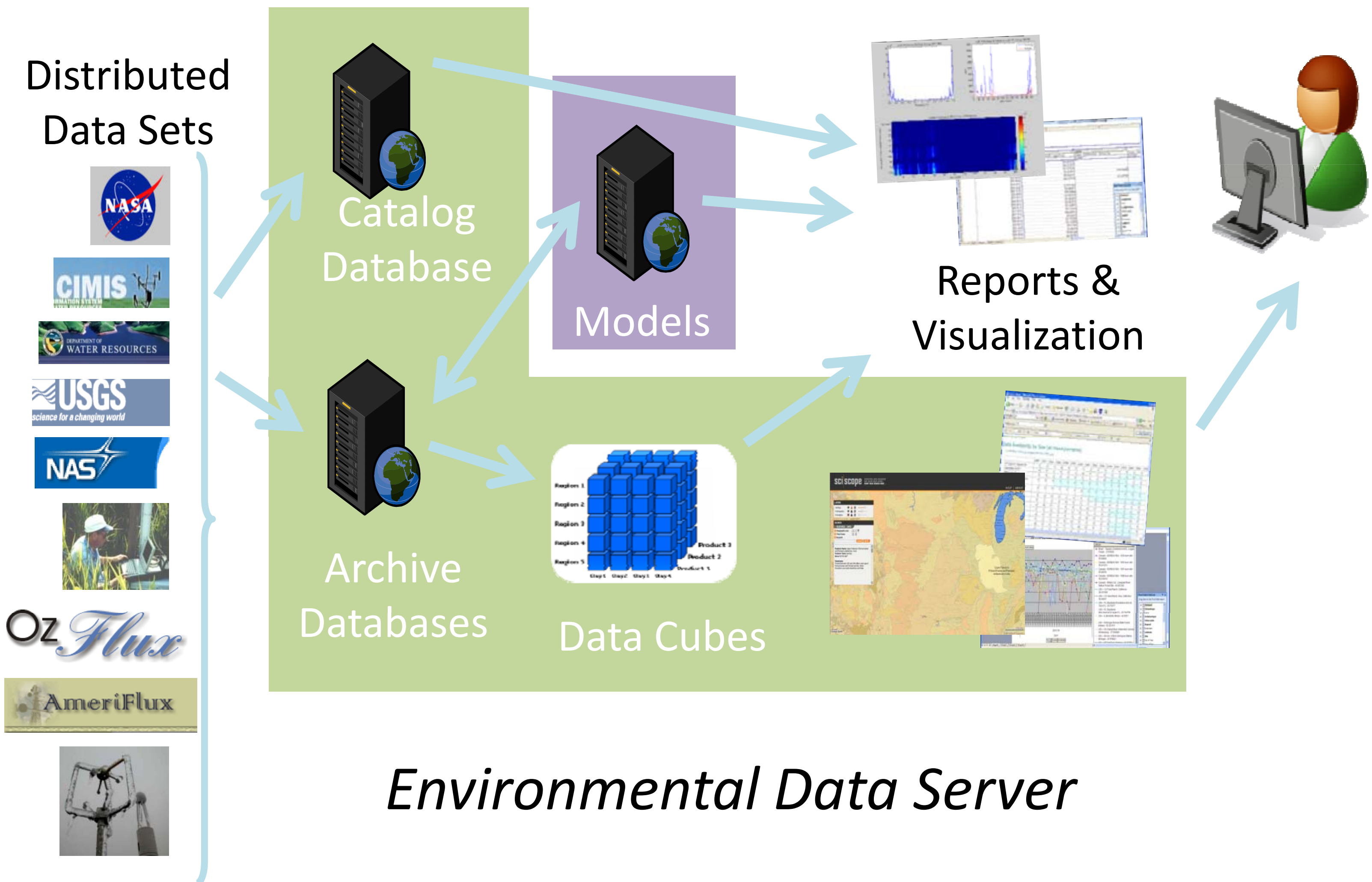
FLUXNET: From Sea to Shining Sea
500+ Sites, *circa* 2009,
960 Site-years, 30 GB
www.fluxdata.org

FLUXNET 2007

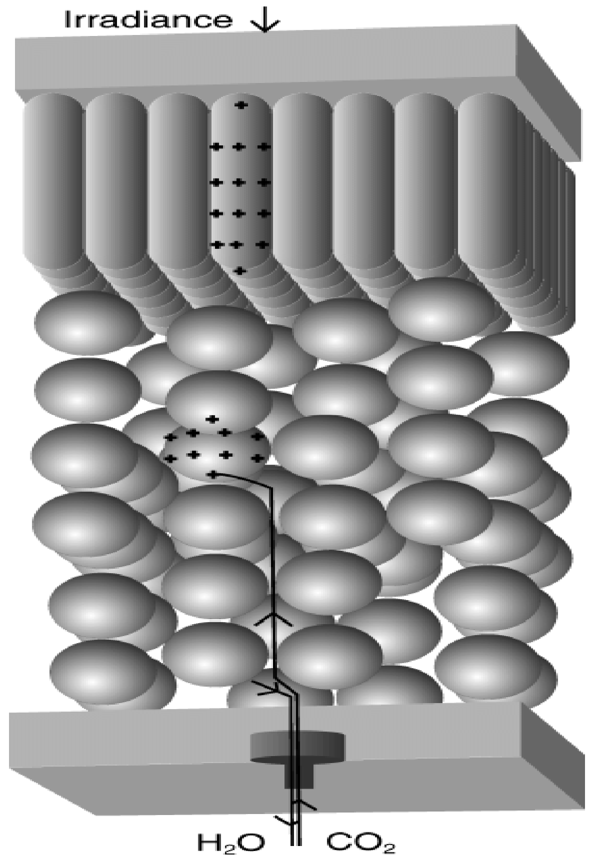


www.fluxdata.org

Connecting People and Data to Produce Information and Knowledge



Part 1, Upscaling Gas Fluxes from Leaves to Landscapes

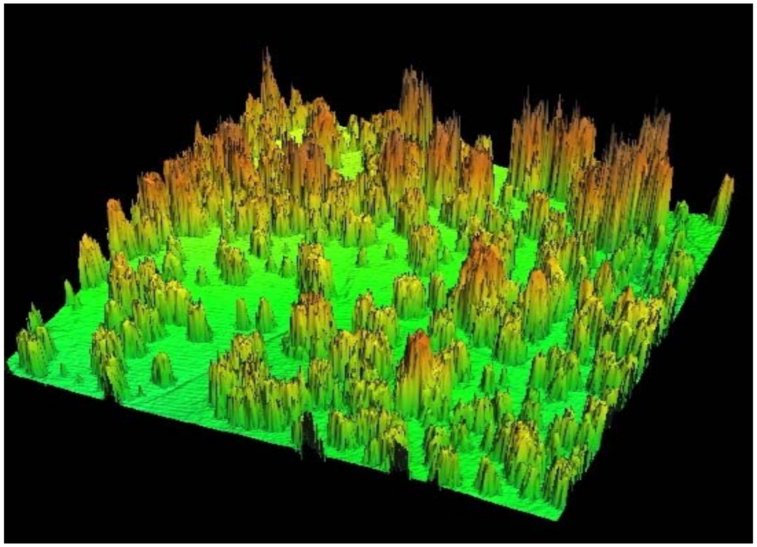


Leaf Photosynthesis Models Depend on the Light, Temperature, Humidity and CO2 Environment at the Vicinity of the Organ

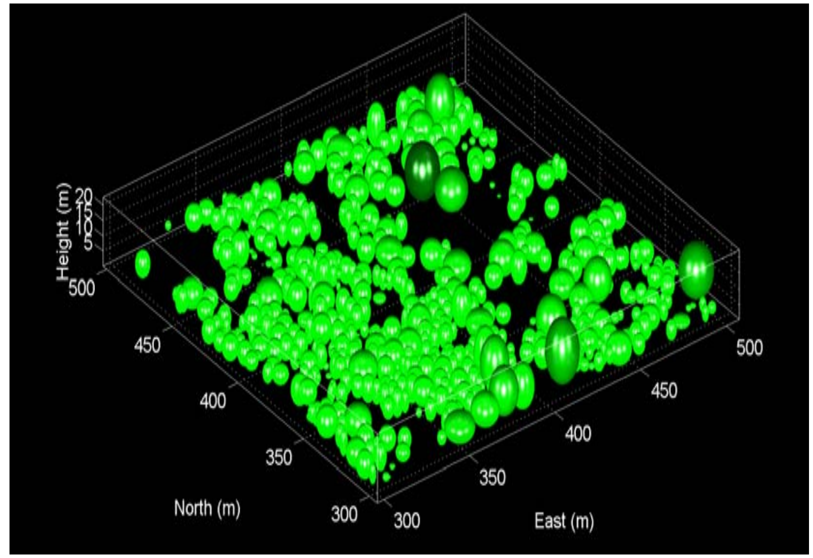
Aalto and Juurala. 2002 PCE

Hierarchy of Canopy Radiative Transfer Models

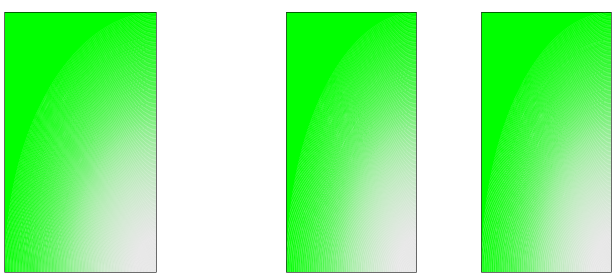
Reality



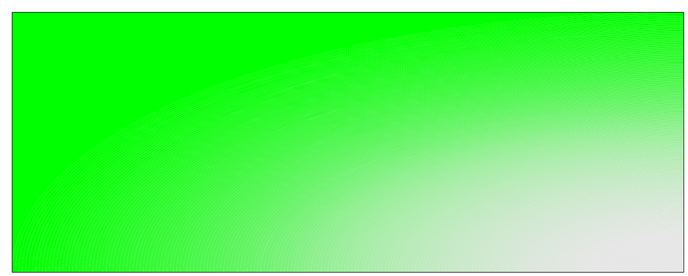
3-D Representation



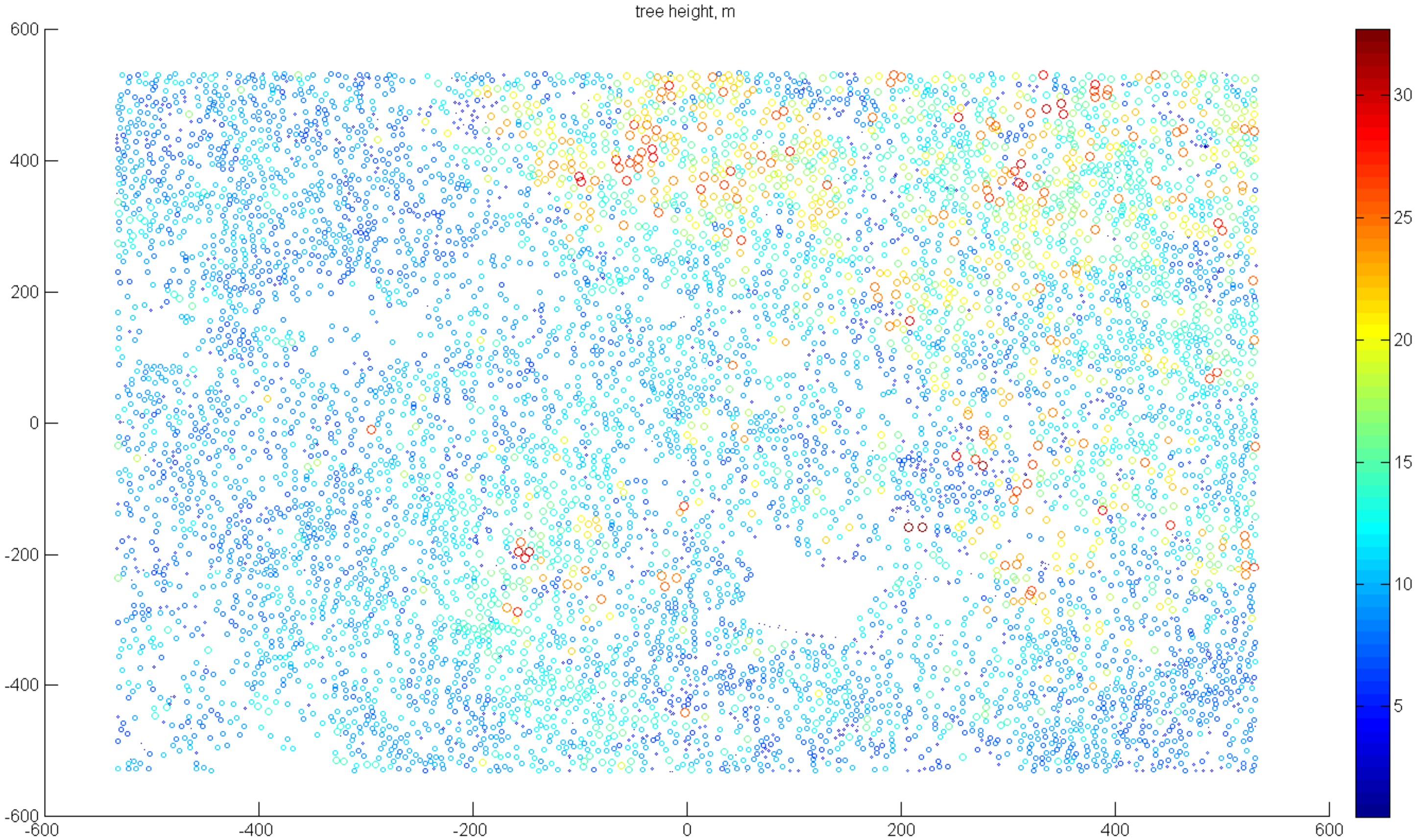
2-D Representation



1-D Representation



Doing the Impossible is Possible Today: We Can Delineate 3-D Canopy Structure via Laser Altimeter Data



CANOAK-3D

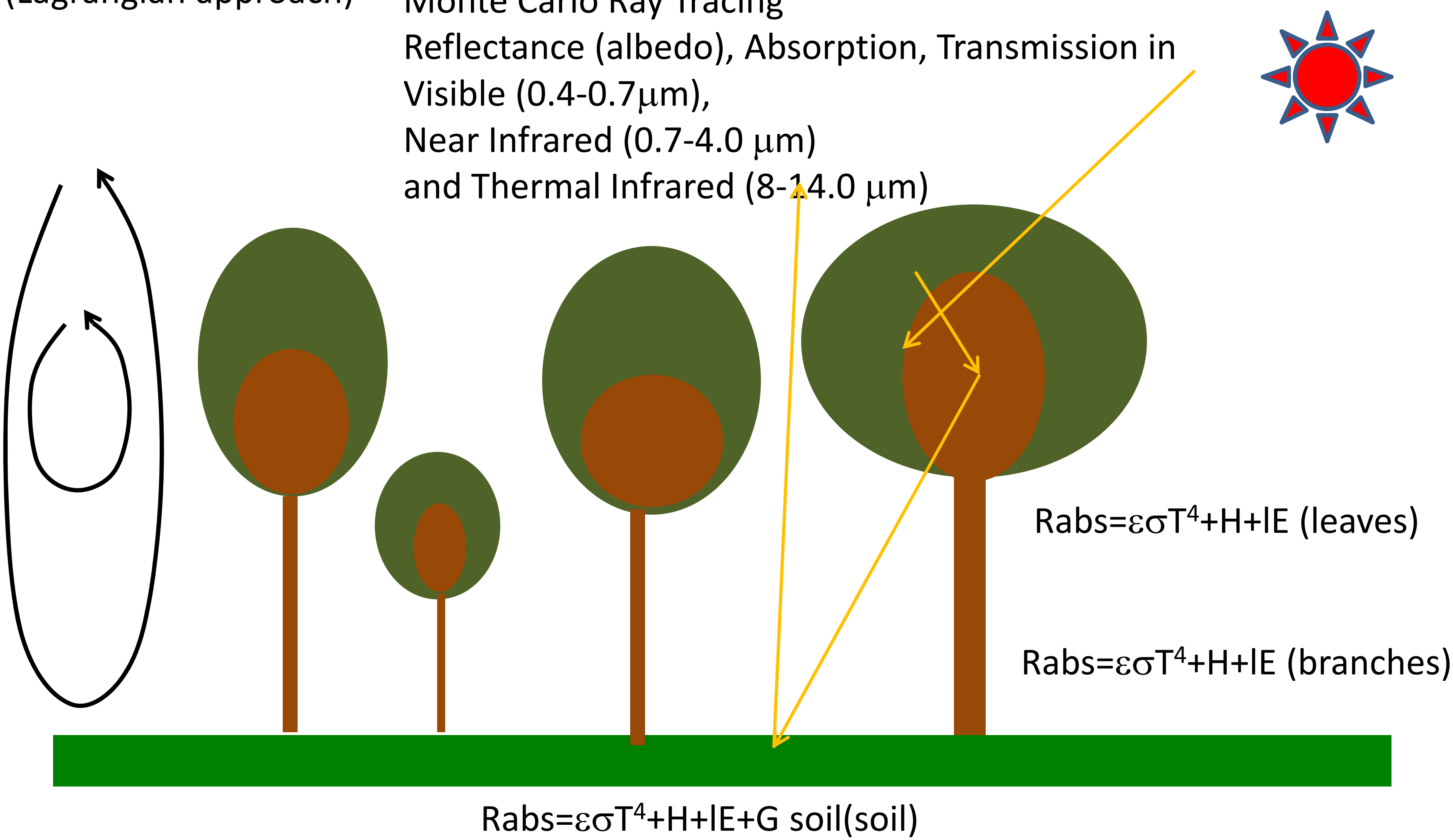
- 3D Radiative transfer + energy exchange model
- Hourly time step
- Input: air temp, humidity, solar radiation wind speed etc..

Turbulent transfer

1D turbulence scheme
(Lagrangian approach)

Radiative Transfer

Monte Carlo Ray Tracing
Reflectance (albedo), Absorption, Transmission in
Visible (0.4-0.7μm),
Near Infrared (0.7-4.0 μm)
and Thermal Infrared (8-14.0 μm)



Computational Challenges

Photons per Hour: 10^7

Trees per Hectare: 144

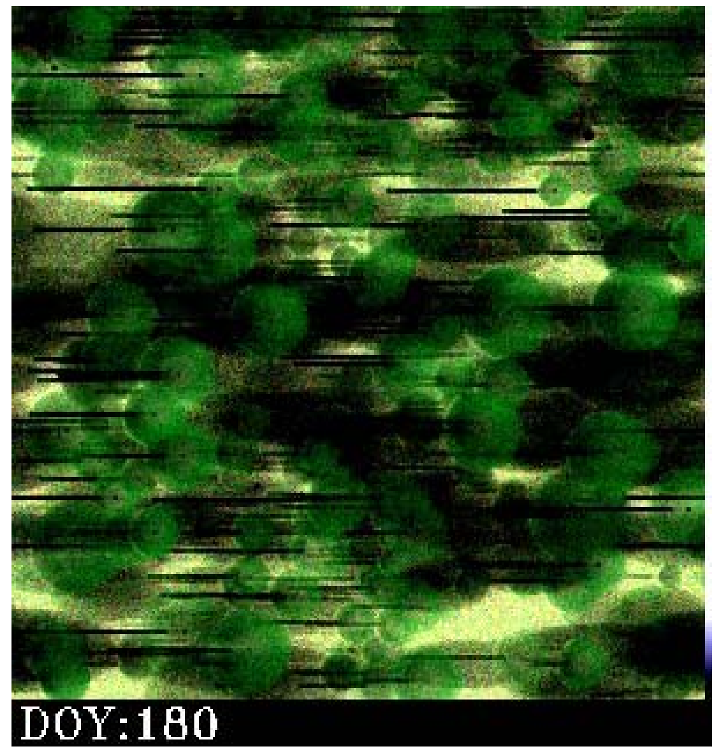
Trees per Landscape (1 km^2): 14,000

Volume per Tree: 200 m^3

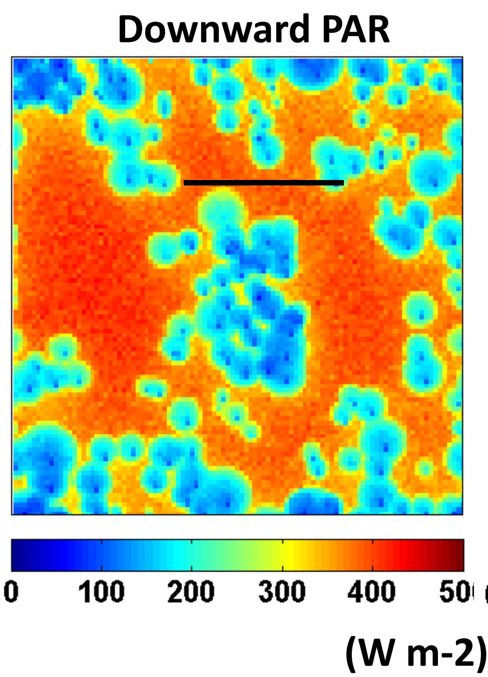
Voxel Size: 1 m^3

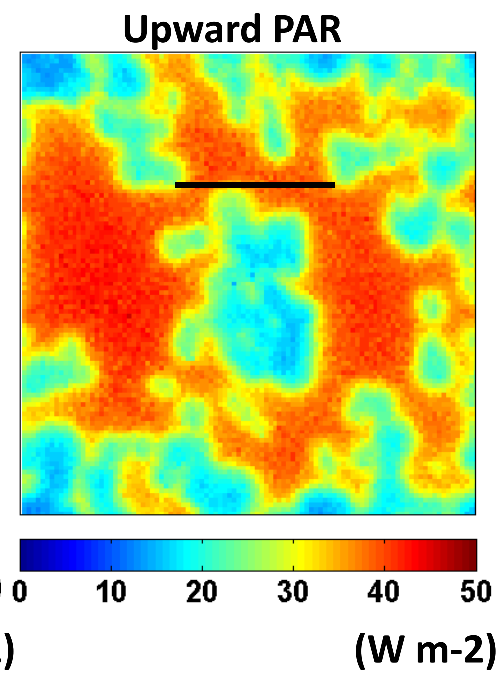
Voxels per Landscape: 2.88 10^6

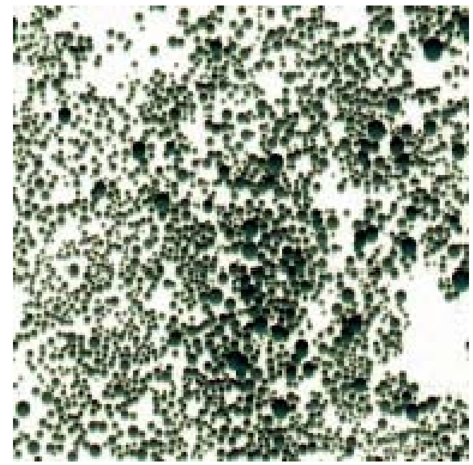
Voxels per Year per Landscape: 2.88 10^{10}

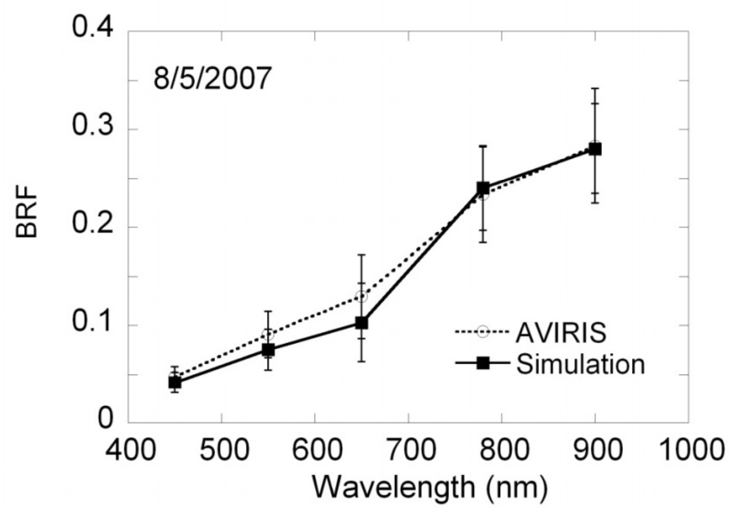


3D Radiation Fields and their Spectral Signatures Can Be Produced in Exquisite Detail



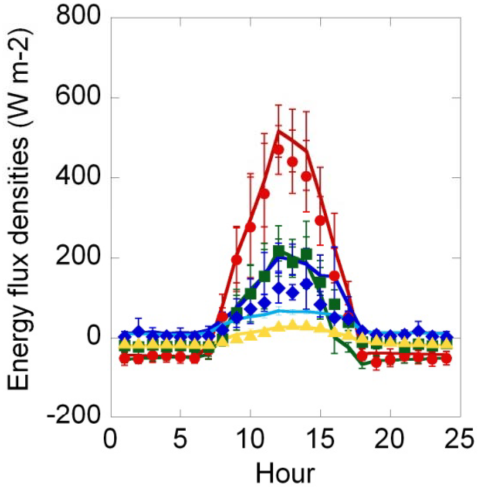




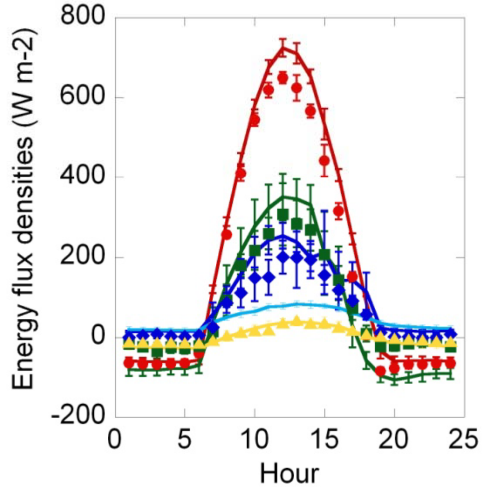


We Can Reproduce Energy Fluxes at a Flux Tower with High Fidelity

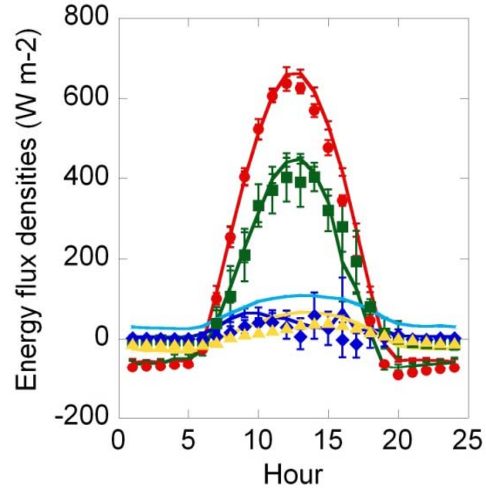
March, D68-74



May, D115-121



July, D 204-210



Net radiation

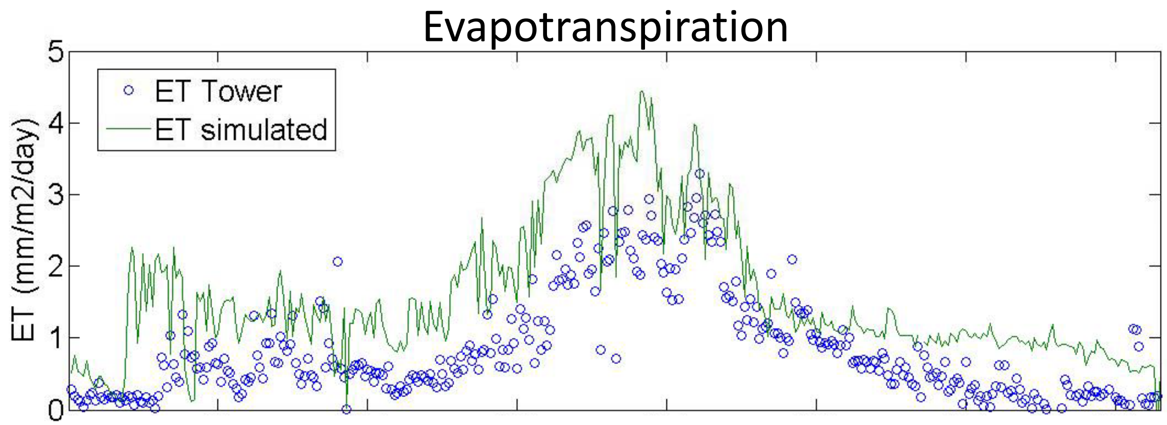
Sensible heat

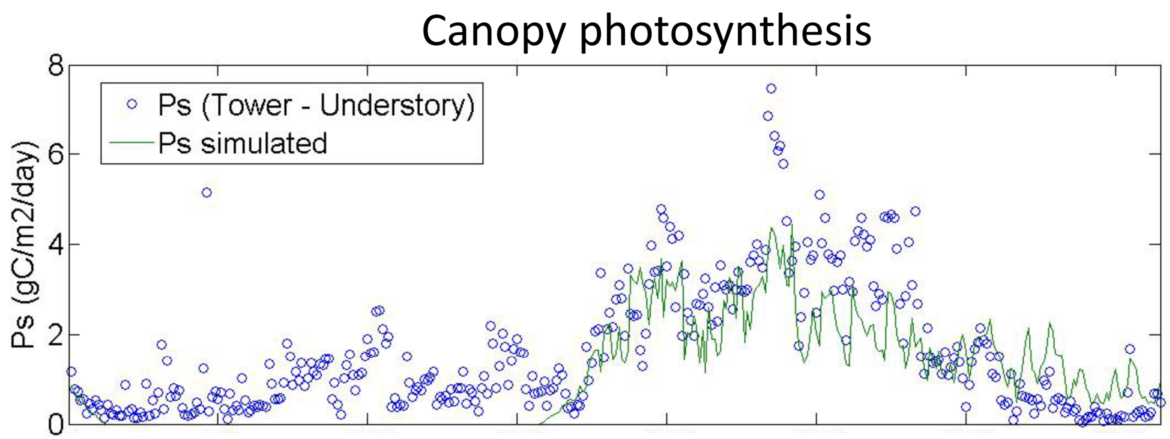
Latent heat

Soil heat flux

Woody storage

Yearly Run for CANOAK=3D





Oct. 2003 Sep. 2004

With 64 CPUs of the NCAR/CISL system, it takes 36 hours for a yearly run.

Part 2, Upscaling from Landscapes to the Globe

‘*Space: The final frontier … To boldly go where no man has gone before*’

Captain James Kirk, Starship Enterprise

- Motivation
  - Current Global-Scale Remote Sensing Products tend to rely on
    - Highly-Tuned Light Use Efficiency Approach
      - GPP=PAR*fPAR*LUE (since Monteith 1960's)
    - Empirical, Data-Driven Approach (machine learning technique)
    - Some Forcings come from Satellite Remote Sensing Snap Shots, at fine Spatial scale (< 1 km)
    - Other Forcings come from coarse reanalysis data (several tens of km resolution)
  - Hypothesis, We can do Better by:
    - Applying the Principles taught in Biometeorology 129 and Ecosystem Ecology 111 which Reflect Intellectual Advances in these Fields over the past Decade
    - Merging Vast Environmental Databases
    - Utilizing Microsoft Cloud Computational Resources

Lessons Learned from the CanOak Model

25+ years of Developing and Testing a Hierarchy of Scaling Models with Flux Measurements at Contrasting Oak Woodland Sites in Tennessee and California

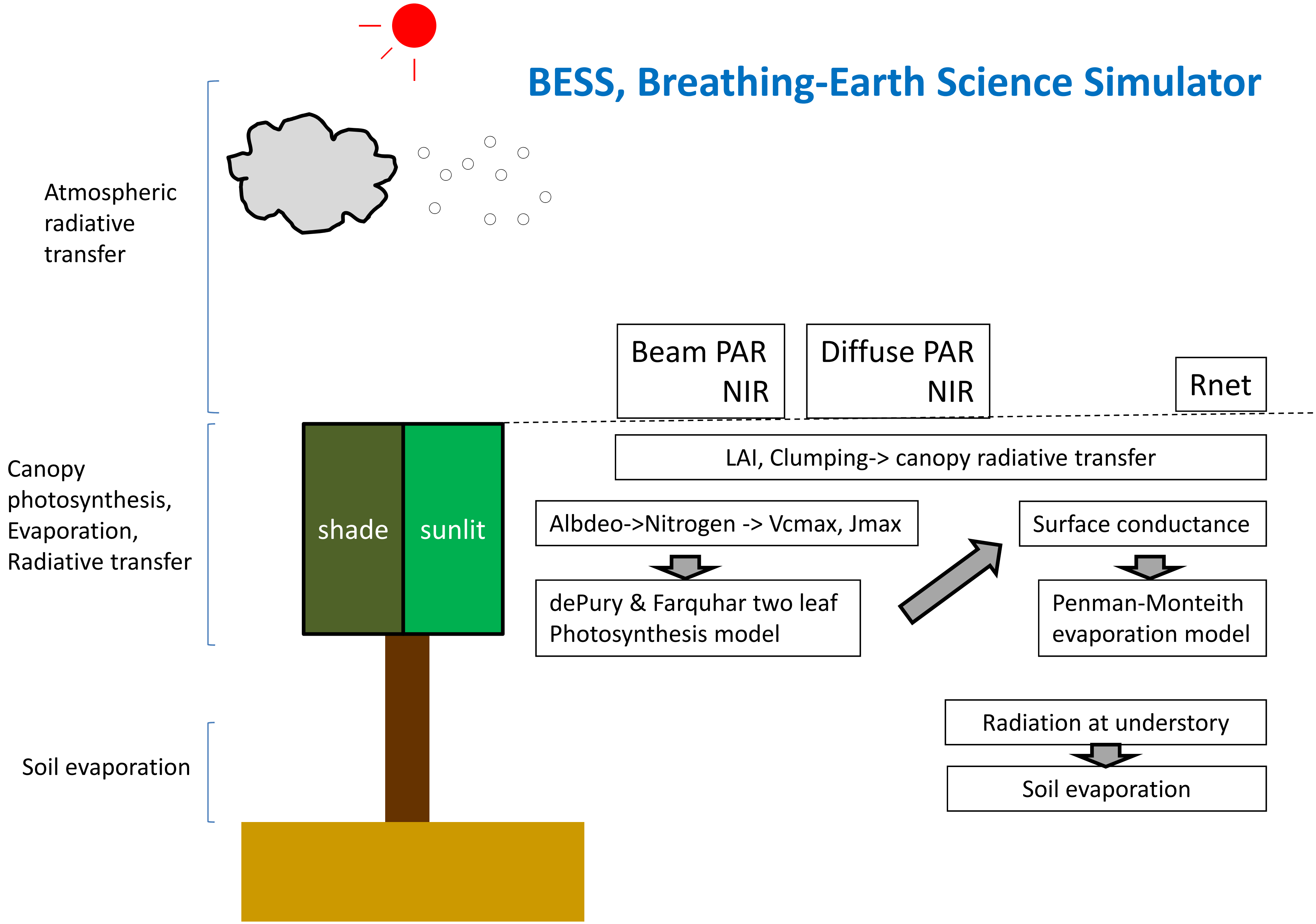
We Must:

- Couple Carbon and Water Fluxes
- Assess Non-Linear Biophysical Functions with Leaf-Level Microclimate Conditions
- Consider Sun and Shade fractions separately
- Consider effects of Clumped Vegetation on Light Transfer
- Consider Seasonal Variations in Physiological Capacity of Leaves and Structure of the Canopy

Necessary Attributes of Global Biophysical ET Model: Applying Lessons from the Berkeley Biomet Class and CANOAK

- Treat Canopy as Dual Source (Sun/Shade), Two-Layer (Vegetation/Soil) system
  - Treat Non-Linear Processes with Statistical Rigor (Norman, 1980s)
- Requires Information on Direct and Diffuse Portions of Sunlight
  - Monte Carlo Atmospheric Radiative Transfer model (Kobayashi + Iwabuchi,, 2008)
- Light transfer through canopies MUST consider Leaf Clumping
  - Apply New Global Clumping Maps of Chen et al./Pisek et al.
- Couple Carbon-Water Fluxes for Constrained Stomatal Conductance Simulations
  - Photosynthesis and Transpiration on Sun/Shade Leaf Fractions (dePury and Farquhar, 1996)
  - Compute Leaf Energy Balance to compute Leaf Saturation Vapor Pressure and Respiration Correctly
  - Photosynthesis of C_3 and C_4 vegetation Must be considered Separately
- Use Emerging Ecosystem Scaling Rules to parameterize models, based on remote sensing spatio-temporal inputs
  - Vcmax=f(N)=f(albedo) (Ollinger et al; Hollinger et al;Schulze et al.; Wright et al.)
  - Seasonality in Vcmax is considered (Wang et al.)

BESS, Breathing-Earth Science Simulator



Challenge for a Computationally-Challenged Biometeorology Lab:
Extracting Data Drivers from Global Remote Sensing to Run the Model

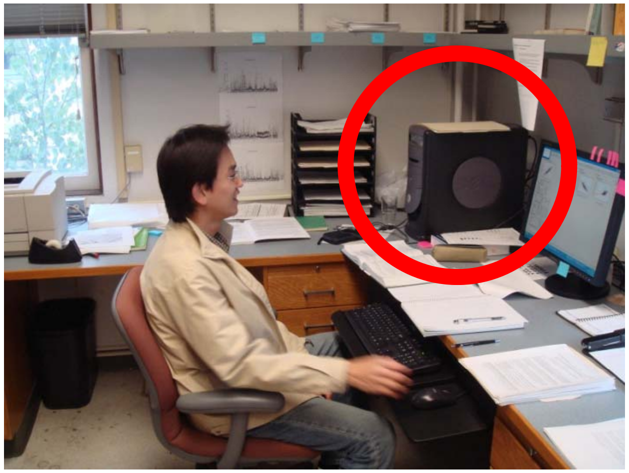
MOD04	aerosol
MOD05	Precipitable water
MOD06	cloud
MOD07	Temperature, ozone
MCD43	albedo
MOD11	Skin temperature
MOD15	LAI
POLDER	Foliage clumping

Atmospheric radiative transfer (MOD04–MOD07)

Net radiation (MOD04–MOD11)

Canopy radiative transfer (MOD15, POLDER)

Youngryel was lonely with 1 PC



Size and Number of Candidate Data Sets is Enormous

	#Source Files	Source Size	# Result Files	Result Size
USA (18)	21850	238 GB	27375	261 GB
FluxTower (3)	80670	993 GB	58400	210 GB
Global (3)	152670	2414 GB	352225	630 GB

US: 15 tiles
FluxTower: 32 tiles
Global: 193 tiles

1. **Global 1-year source data: 2.4 TB (10 yr: 24 TB)**
2. **How to know which source files are missed among >0.1 million files**

Barriers to Global Remote Sensing by the Berkeley Biometeorology Lab

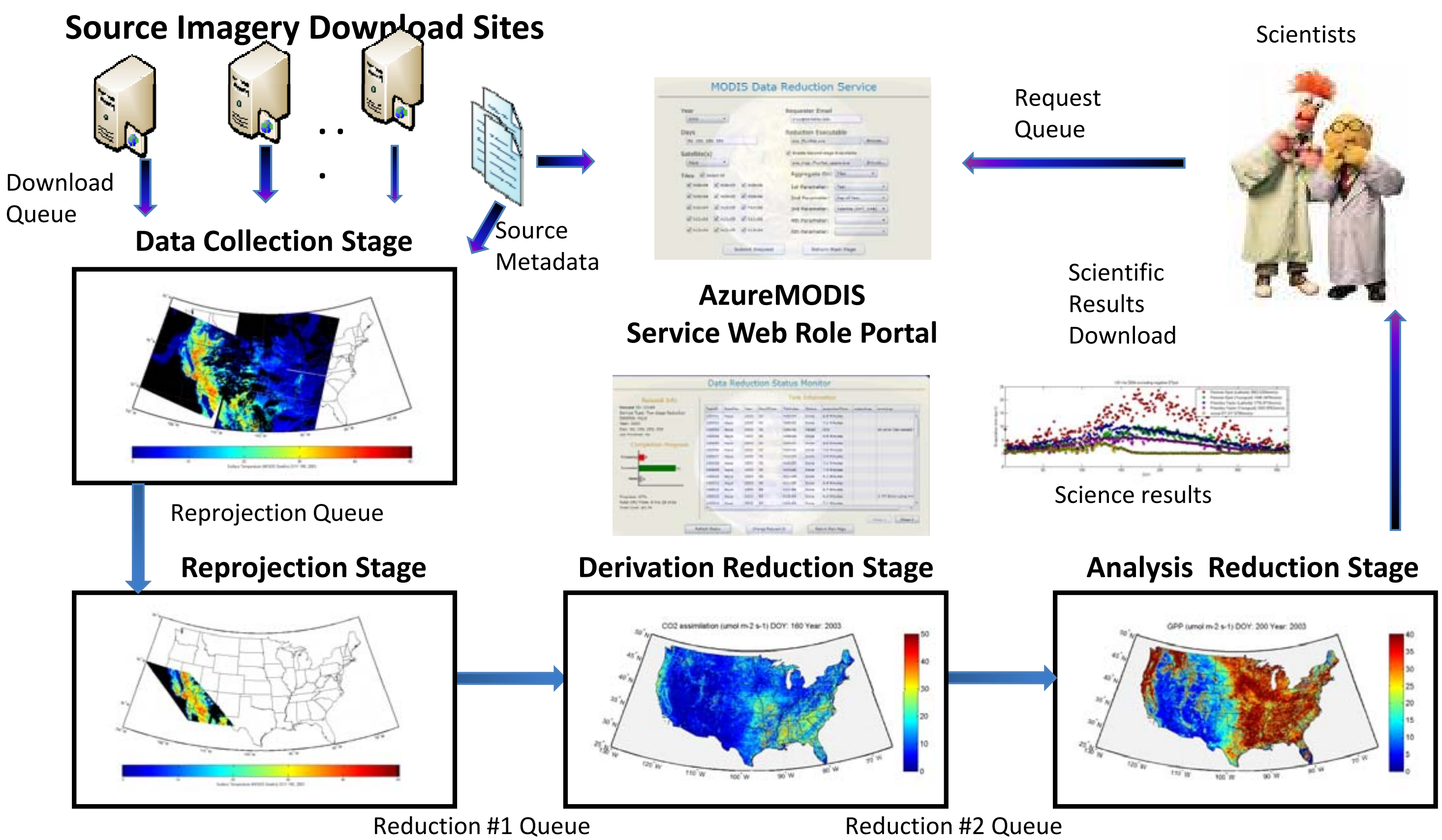
- Data processing
  - Global 1-year calculation: 9000 CPU hours
  - That is, 375 days.
  - 1-year calculation takes 1 year!







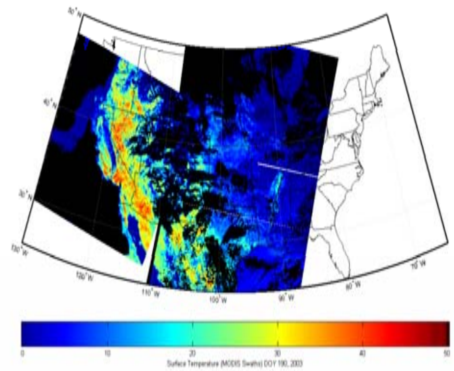
Help from ModisAzure -Azure Service for Remote Sensing Geoscience



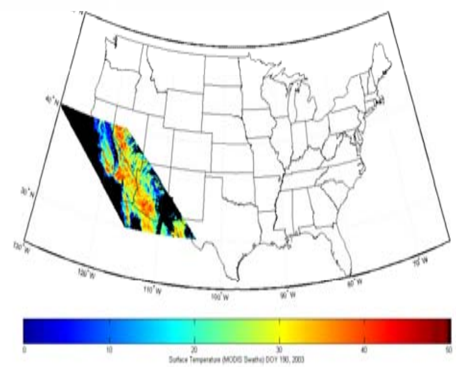
AZURE Cloud with 200 CPUs cuts 1 Year of Processing to <2 days

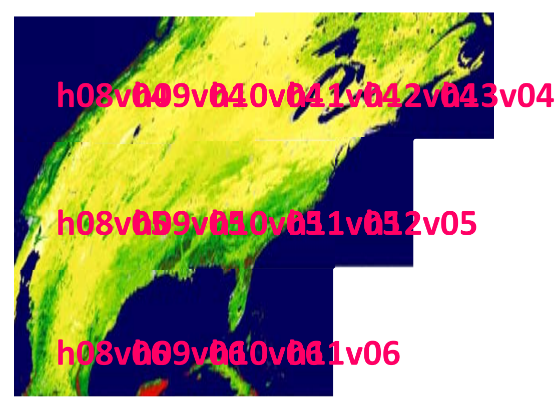
Tasked Performed with MODIS-AZURE

- Automation
  - Downloads thousands of files of MODIS data from NASA ftp
- Reprojection
  - Converts one geo-spatial representation to another.
  - Example: latitude-longitude swaths converted to sinusoidal cells to merge MODIS Land and Atmosphere Products
- Spatial resampling
  - Converts one spatial resolution to another.
  - Example is converting from 1 km to 5 km pixels.
- Temporal resampling
  - Converts one temporal resolution to another.
  - Converts daily observation to 8 day averages.
- Gap filling
  - Assigns values to pixels without data either due to inherent data issues such as clouds or missing pixels.
- Masking
  - Eliminates uninteresting or unneeded pixels.
  - Examples are eliminating pixels over the ocean when computing a land product or outside a spatial feature such as a watershed.

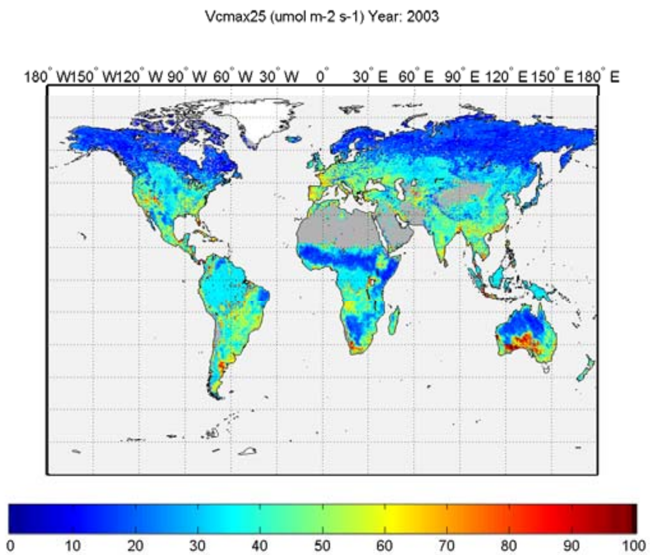


Reprojected Data (Sinusoidal format – equal land area pixel)

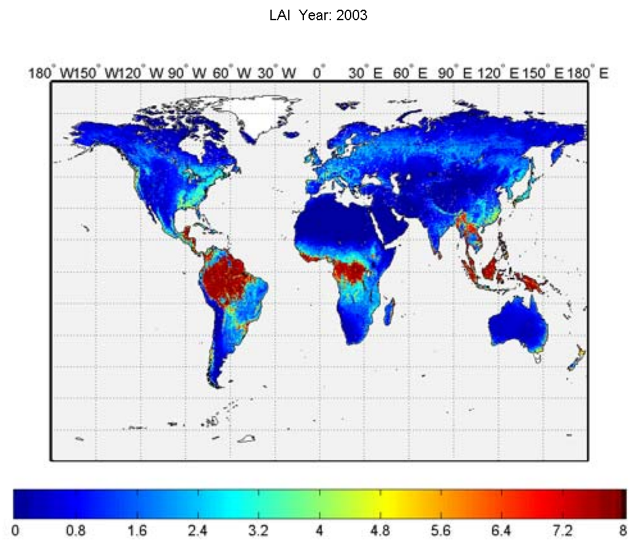




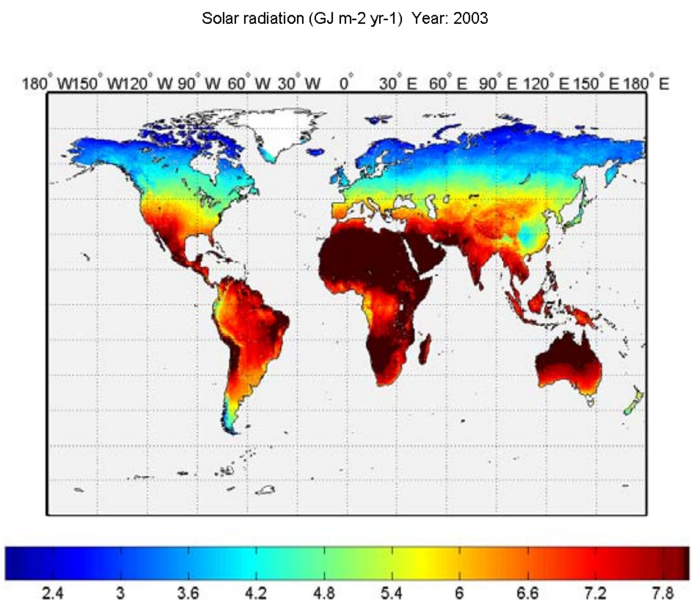
Photosynthetic Capacity



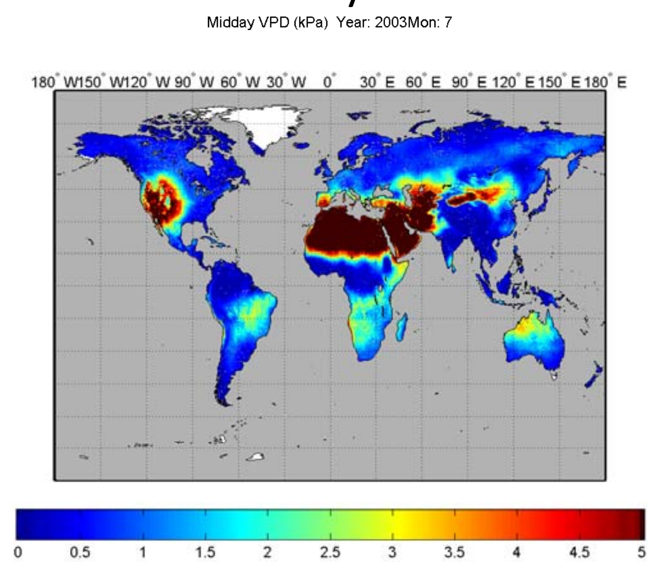
Leaf Area Index



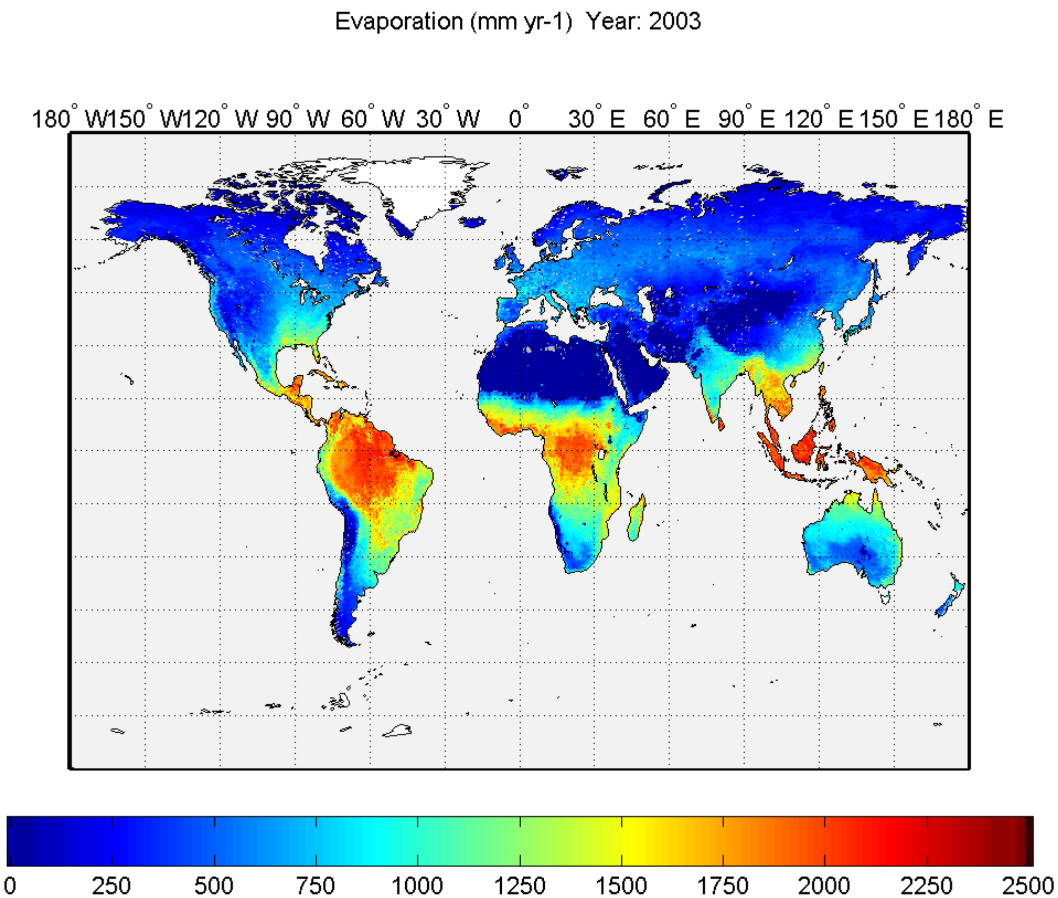
Solar Radiation



Humidity Deficits

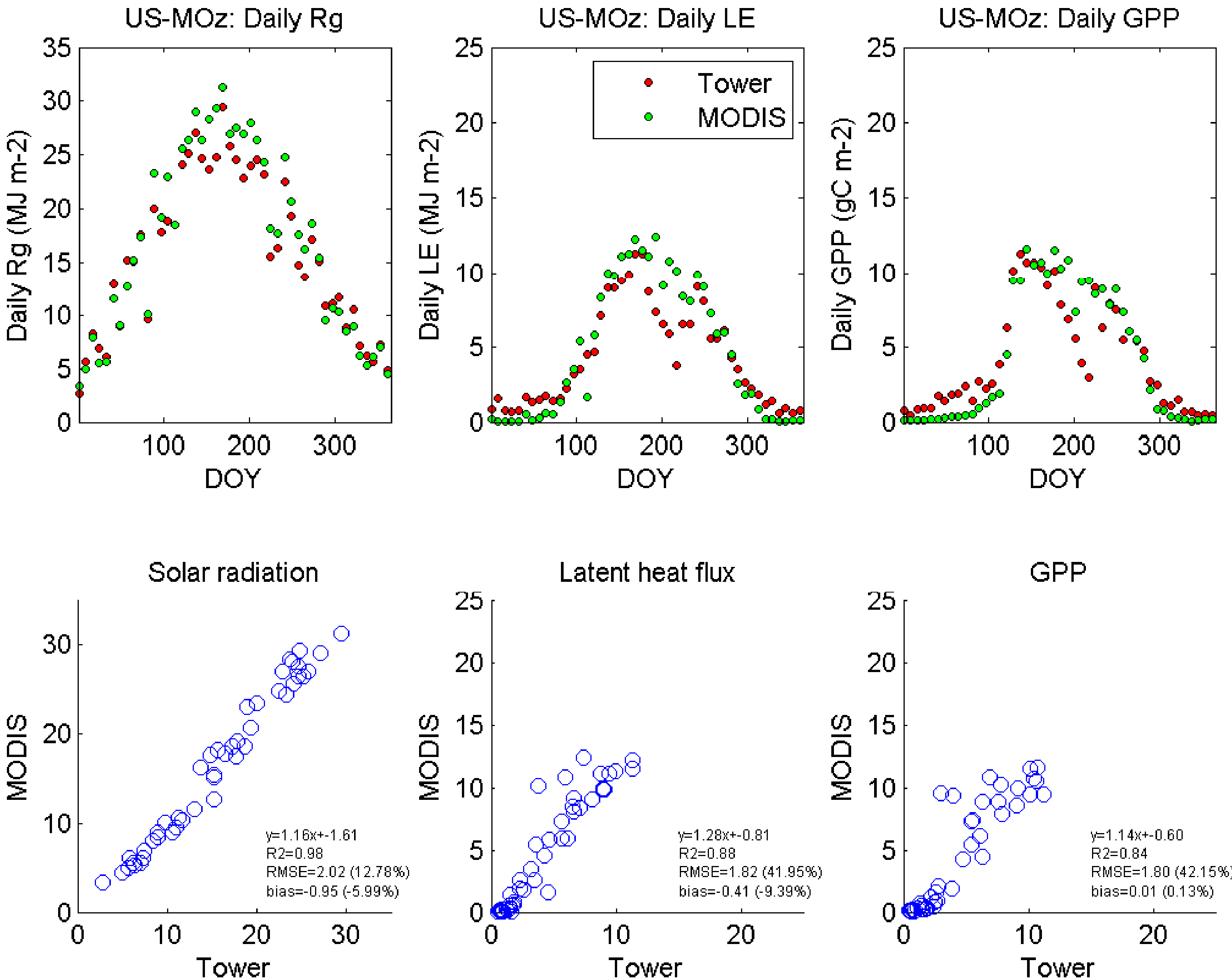


Global Evaporation at 1 to 5 km scale



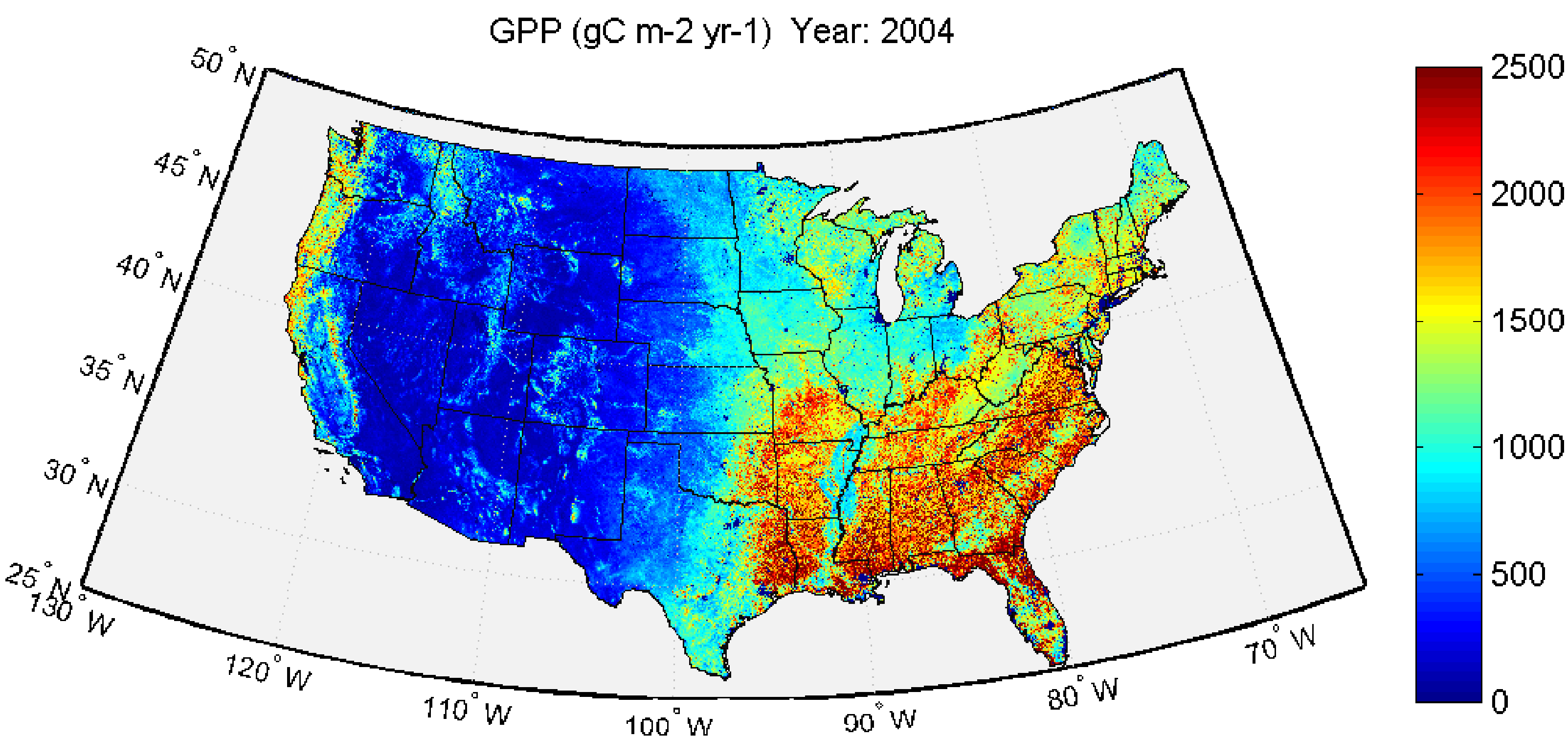
<ET> = 503 mm/y == 7.2 10^{13} m^3/y

An Independent, Bottom-Up Alternative to Residuals
based on the Global Water Balance, ET = Precipitation - Runoff



US-MOz: Daily Rg
Daily Rg (MJ m-2)
DOY
US-MOz: Daily LE
Daily LE (MJ m-2)
DOY
Tower
MODIS
US-MOz: Daily GPP
Daily GPP (gC m-2)
DOY
Solar radiation
MODIS
Tower
y=1.16x+-1.61
R2=0.98
RMSE=2.02 (12.78%)
bias=-0.95 (-5.99%)
Latent heat flux
MODIS
Tower
y=1.28x+-0.81
R2=0.88
RMSE=1.82 (41.95%)
bias=-0.41 (-9.39%)
GPP
MODIS
Tower
y=1.14x+-0.60
R2=0.84
RMSE=1.80 (42.15%)
bias=0.01 (0.13%)

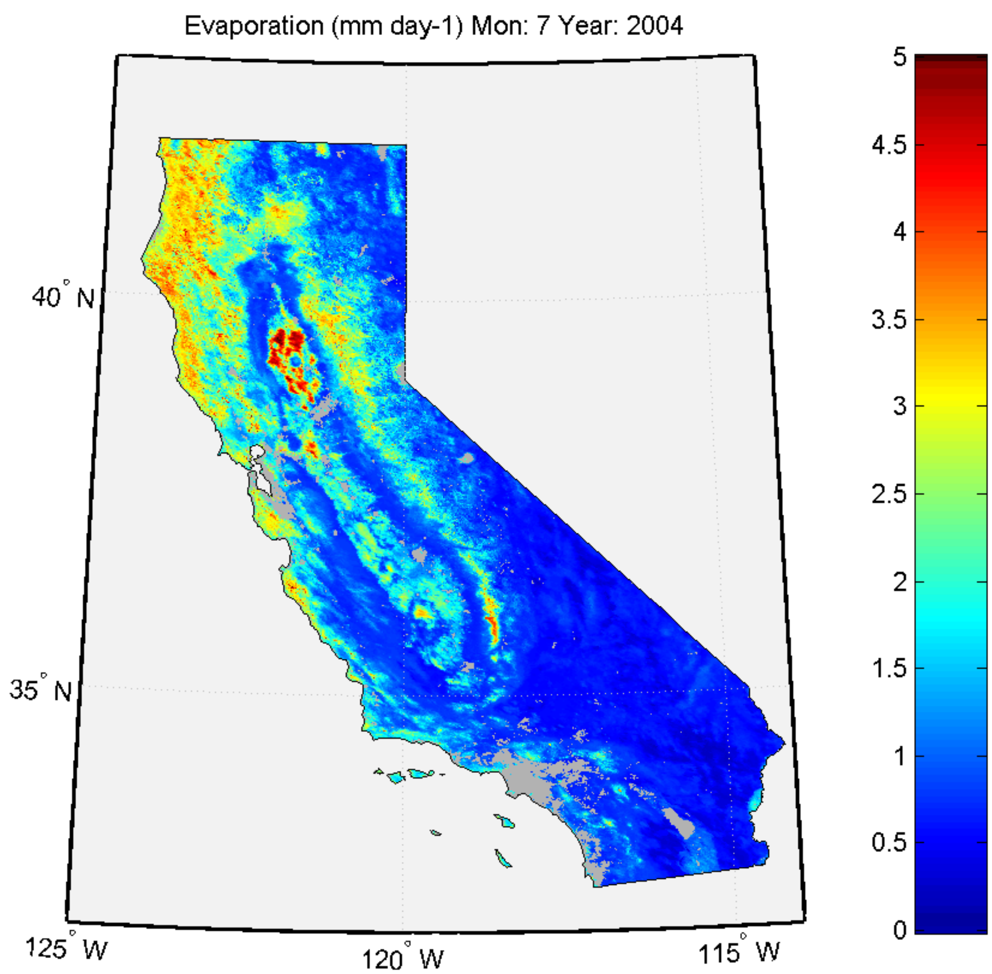
Gross Photosynthesis, GPP, Across the US



Lessons for Biofuel Production

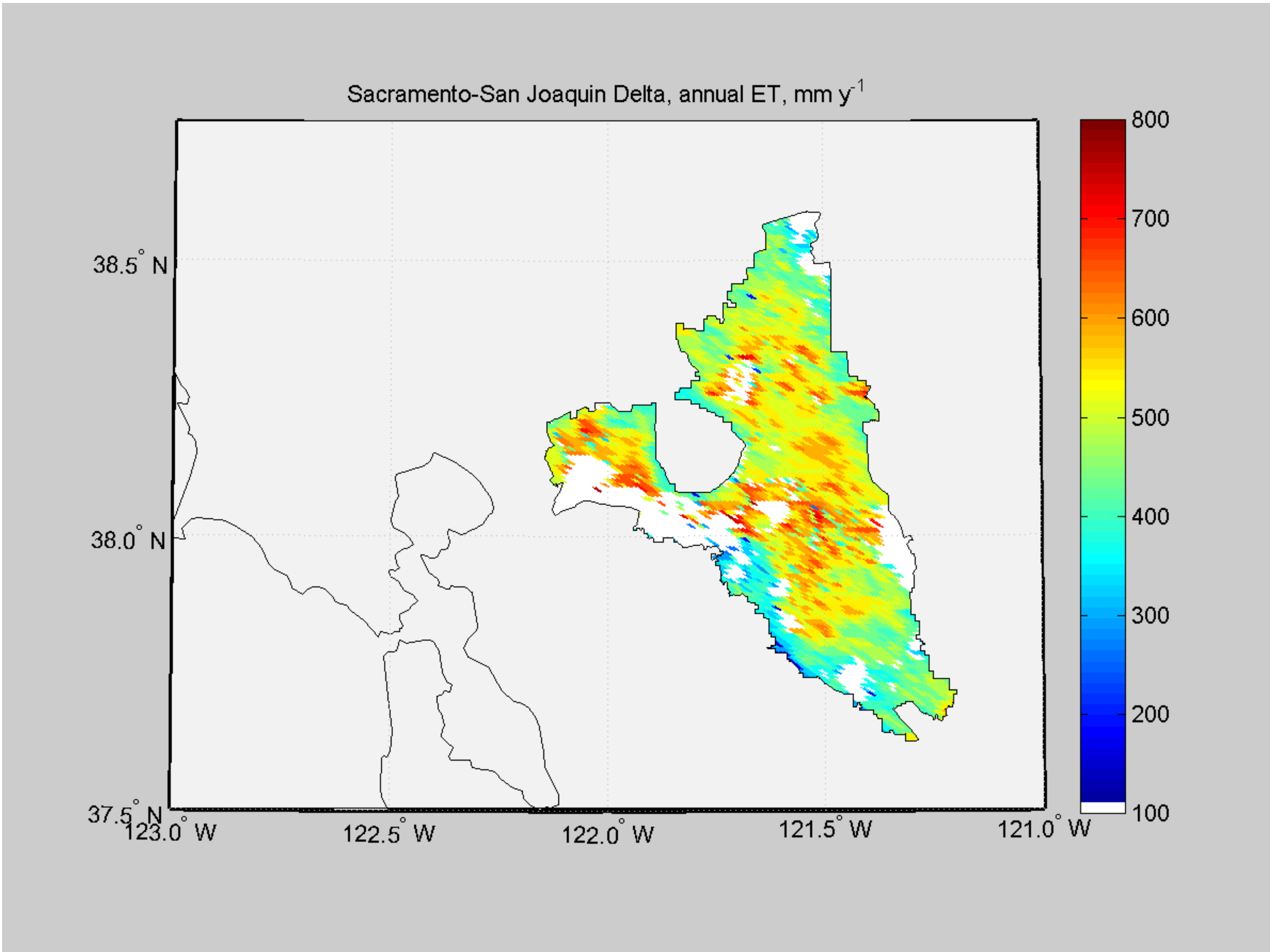
Indicates Less GPP in the Corn Belt, than the Adjacent Temperate Forests

California Evaporation at Peak of Summer Evaporative Demand

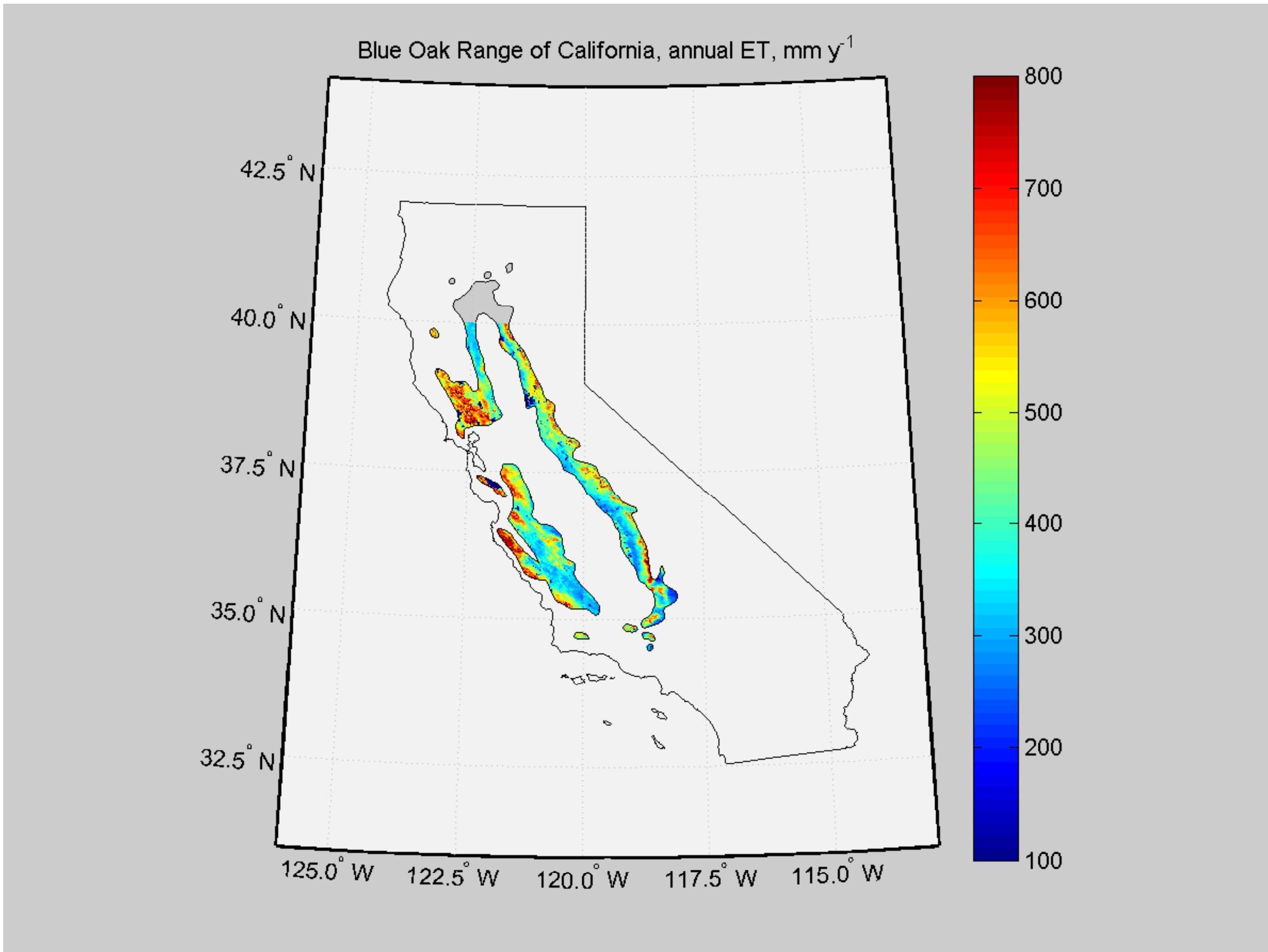


Shows ET ‘Hot-Spot’ in the Rice Growing Region of the Sacramento Valley

Water Management Issues: How Much Water is Lost from the Delta?



Regional Questions,
How Water is Used by Deciduous vs Evergreen Oak Woodlands?

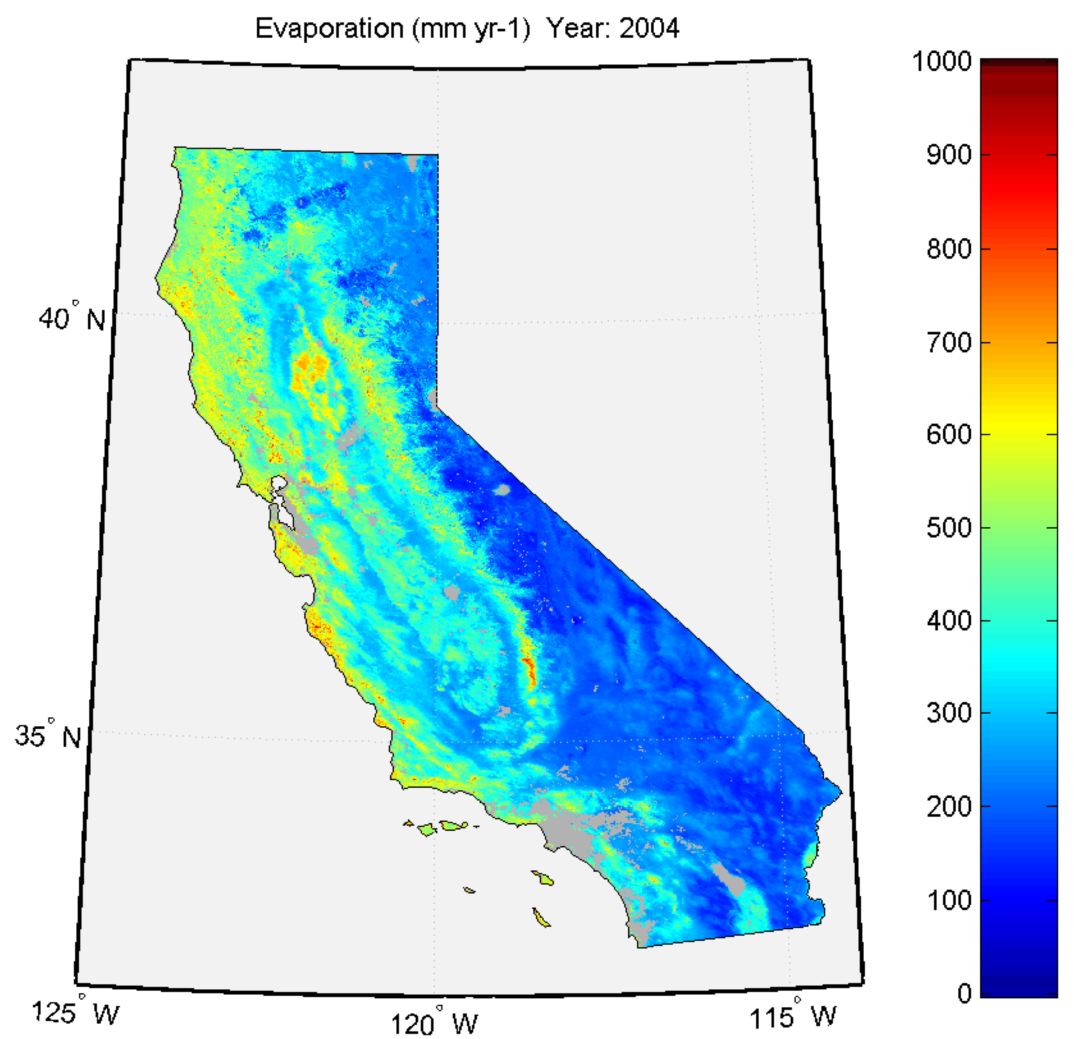


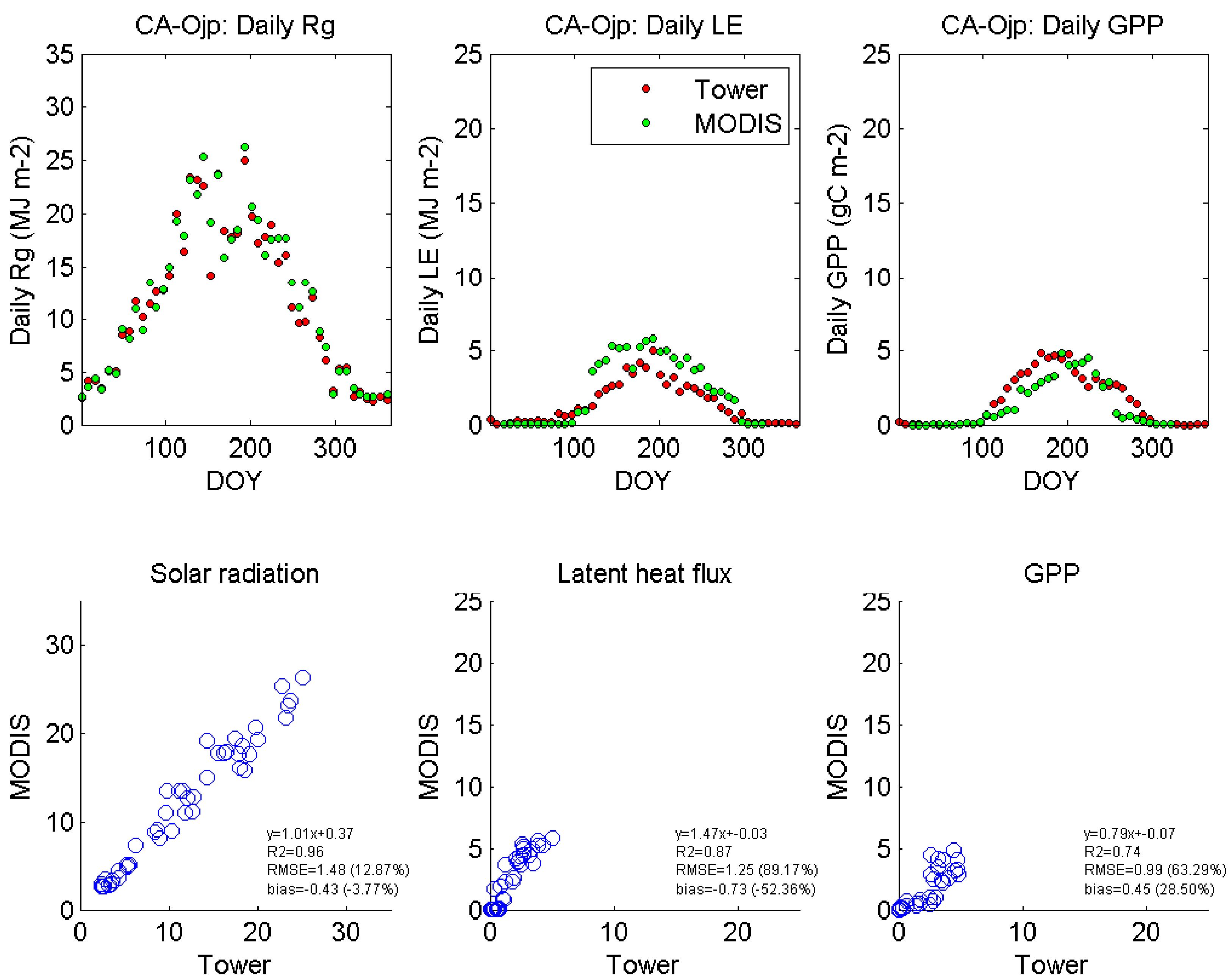
Conclusion

- Advances in Theory, Data Availability, Data Sharing and Computational Systems Enable us to Produce the Next-Generation of Globally-Integrated Products on the ‘Breathing of the Earth’
- Data-Mining these Products has Much Potential for Regional and Locale Decision making on Environmental and Agricultural Management

Extra unused

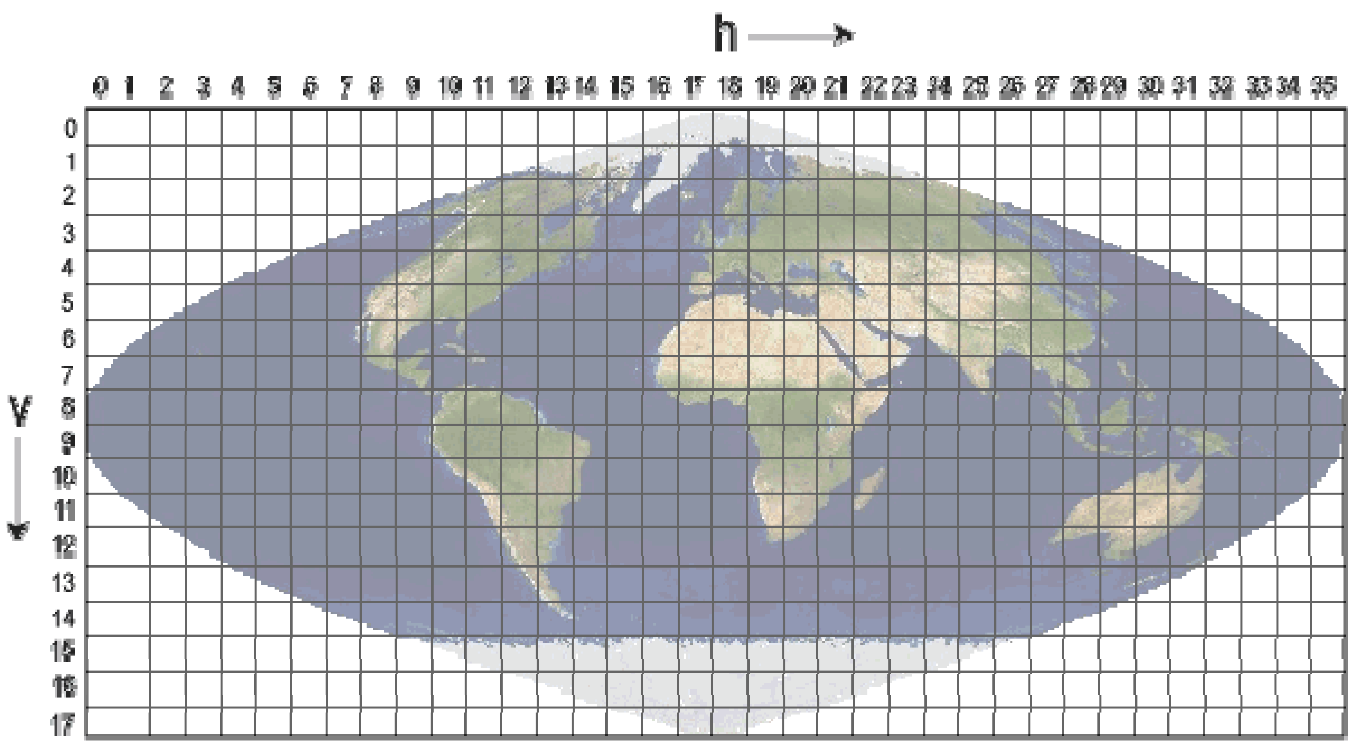
California Evaporation at 1 km Scale, annual Integral





CA-Ojp: Daily Rg
Daily Rg (MJ m-2)
DOY
CA-Ojp: Daily LE
Tower
MODIS
Daily LE (MJ m-2)
DOY
CA-Ojp: Daily GPP
Daily GPP (gC m-2)
DOY
Solar radiation
MODIS
Tower
y=1.01x+0.37
R2=0.96
RMSE=1.48 (12.87%)
bias=-0.43 (-3.77%)
Latent heat flux
MODIS
Tower
y=1.47x+-0.03
R2=0.87
RMSE=1.25 (89.17%)
bias=-0.73 (-52.36%)
GPP
MODIS
Tower
y=0.79x+-0.07
R2=0.74
RMSE=0.99 (63.29%)
bias=0.45 (28.50%)

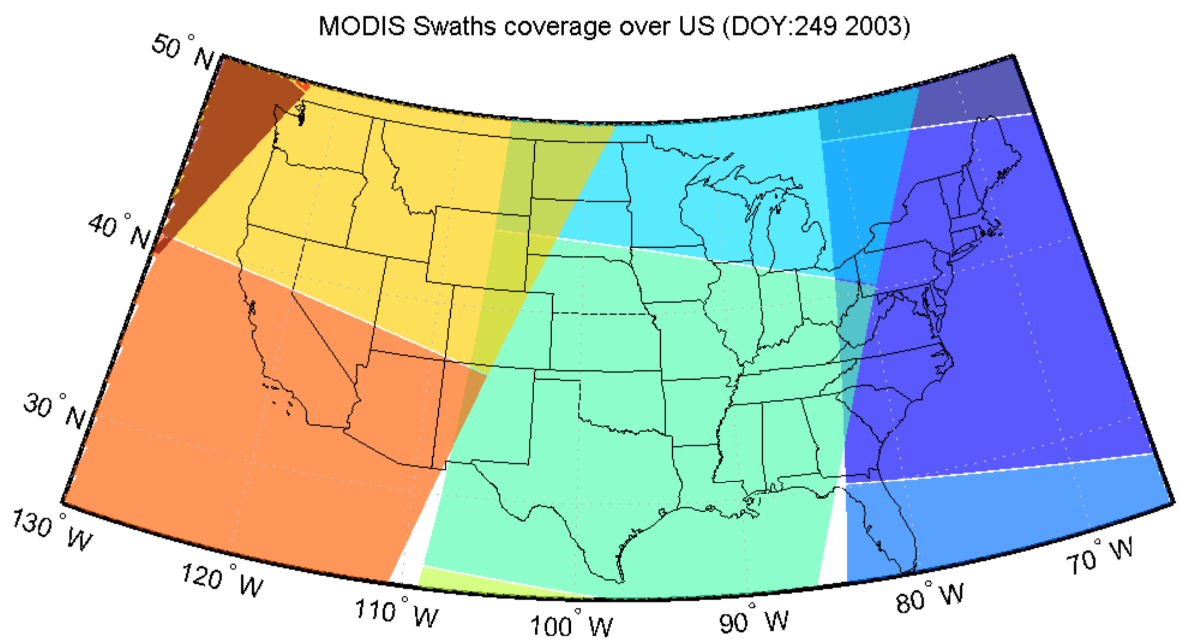
- Data standardization



MODIS Land products: standardized tiles (sinusoidal projection)

Barriers for global RS study

- 2. Data standardization



MODIS Atmospheric products: swath
=> Should be gridded to overlay with the land products

Current status

- The Cloud includes
  – 10-year MODIS Terra and Aqua data over the US (1 km resolution)
  – 3-year MODIS Terra for the global land (5 km resolution)
- Quota:
  – 200 CPUs
  – 100TB storage

Help from MODIS-AZURE

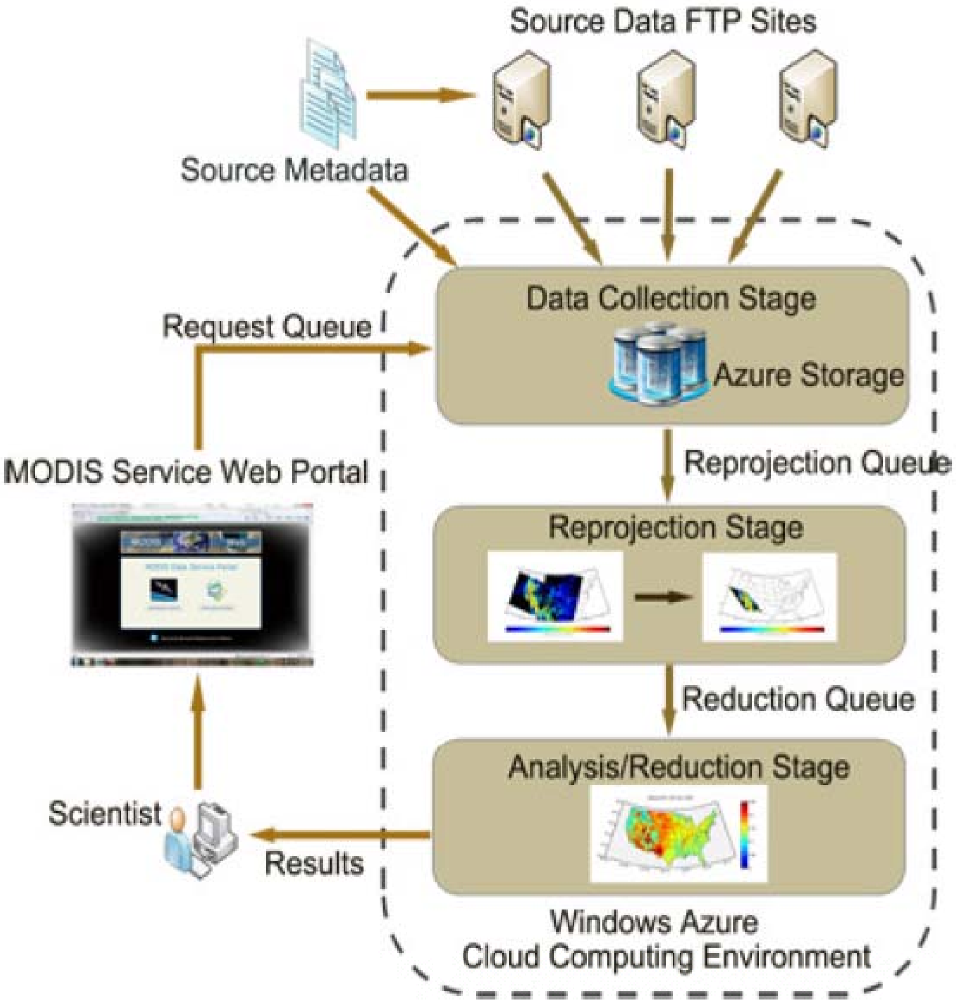


Figure 1. MODISAzure System Overview.

Necessary Attributes of the Next-Generation Global Biophysical Model, BESS

- Direct and Diffuse Sunlight
  - Monte Carlo Atmospheric Radiative Transfer model (Kobayashi, xxxx)
  - Light transfer through canopies consider leaf clumping
- Coupled Carbon-Water for Better Stomatal Conductance Simlulations
  - Photosynthesis and Transpiration on Sun/Shade Leaf Fractions (dePury and Farquhar, 1996)
  - Photosynthesis of C3 and C4 vegetation considered
- Ecosystem Scaling Relations to parameterize models, based on remote sensing spatio-temporal inputs
  - Vcmax=f(N)=f(albedo) (Ollinger et al; Hollinger et al;Schulze et al.; Wright et al.
  - Seasonality in Vcmax is considered
- Model Predictions should Match Fluxes Measured at Ecosystem Scale hourly and seasonally.