Land Cover, Phenology, Leaf-Area Index, PAR Fraction Absorption, and Bidirectional Reflectance/Vegetation Structure from MODIS for Terrestrial Carbon Modelling

Alan Strahler, Boston University,
And the BU MODIS Land Team

CSIRO Center for Atmospheric Research,

Melbourne, 22 October 2003

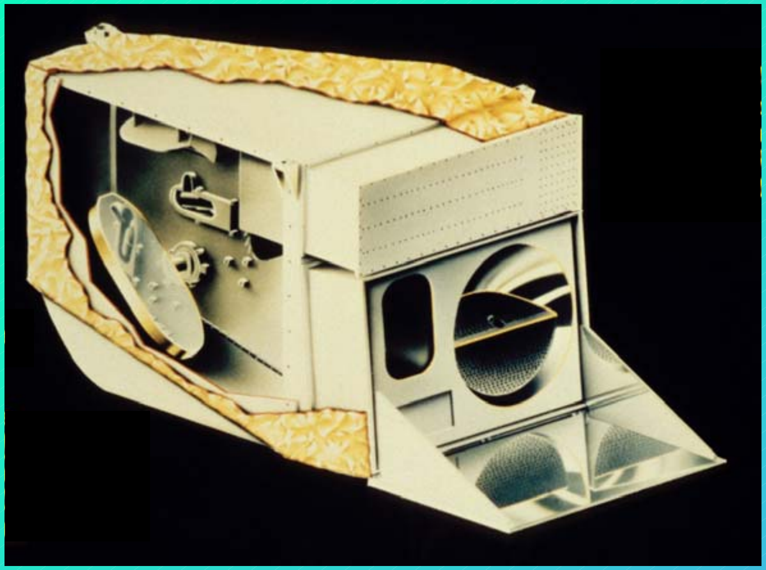
Terra Orbit and MODIS Swath

Terra Instrument Swaths

MODIS: System Characteristics

- Orbit—EOS-Terra Platform
  - Sun-synchronous, near-polar, 705.3 km, 98.21° inclination
  - 10:30 AM local solar equatorial crossing time (descending node)
  - Launched December 18, 1999
- Orbit—EOS-Aqua Platform
  - Sun-synchronous, near-polar, 705.3 km, 98.21° inclination
  - 1:30 PM local solar equatorial crossing time (ascending node)
  - Launched May 4, 2002

MODIS: System Characteristics, Cont.



Instrument Characteristics

- 36 spectral bands, VNIR, SWIR, TIR (0.4–14 µm)
- Spatial resolutions at 250-, 500-, and 1000-m (nadir) depending on waveband

- Whiskbroom design—double-sided rotating mirror, 10 lines per scan
- Four focal planes: Visible, NIR, SW-Midwave IR, Thermal IR
- Scan angle: ±55°, 2330–km swath
- Repeat: 2-day global repeat, 1-day or less poleward of 30° lat.

MODIS: System Characteristics,Cont.

- Onboard calibration
  - Spectroradiometric calibration assembly (SRCA)
  - Solar diffuser with stability monitor
  - Blackbody
  - Space view, lunar view
- Registration
  - Band-to-band registration ≤0.2 IFOV across focal planes, ≤0.1 IFOV within focal planes
- Geolocation
  - Present performance is ±50 m 1-σ (Terra)

The MODIS Land Cover Product

MOD12Q1: Global Land Cover
MOD12Q2: Land Cover Dynamics

A.H. Strahler (PI), Mark Friedl, Xiaoyang Zhang, and John Hodges

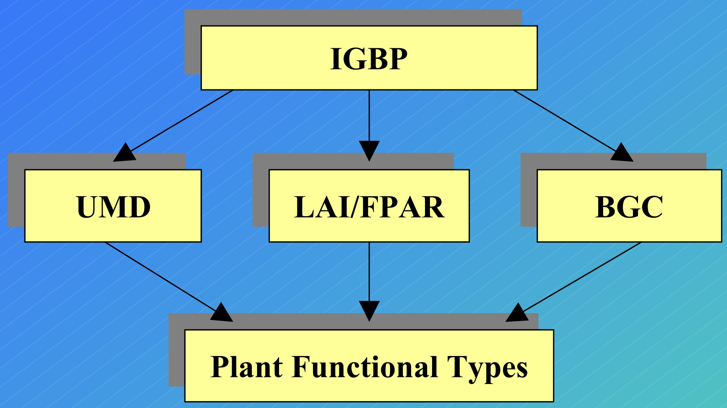
http://geography.bu.edu/landcover/

Center for Remote Sensing and Dept. of Geography
Boston University

MOD12Q1: What Is It?

- *A global database of land cover type classes*
  - Prepared at 1-km spatial resolution from MODIS data
  - Remade at 3-month intervals using data from the preceding 12 months
  - Includes 5 different sets of land cover labels, label confidences, and quality assessment information
  - Available in coarser resolutions of 1/20°, 1/4°

MOD12Q1: What Is It?



- *IGBP:International Geosphere-Biosphere Project labels*
  - 17 classes of vegetation life-form
- *UMD: University of Maryland land cover class labels*
  - 14 classes without mosaic classes
- *LAI/FPAR: Classes for LAI/FPAR Production*
  - 6 labels including broadleaf and cereal crops
- *BGC: Biome BGC Model Classes*
  - 6 labels: leaf type, leaf longevity, plant persistence

- *Plant Functional Types (Future)*
  - Plant functional types to be used with the community land model (NCAR, Bonan)
  - Exact classes TBD

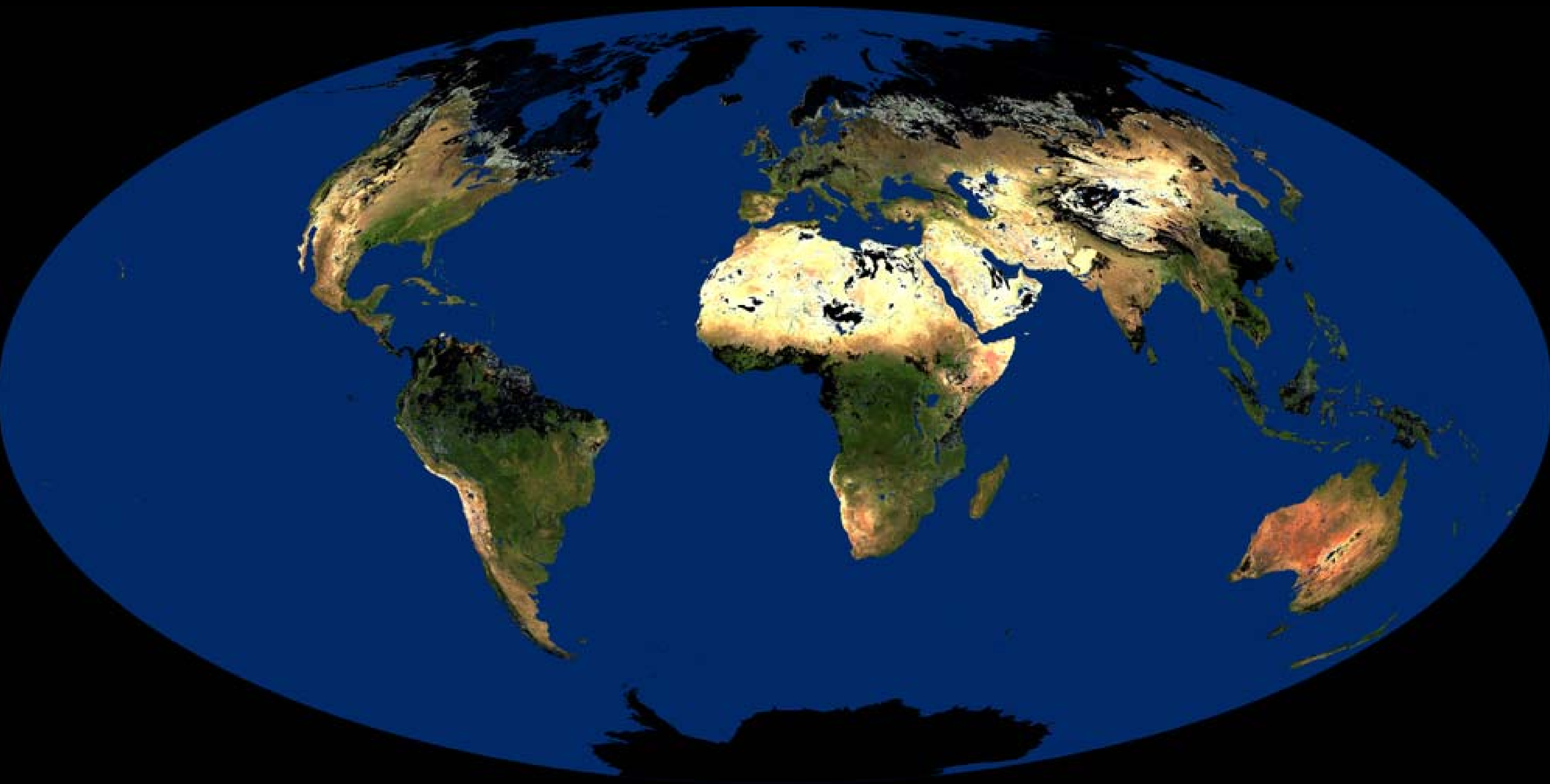
MOD12Q1: What Is It?

- *Land Cover Types*
  - **IGBP, UMD, LAI/FPAR, BGC, PFT (Future)**
- *Confidences*
  - **Classification confidence (percent scale) for each pixel for each label**
- *Secondary IGBP Label and Confidence*
  - **For IGBP, a secondary class label and confidence value for each pixel**
- *Quality Assurance Flag Word for each Pixel*
  - **Includes quality for pixel, last update, and embedded land/water mask**

MOD12Q1: Where Does it Come From?

- *MODIS Data*
  - 16-day Nadir BRDF-Adjusted Reflectances (NBARs) assembled over one year of observations
  - 7 spectral bands, 0.4–2.4 μm, similar to Landsat
  - 16-day Enhanced Vegetation Index (EVI)
- *Training Data*
  - >1,500 training sites delineated from high resolution satellite imagery (largely Landsat)
- *Classifier*
  - Uses decision tree classifier with boosting

Global Composite Map of Nadir BRDF-Adjusted Reflectance (NBAR)
April 7–22 2001



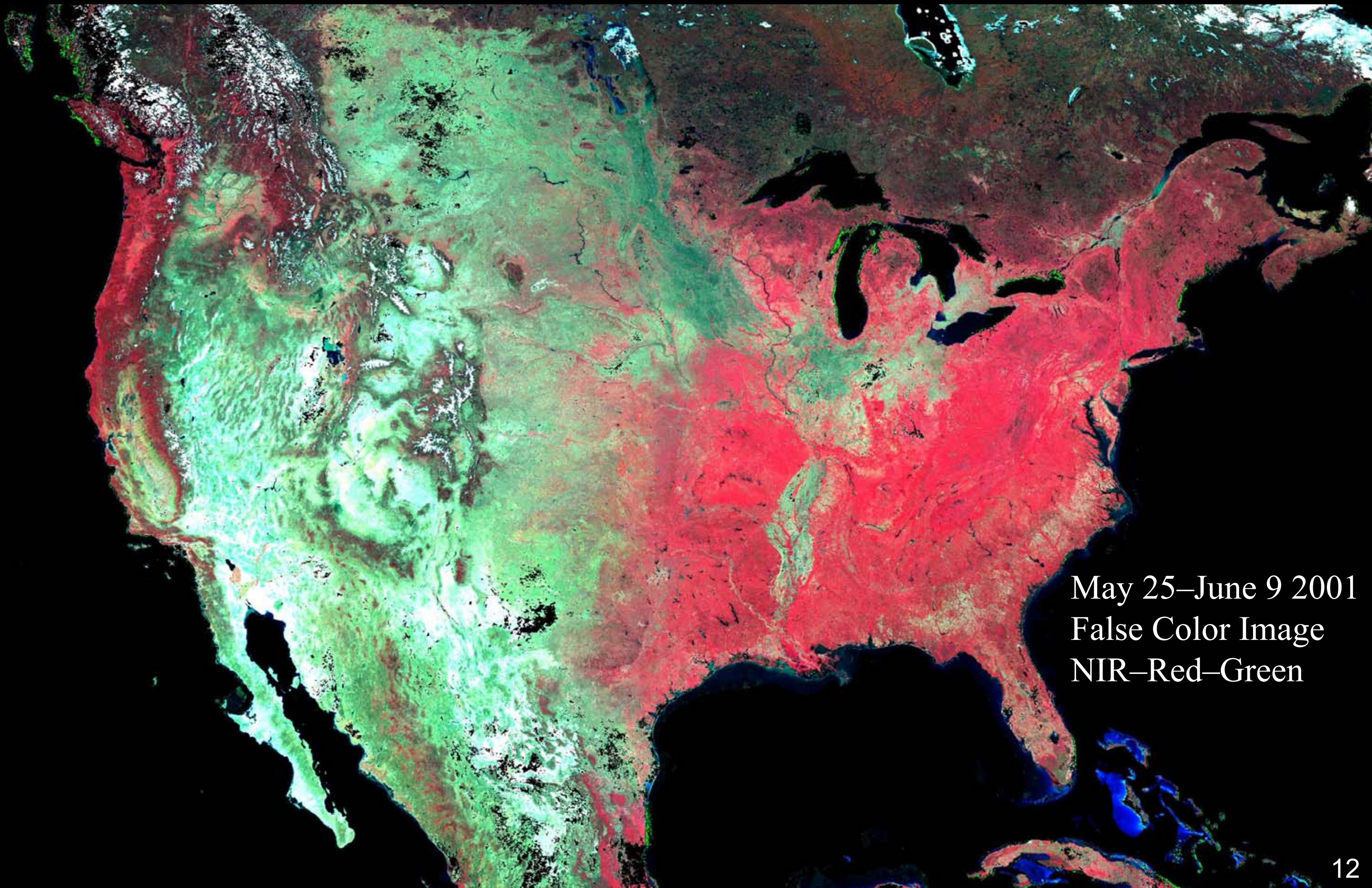
No data

True color, MODIS Bands 2, 4, 3

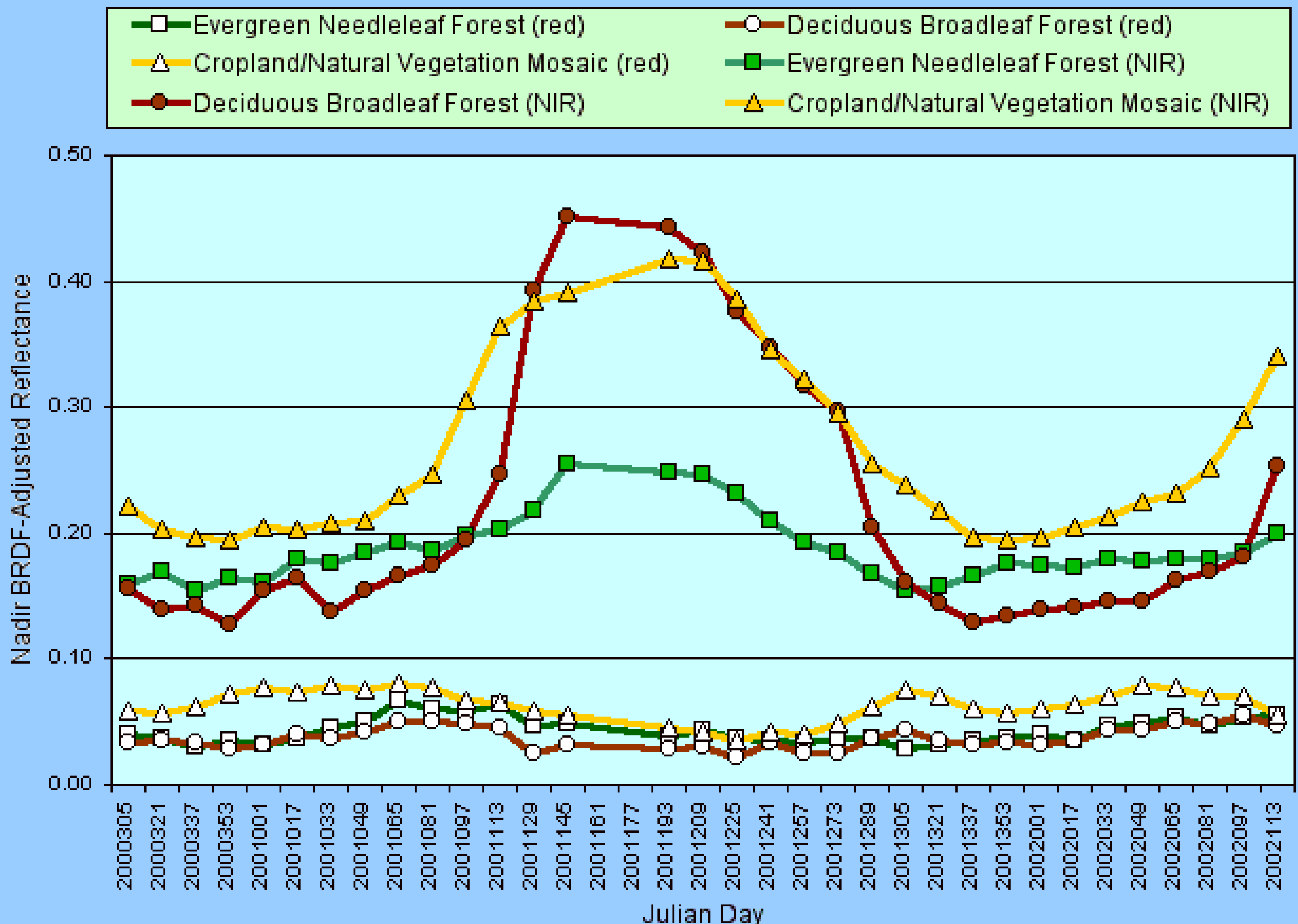
10 km resolution, Hammer-Aitoff projection, produced by MODIS BRDF/Albedo Team

MODLAND/Strahler et al.

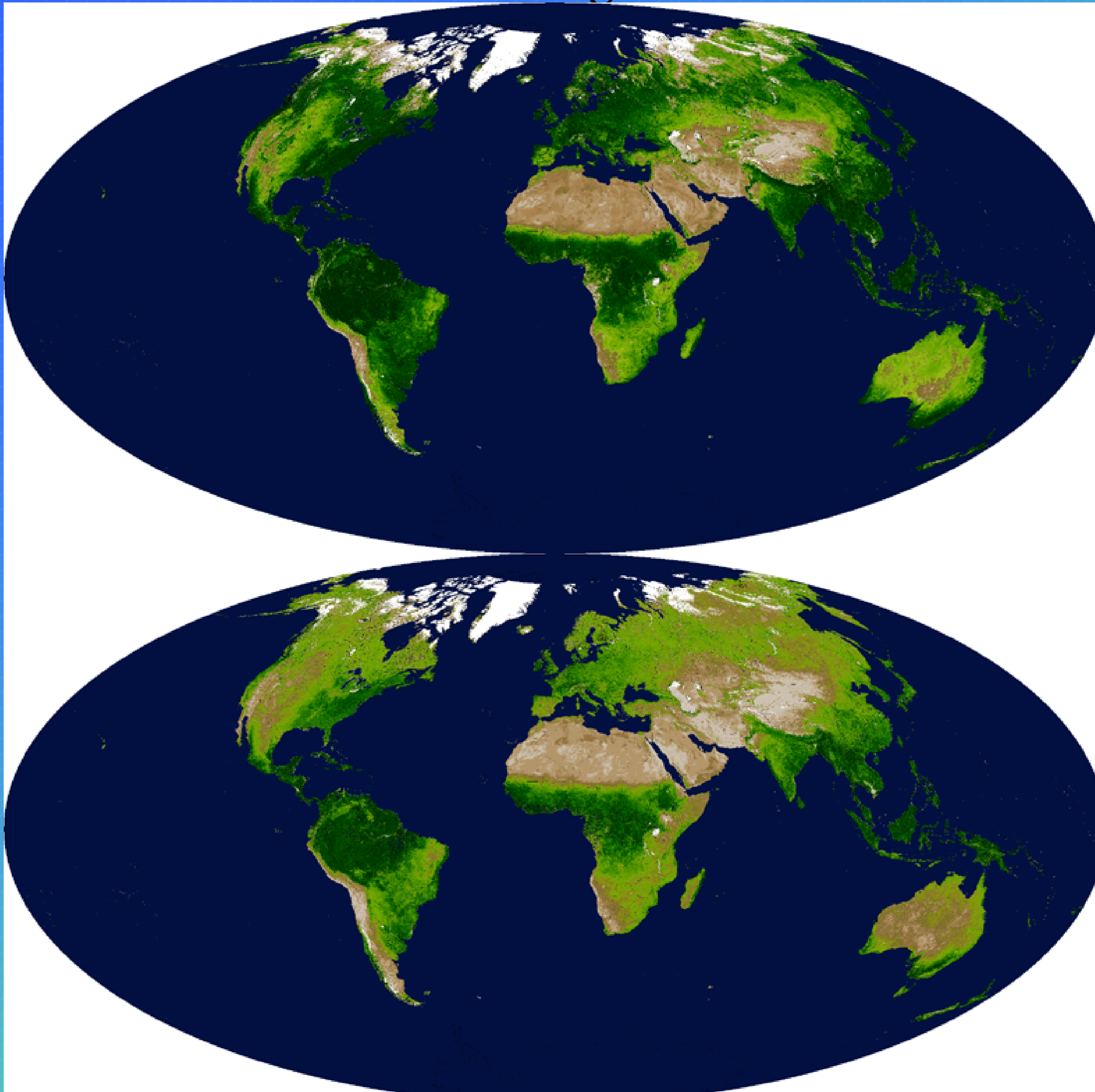
MODIS Nadir BRDF-Adjusted Reflectance



NBAR Time Trajectories

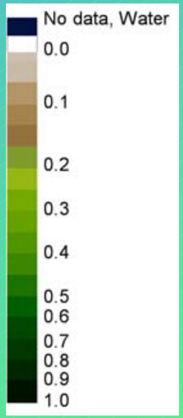


MODIS 500 m Vegetation Indices



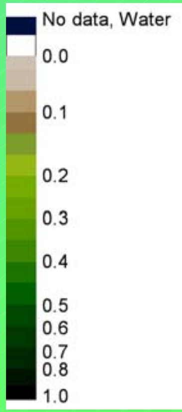
September 30 –
October 15, 2000

NDVI



MOD13A1 16 day Composite

EVI



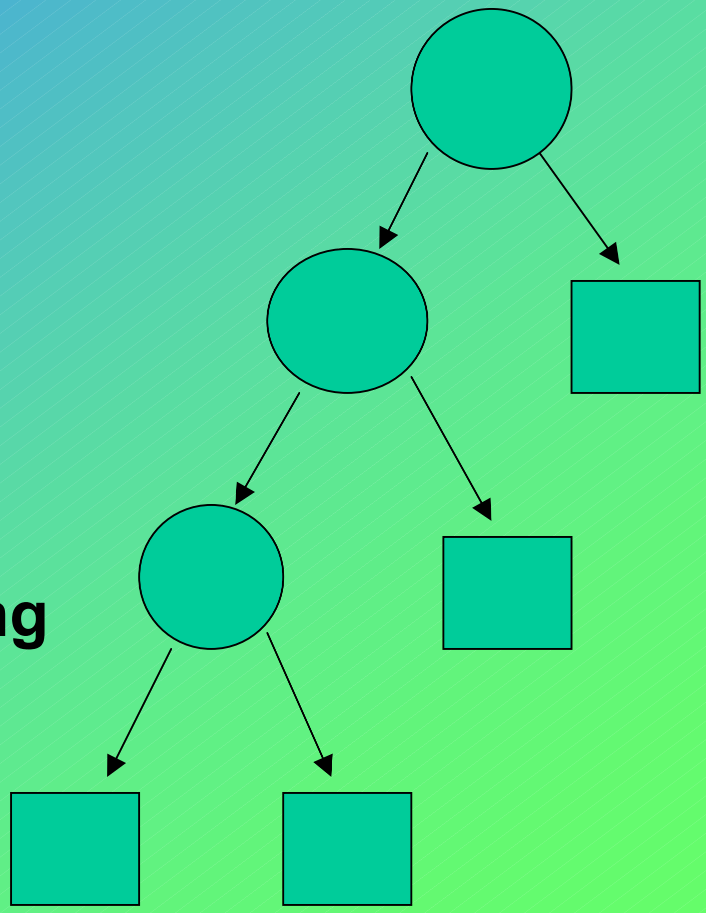
NDVI

EVI

EVI shows better dynamic range, less saturation

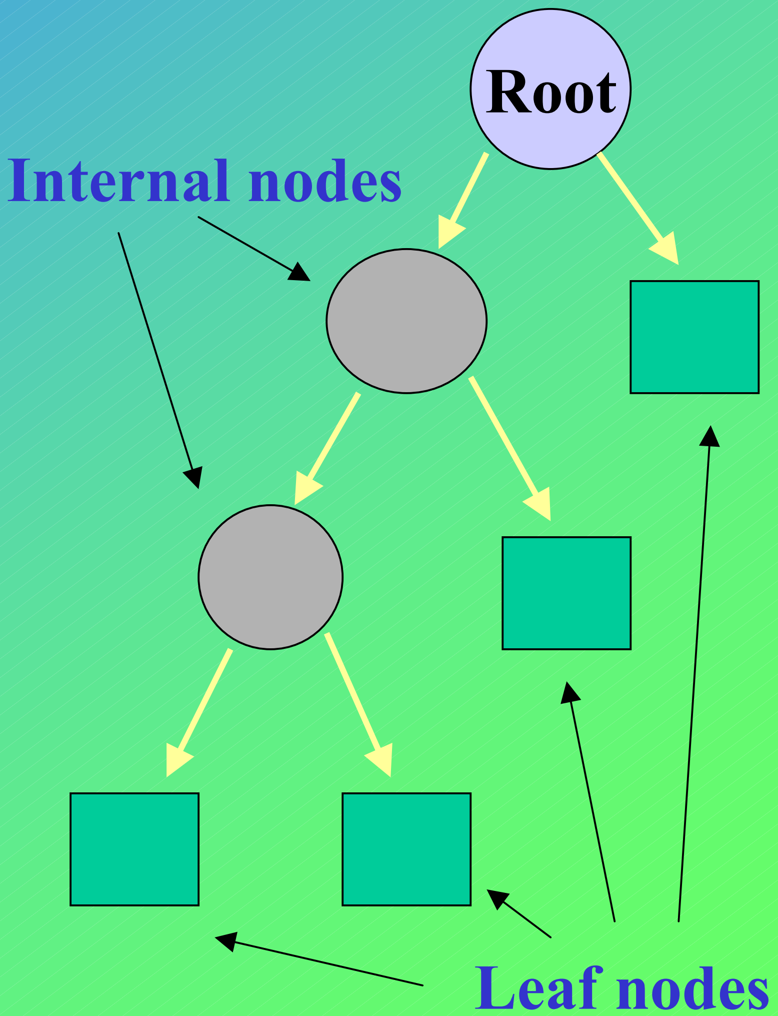
Decision Tree Classification

- *Goal:*
  - Optimal prediction of class labels from a set of feature values
- *Basic approach*
  - Supervised learning using training data
- *Key attributes:*
  - Nonparametric
  - Able to handle noisy or missing features
  - Adept at capturing non-linear, hierarchical patterns

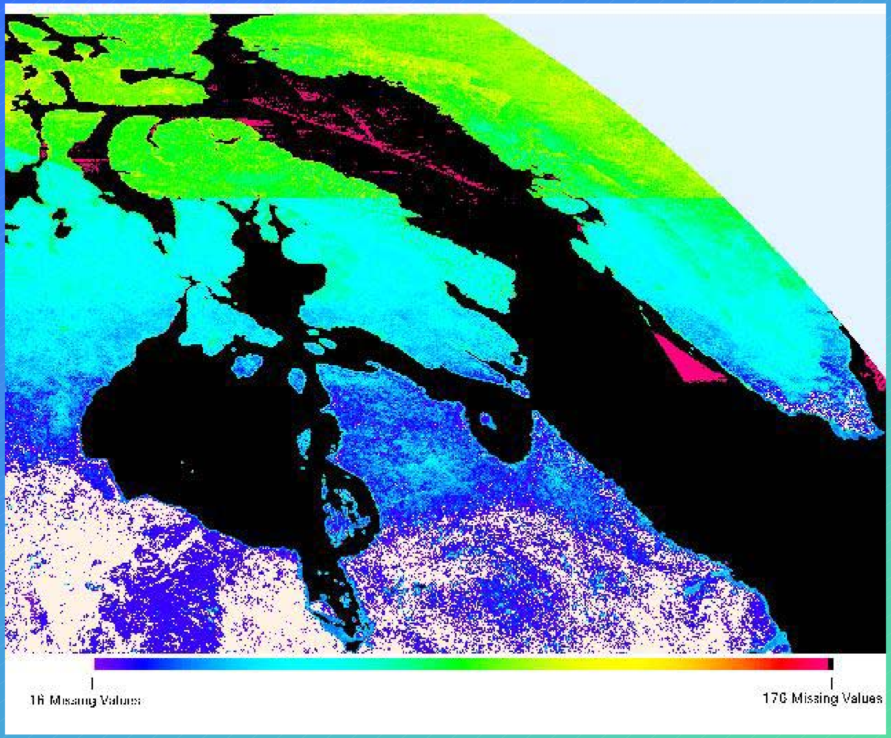


DTs: Basic Theory

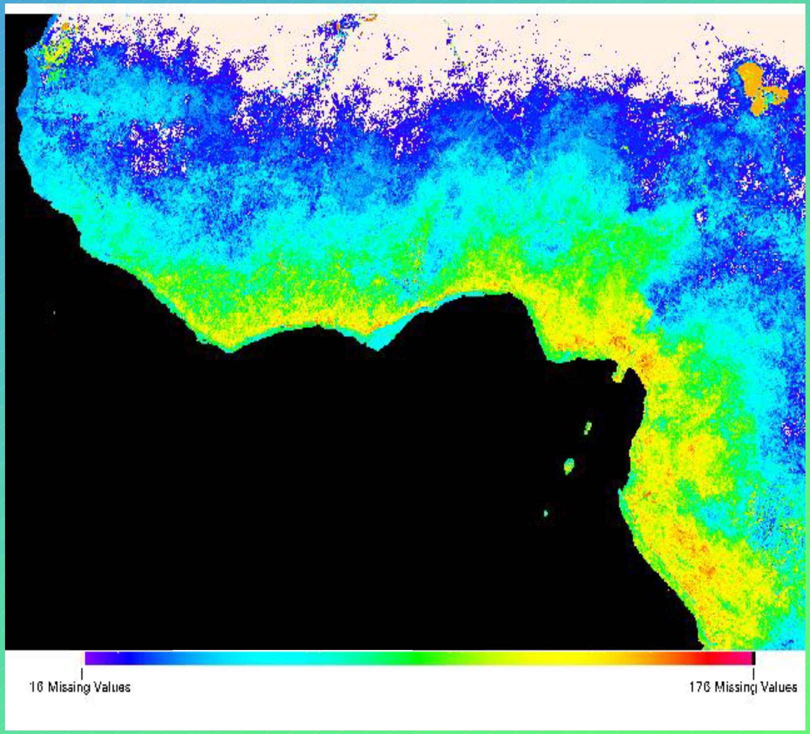
- *Terminology*
  - Root node (all data), internal nodes and terminal or leaf nodes (predictions)
- *DT Estimation:*
  - Recursive partitioning of training data into successively more homogeneous subsets
- *Multiple leaf nodes per class*
  - Leaf nodes identify class assignment
  - Sub-classes allocated individual leaves



Key Advantage of DTs: Ability to Handle Noisy and Missing Data



High Latitudes



Cloud Cover in Tropics

Color scale indicates number of missing values

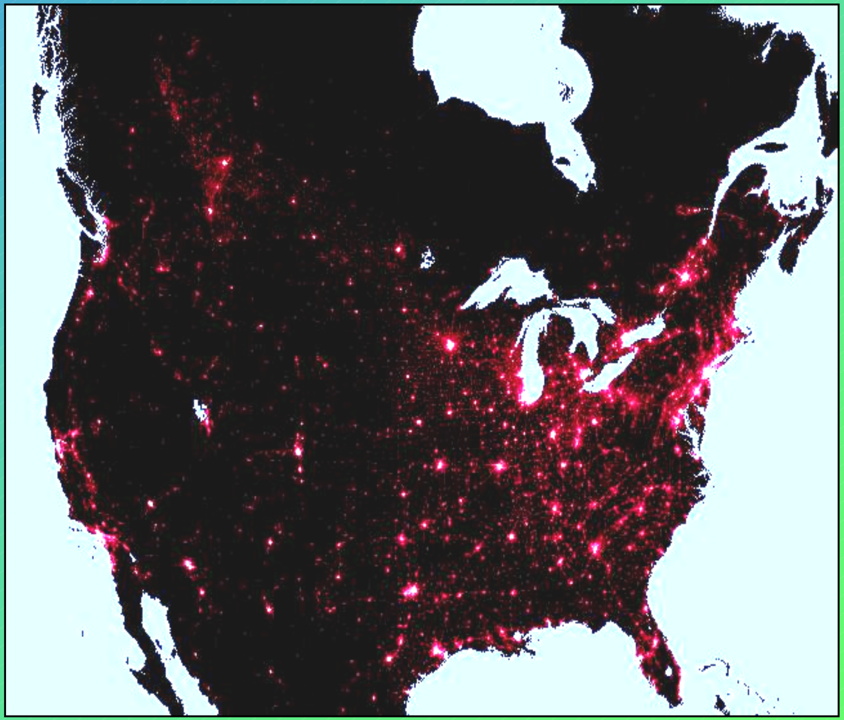
16 176

Postclassification Processing

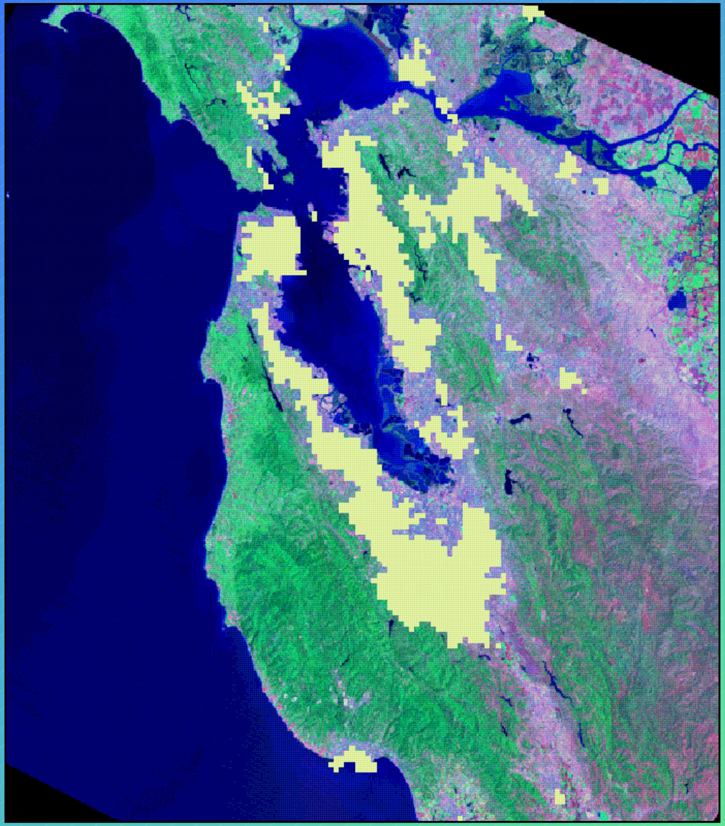
- *Application of Prior Probabilities*
  - Use of priors to remove training site count biases (sample equalization)
  - Application of global and moving-window priors from earlier products
    - Increases accuracies, reduces speckle
  - Use of external maps of prior probabilities to resolve confusions
    - Agriculture/natural vegetation confusion in some regions
    - Use of city lights DMSP data to enhance urban class accuracy (to come)
- *Filling of Cloud-Covered Pixels from Earlier Maps*
  - Use of at-launch (EDC DISCover v. 2) or provisional product when there are not sufficient values to classify a pixel with confidence

Urban Class Mapping

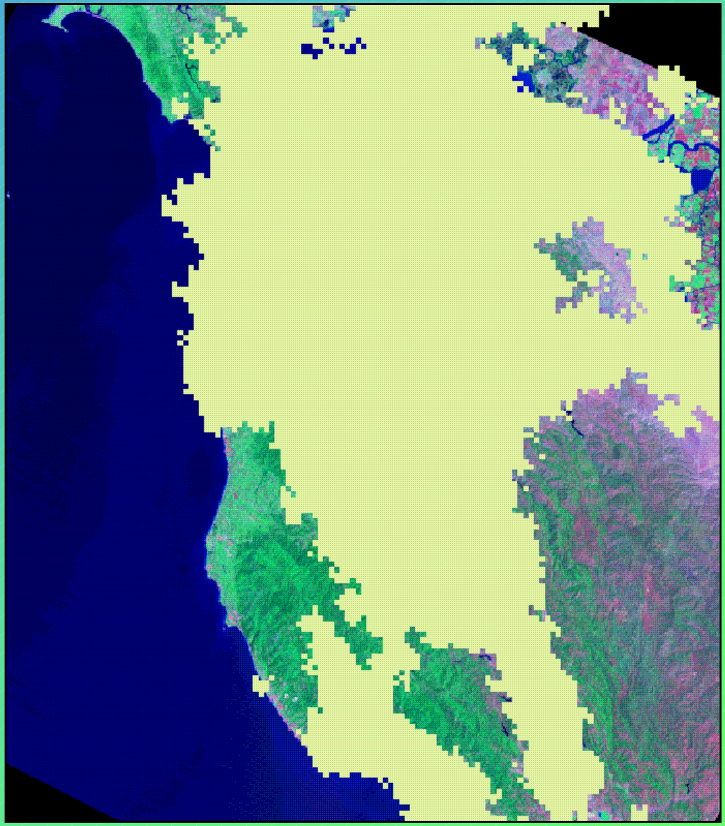
- *Small but important class*
- *Features do not distinguish urban areas*
  - Confusion with barren
- *Solution:*
  - Exploit DMSP OLS data



DMSP-OLS Data Greatly Overestimates Urban Areas

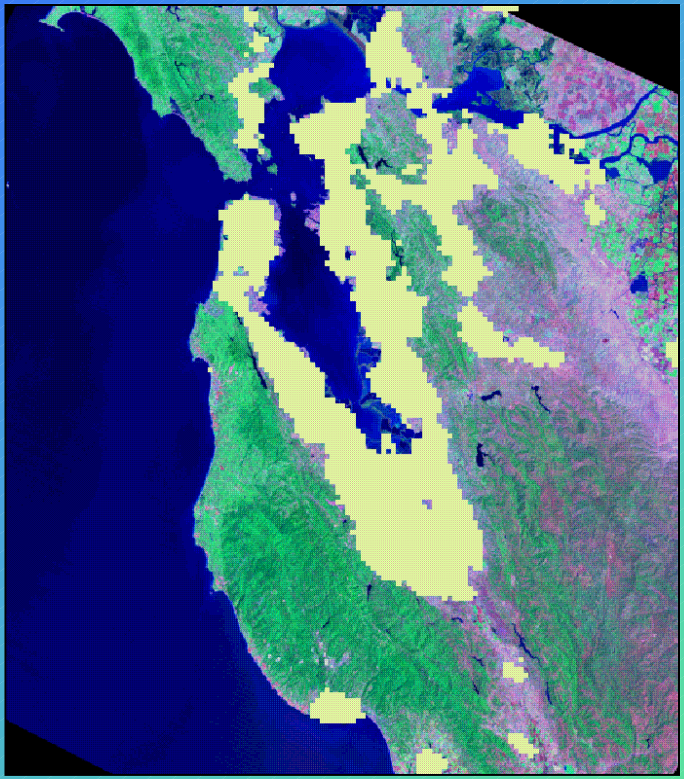


Digital chart of the world urban layer

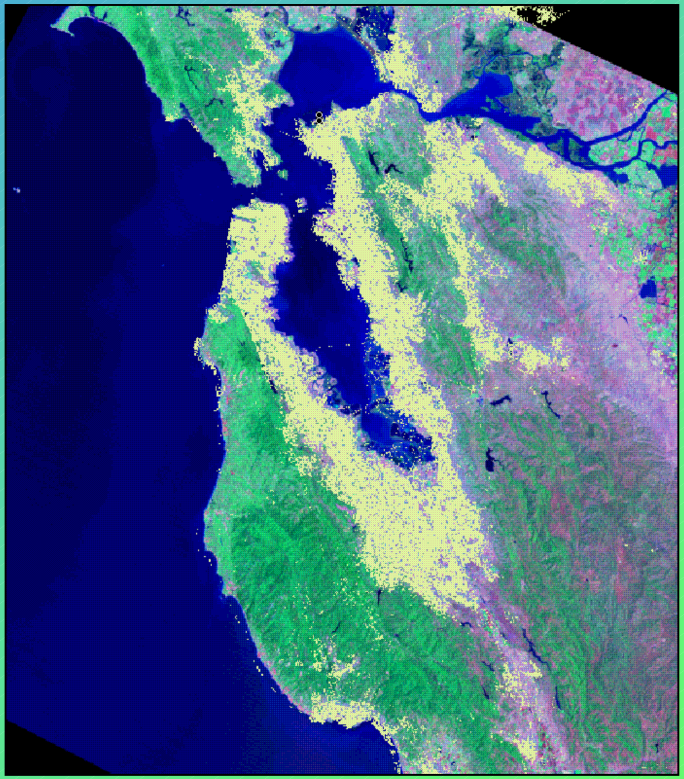


Thresholded DMSP data

Solution:Train Separate Logistic Regression Using DMSP Data & Use Results for Priors



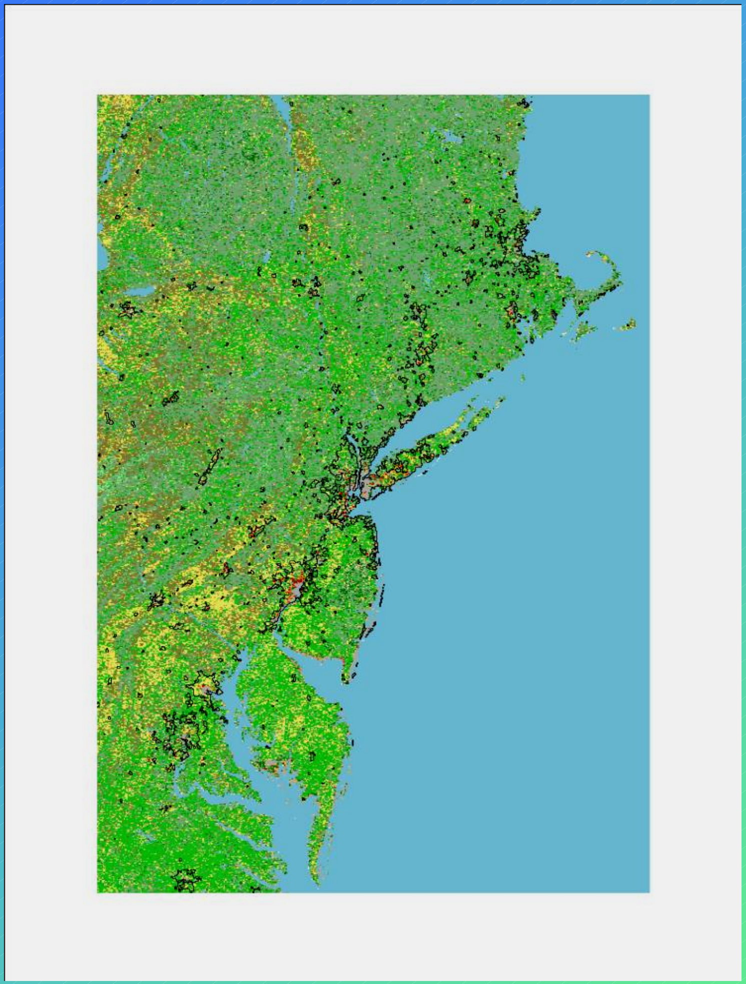
MODIS Classification
(draped on TM image)



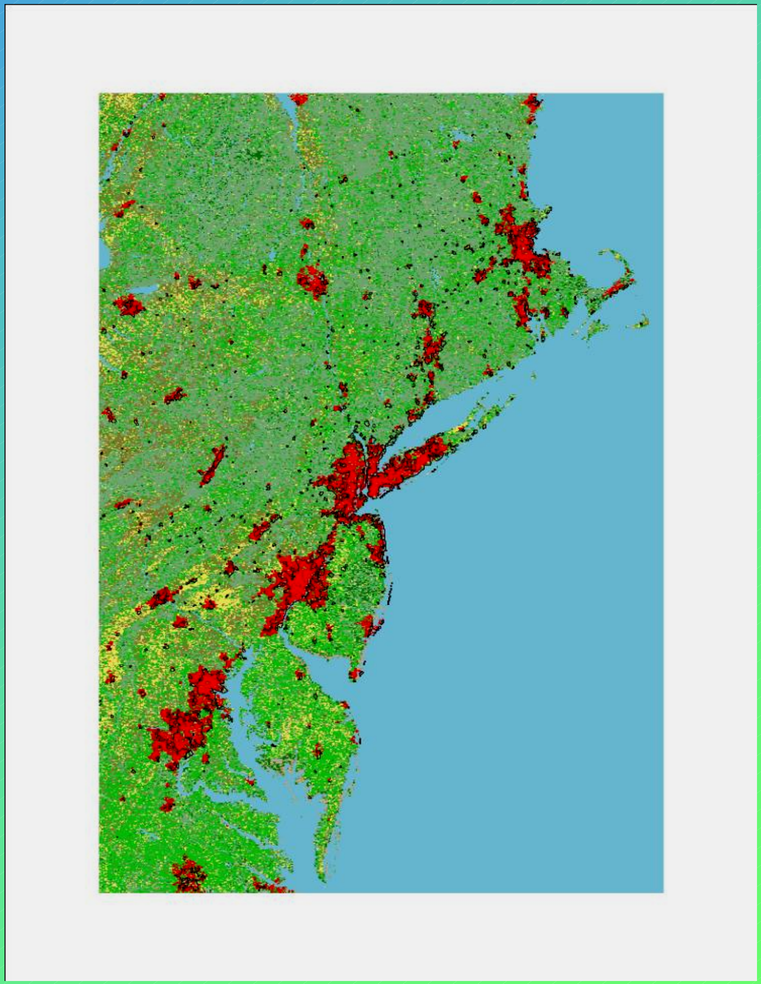
NLCD classification from
TM data

Mapping Urban Areas –
Merging City Lights Data with MODIS

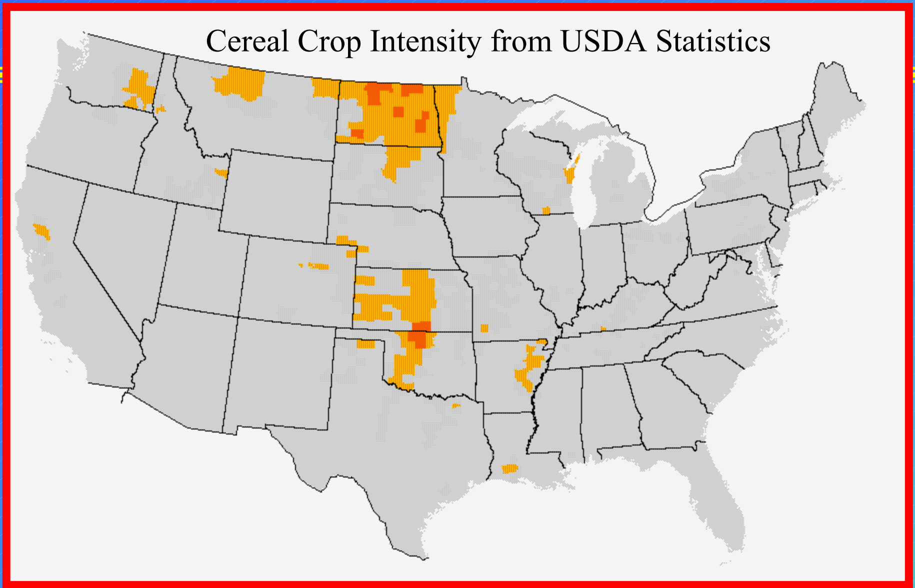
- ***MODIS Data Only***

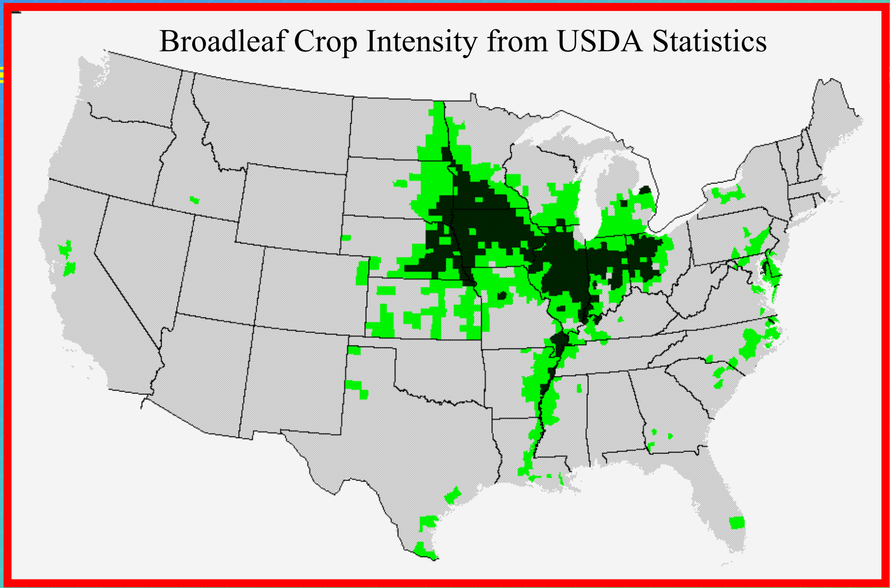


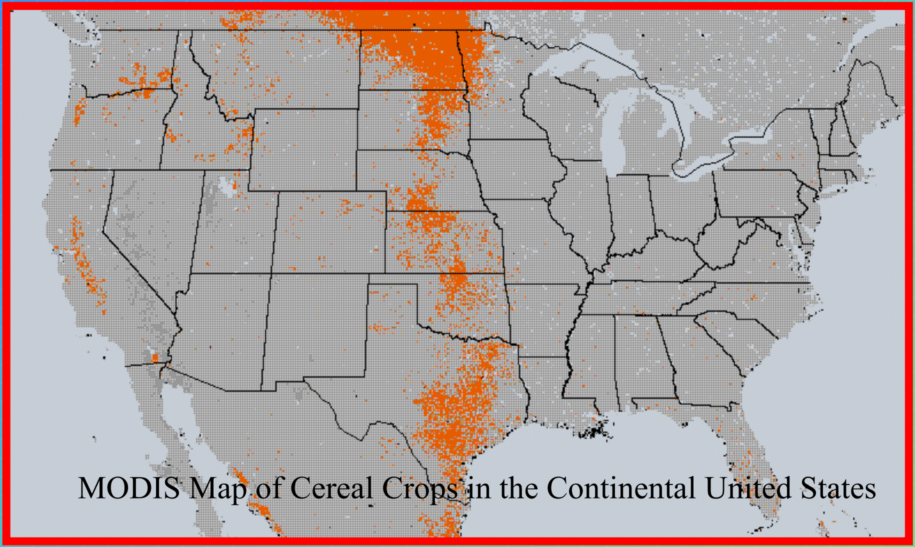
- ***MODIS + City Lights***

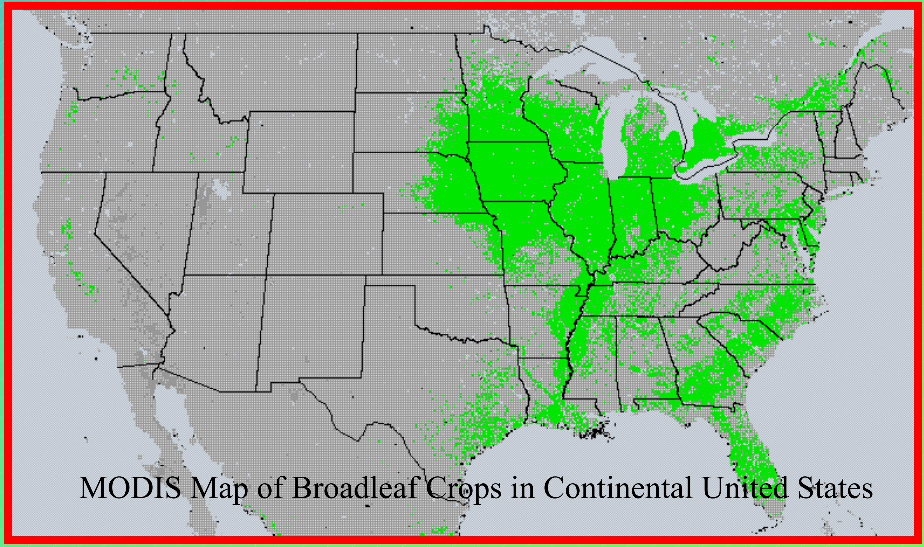


Using Priors to Classify Cereal and Broadleaf Crops





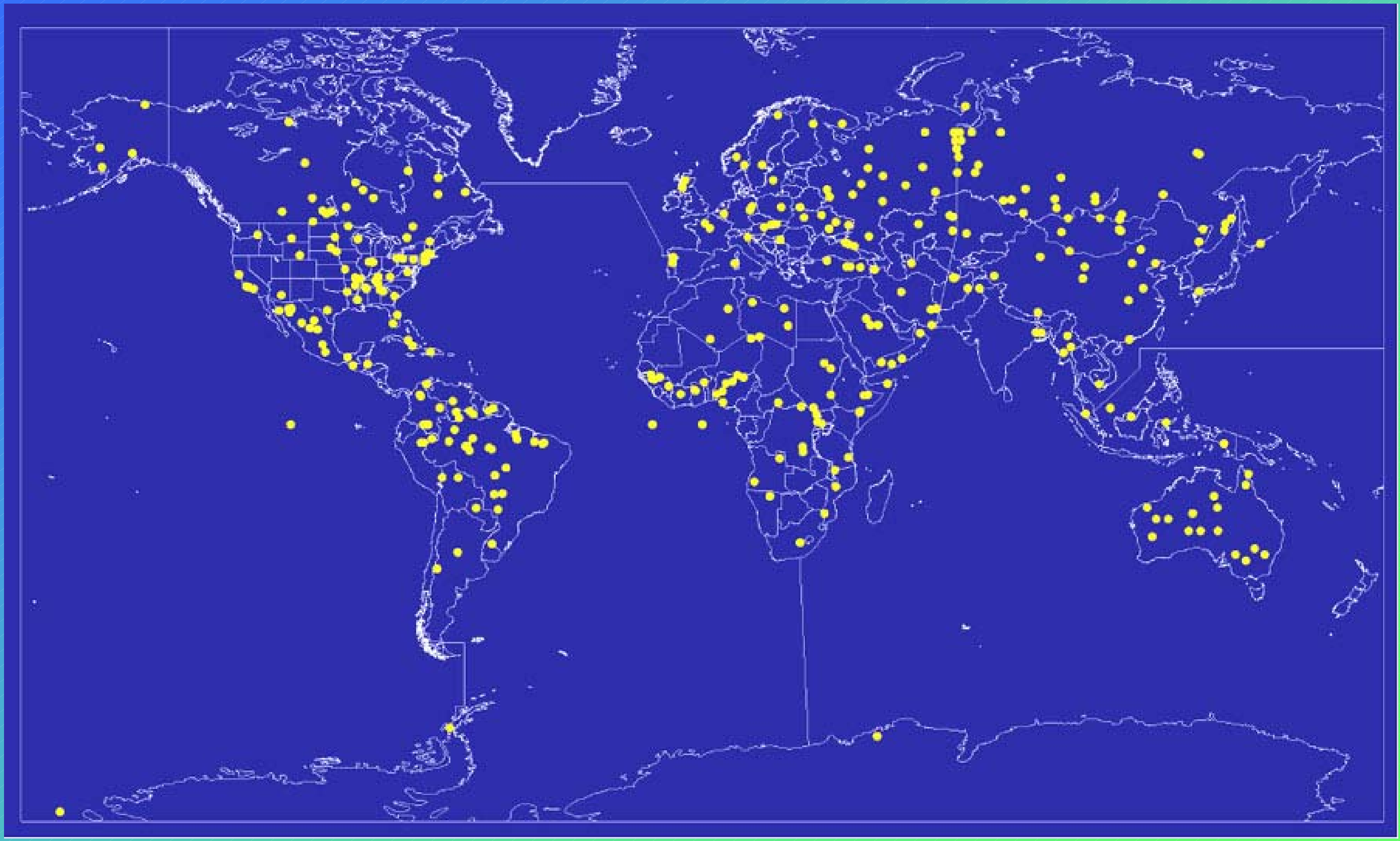




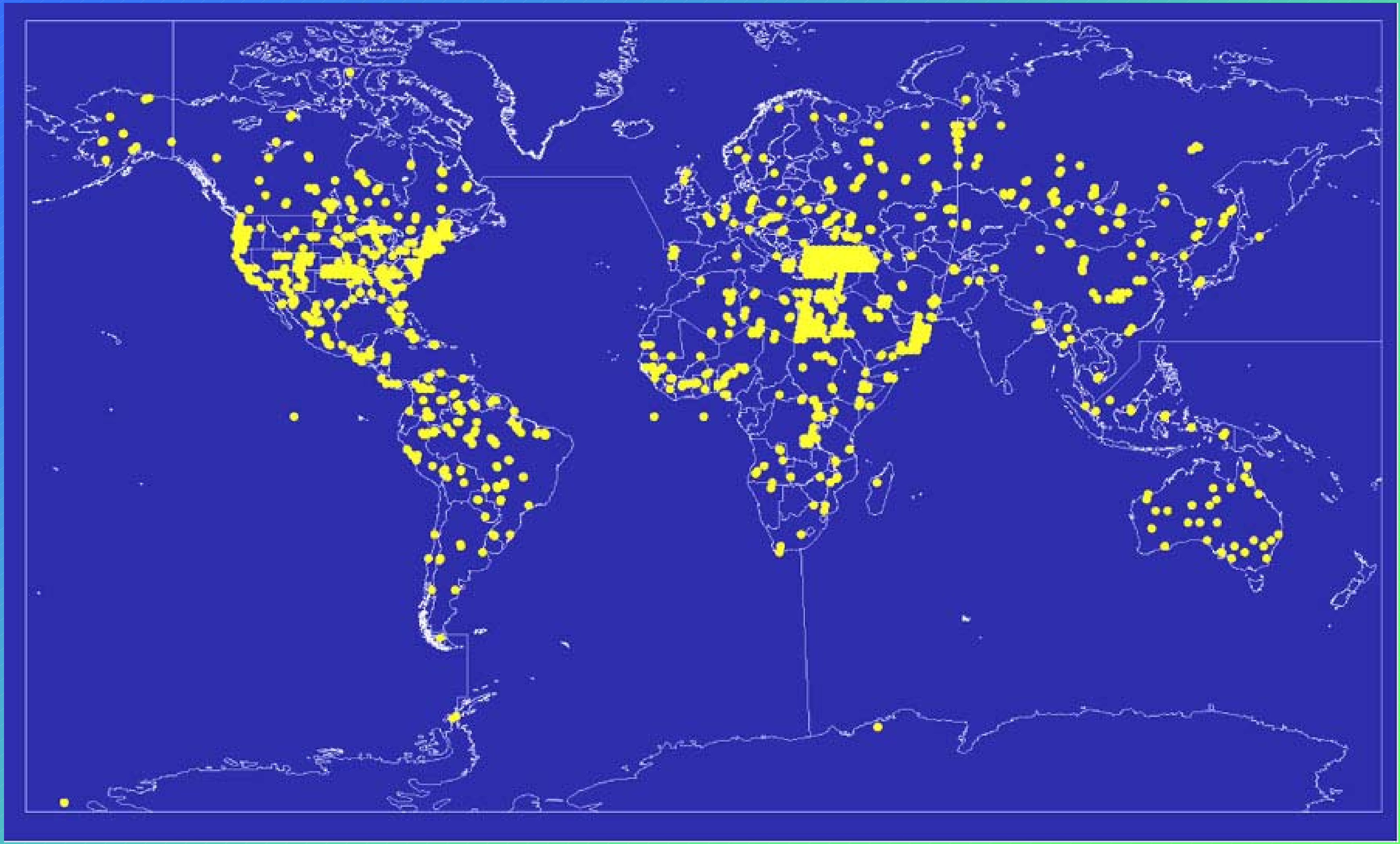
Training Sites—STEP Database

- *STEP:*
  - **System for Terrestrial Ecosystem Parameterization**
- *Key STEP Parameters*
  - **Life form, height, cover fraction, of layers**
  - **Leaf type, phenology, periodicity, physiognomy of dominants in layers**
  - **Elevation, moisture regime, perturbation**
  - **Classifications: IGBP, BU, EDC SLCRs, and others**
  - **Simple description of site and type (words)**
- *STEP Flexibility*
  - **Allows application of many different land cover labeling schemes by inference of label from parameters in database**

DISCover Core Validation Sites



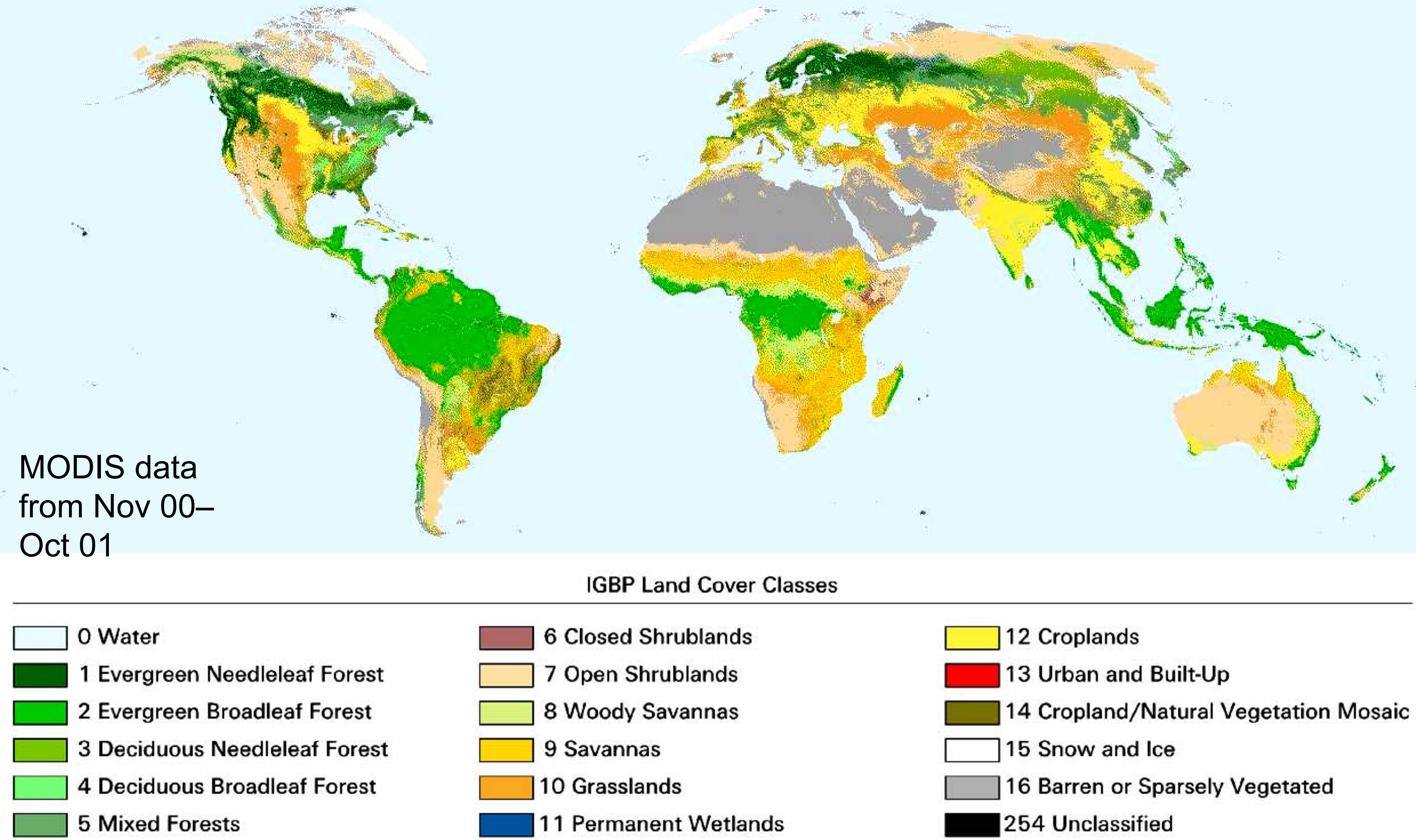
Supplemental BU Training Sites



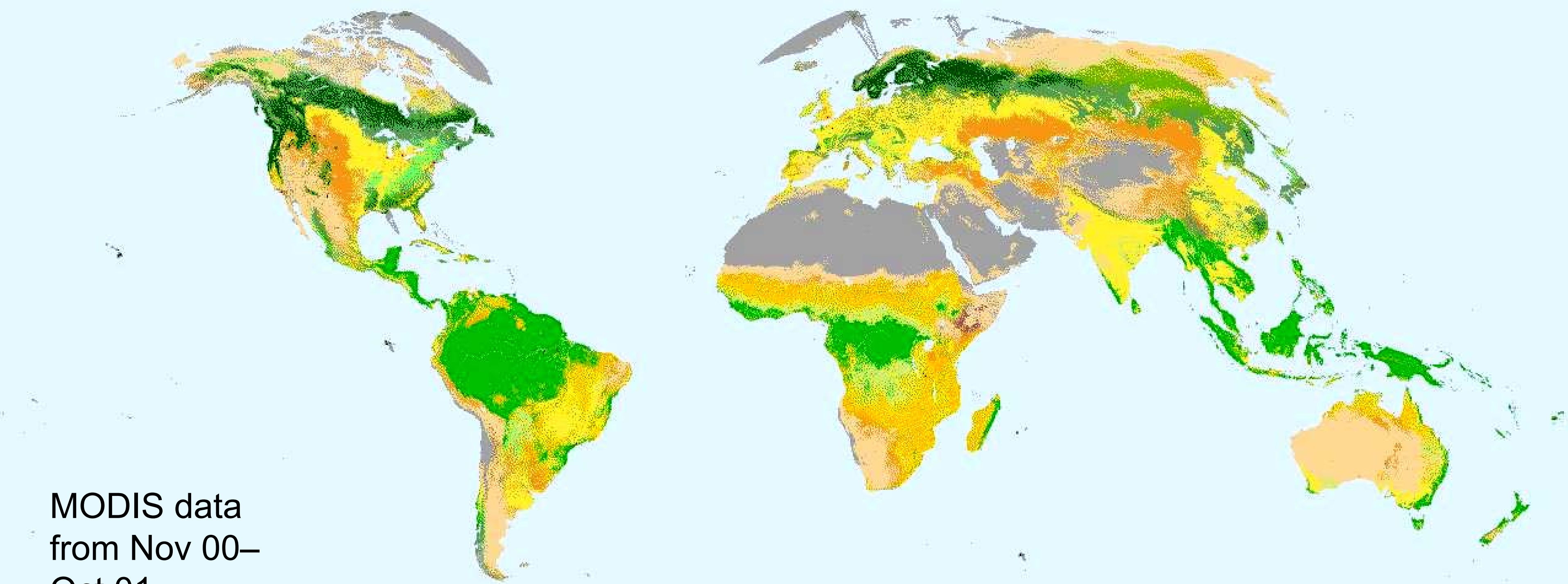
Recent Global Land Cover Products, Cont.

- *First “Consistent Year” Land Cover Product*
  - Completed June 1, 2002
  - Based on MODIS v003 data from Nov 2000–Oct 2001 (consistent year of reprocessed data)
  - Assigned “Validated Levels 1 and 2” status

Consistent Year Land Cover Product
June 02—IGBP



Consistent Year Land Cover Product
June 02—UMd

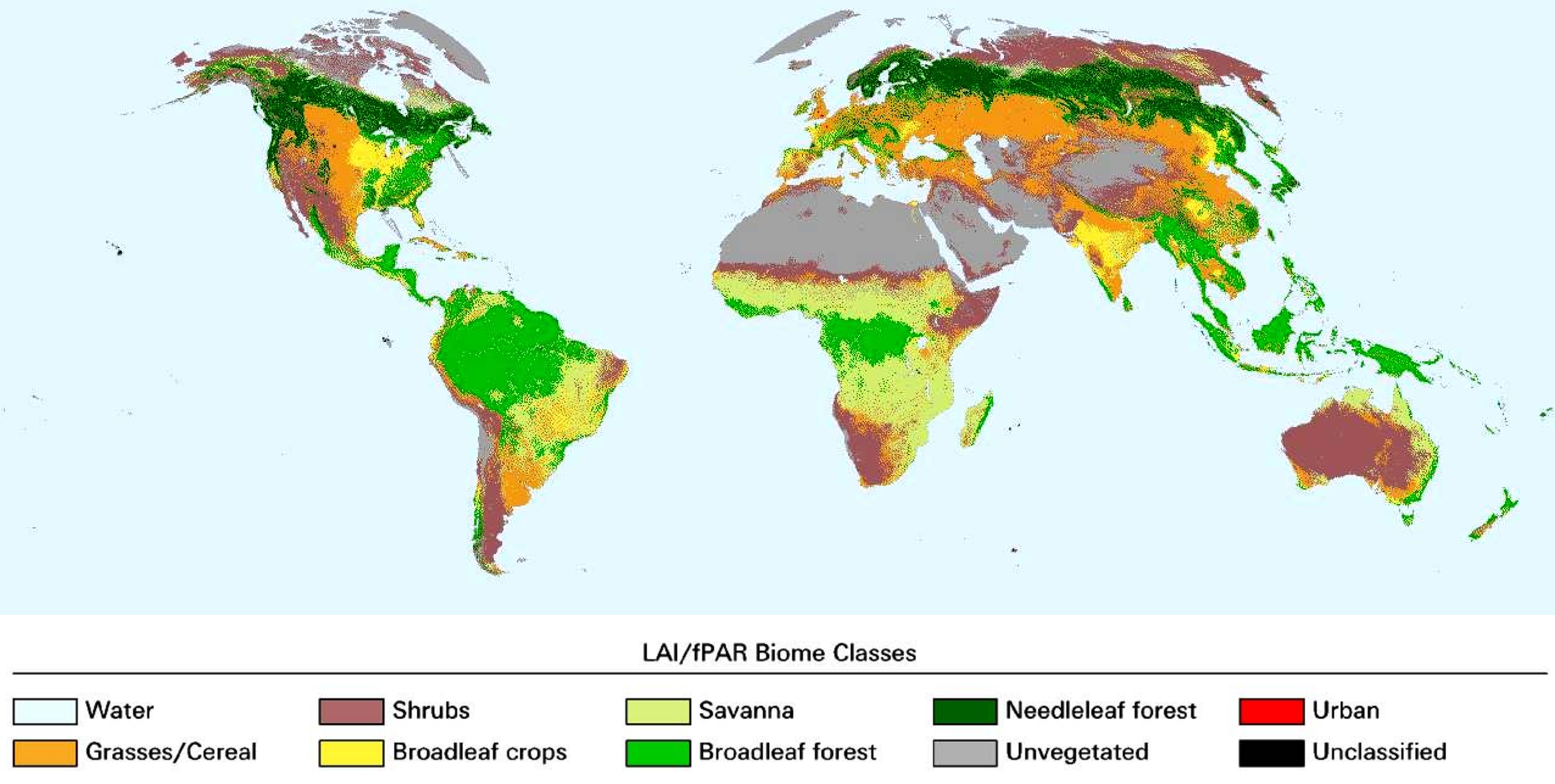


MODIS data from Nov 00–Oct 01

UMD Land Cover Classes

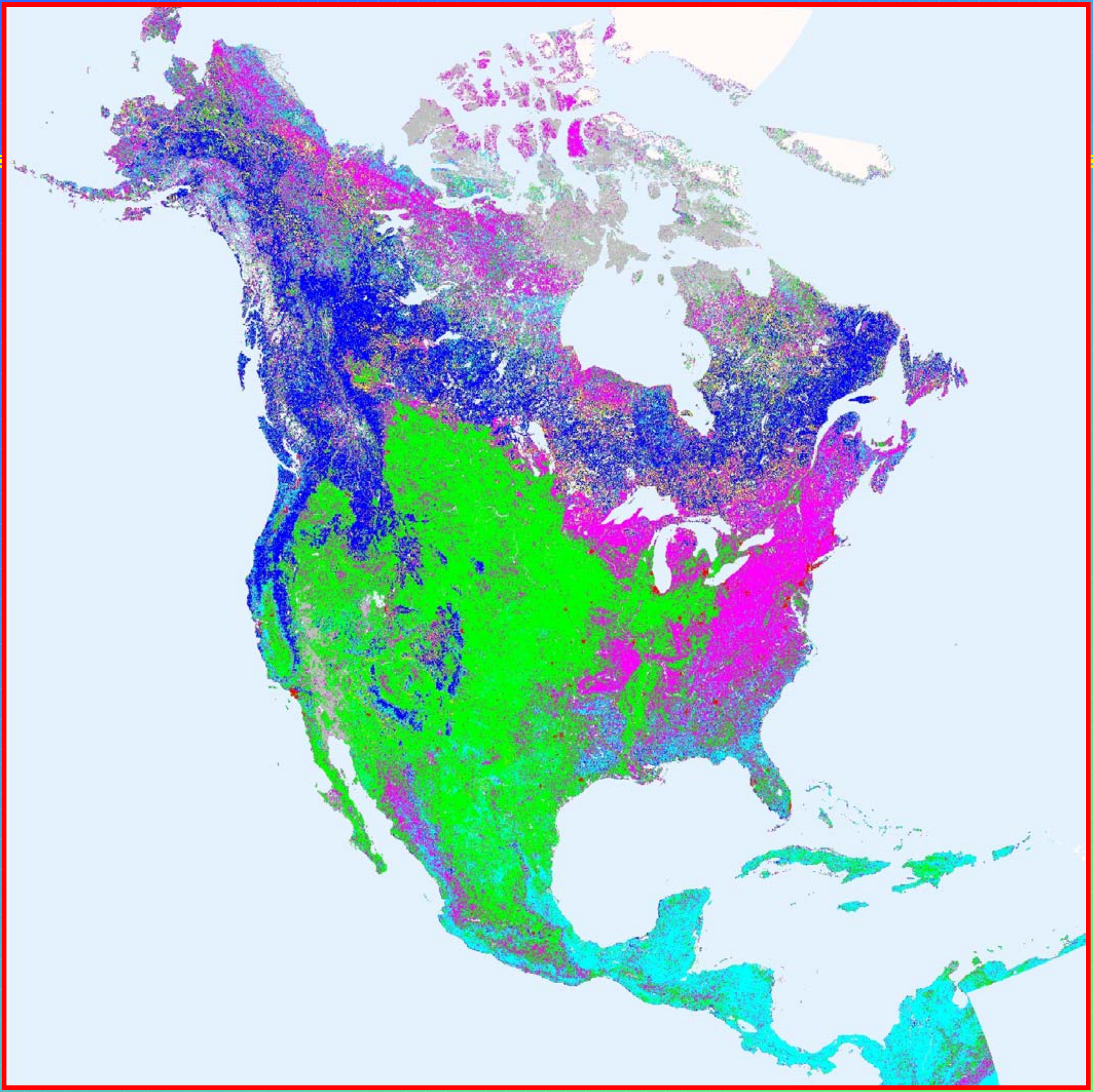
- 0 Water
- 1 Evergreen Needleleaf Forest
- 2 Evergreen Broadleaf Forest
- 3 Deciduous Needleleaf Forest
- 4 Deciduous Broadleaf Forest
- 5 Mixed Forests
- 6 Closed Shrublands
- 7 Open Shrublands
- 8 Woody Savannas
- 9 Savannas
- 10 Grasslands
- 12 Croplands
- 13 Urban and Built-Up
- 16 Barren or Sparsely Vegetated
- 254 Unclassified

Consistent Year Land Cover Product
June 02—LAI/FPAR Biomes

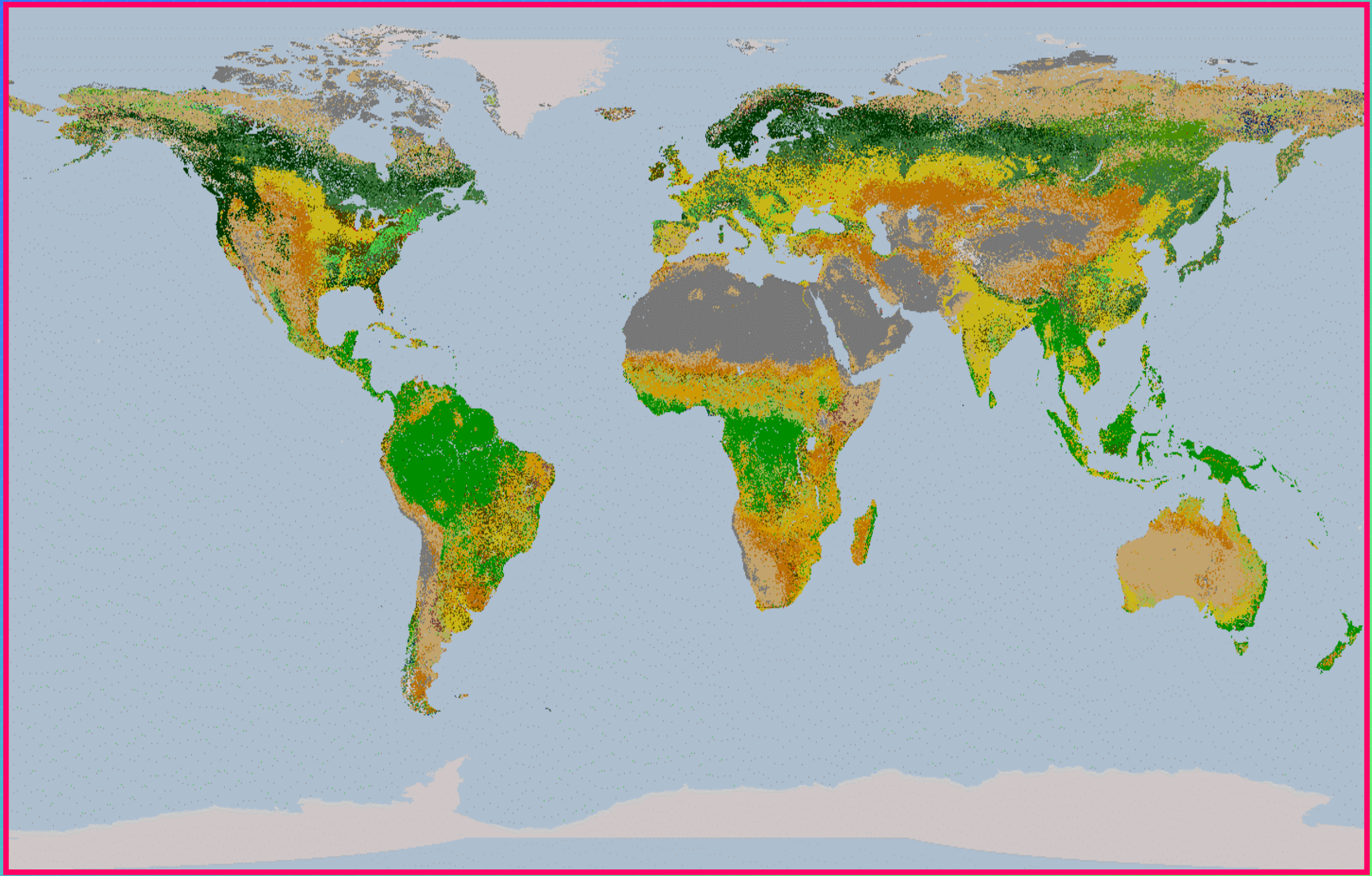


Biome-BGC Labels

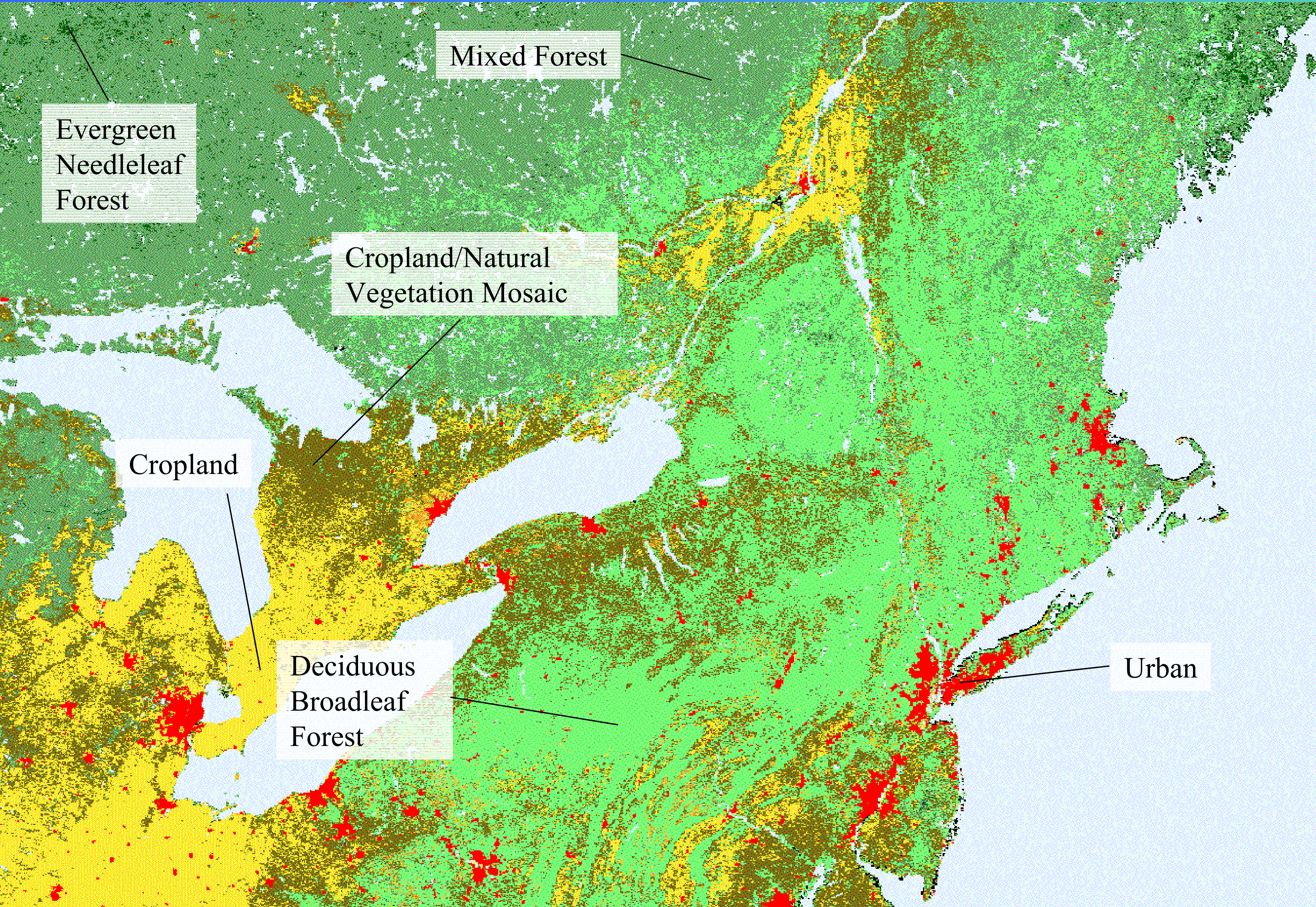
(Provisional Product)



Climate Modeling Grid (CMG)

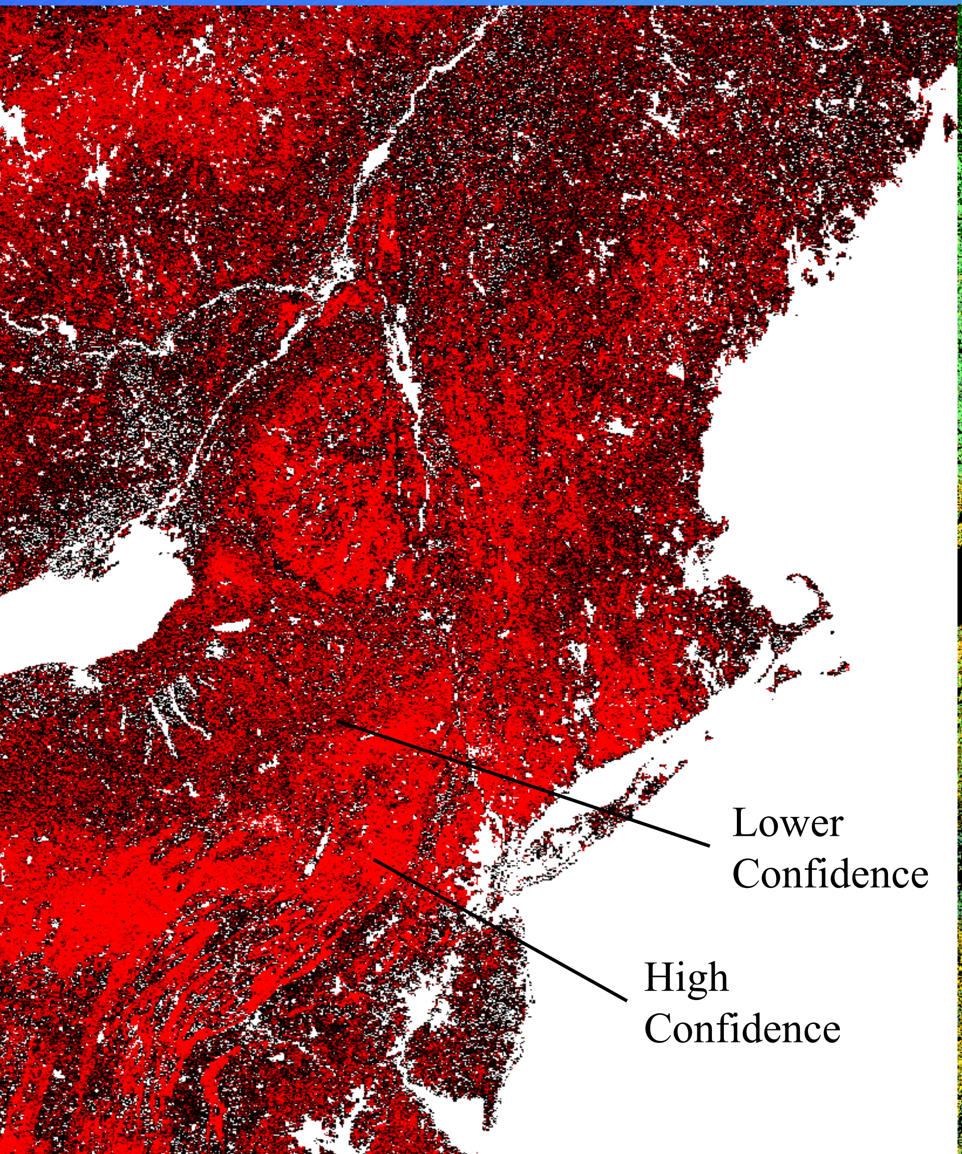


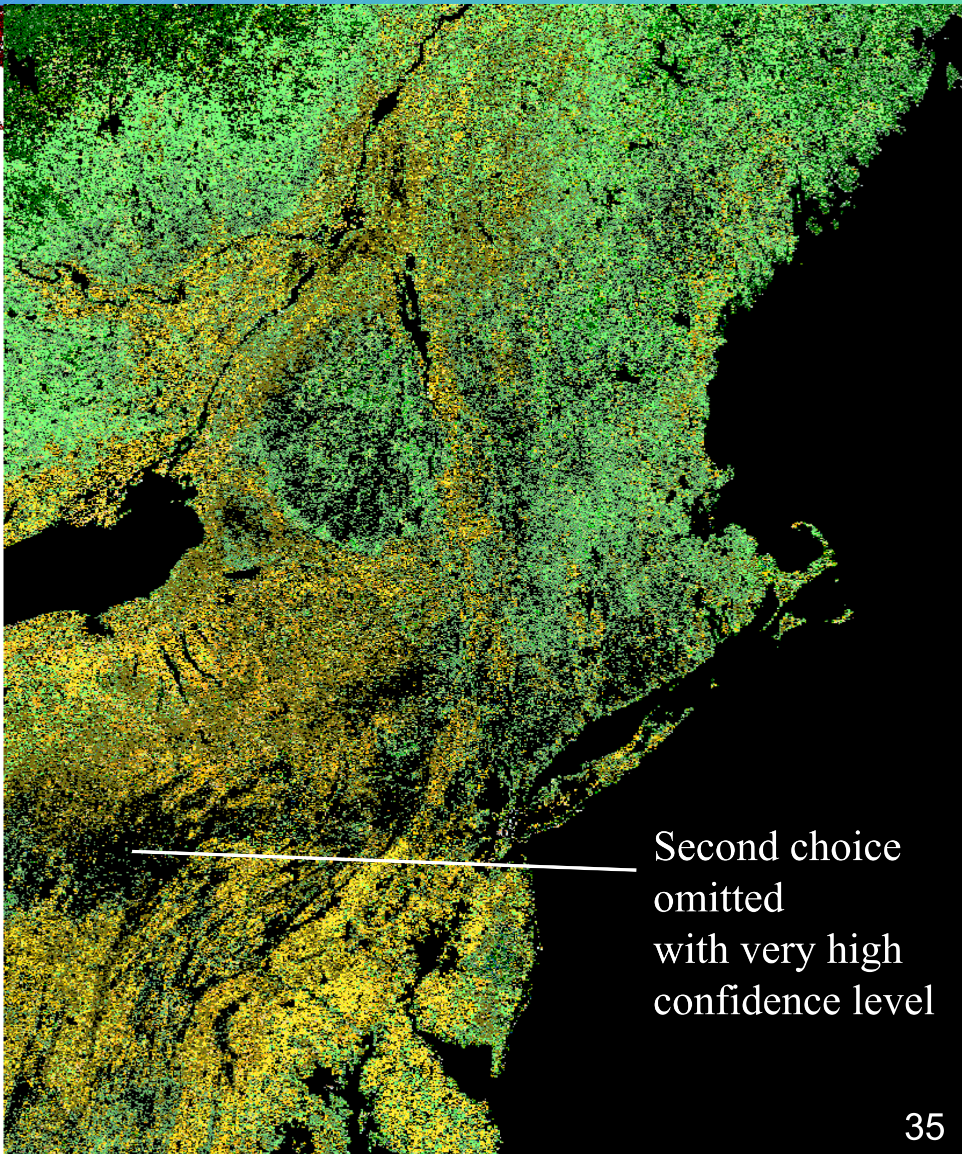
Consistent Year Land Cover Product, Nov 00–Oct 01



Classification Confidence Map

Second Most-Likely Class





Land Cover Validation

- *Validation Plan Utilizes Multiple Approaches*
- *Level 1: Comparisons with existing data sources*
  - Examples
    - Global AVHRR land cover datasets: DISCover, UMd
    - Humid Tropics: Landsat Pathfinder
    - Forest Cover: FAO Forest Resources Assessment
    - Western Europe: CORINE
    - United States: USGS/EPA MLRC
    - United States: California Timber Maps (McIver and Woodcock)
    - MODIS and Bigfoot test site comparisons

Validation Levels, Cont.

- *Level 2: Quantitative studies of output and training data*
  - Per-pixel confidence statistics
    - Aggregated by land cover type and region
    - Describe the accuracy of the classification process
  - Test site cross-comparisons
    - Confusion matrices globally and by region
    - Provides estimates of errors of omission and commission
- *Level 3: Sample-based statistical studies*
  - Random stratified sampling according to proper statistical principles
  - Costly, but needed for making proper accuracy statements
- *CEOS Cal-Val Land Product Validation Land Cover Activity*

Cross Validation with Training Sites

- *Cross-Validation Procedure*
  - Hide 10 percent of training sites, classify with remaining 90 percent; repeat ten times for ten unique sets of all sites
  - Provides “confusion matrix” based on unseen pixels where whole training site is unseen
  - NOT a stratified random sample, but a useful indication of global and within-class accuracies

Confusion Matrix

Site Class	Class Name	Classification Outcome																
		1	2	3	4	5	6	7	8	9	10	11	12	14	15	16	17	Total
1	Evergreen Needleleaf	1323	13	65	23	407	7	35	96	11	6	20	35	7	0	2	6	2056
2	Evergreen Broadleaf	12	5139	0	3	3	2	7	141	48	14	1	18	14	0	3	4	5409
3	Deciduous Needleleaf	20	0	102	3	85	0	5	38	1	3	1	3	0	0	0	0	261
4	Deciduous Broadleaf	7	11	15	381	243	1	10	34	10	9	0	16	11	0	2	8	758
5	Mixed Forest	167	3	50	178	1370	1	9	59	7	29	70	71	52	0	0	11	2077
6	Closed Shrubland	24	18	0	0	6	129	154	37	55	14	0	29	0	0	0	0	466
7	Open Shrubland	4	4	2	17	9	53	1204	27	9	170	3	5	0	1	168	3	1679
8	Woody Savanna	76	56	0	6	61	3	97	617	154	47	0	36	12	0	0	2	1167
9	Savanna	1	53	3	0	4	25	84	303	504	49	7	13	49	0	3	0	1098
10	Grasslands	5	36	0	1	4	1	161	15	69	1028	0	78	20	0	54	2	1474
11	Pmnt Wtlnd	60	15	0	1	7	0	9	9	2	8	174	3	1	0	0	0	289
12	Cropland	23	46	3	33	21	15	243	142	252	365	0	4775	299	0	13	10	6240
14	Cropland/Natural Vegn	2	134	0	195	62	3	9	113	150	29	0	197	546	0	3	4	1447
15	Snow+ice	1	0	0	0	0	0	31	0	0	3	0	2	0	1261	47	1	1346
16	Barren	2	6	0	2	12	38	491	10	10	56	0	9	2	0	3853	1	4492
17	Water	7	5	0	9	11	1	2	0	2	6	0	12	3	0	0	9155	9213
	Total	1734	5539	240	852	2305	279	2551	1641	1284	1836	276	5302	1016	1262	4148	9207	39472

Global and Regional Accuracy

Table 4. Global accuracy and accuracy of continental regions (percent).				
Region	Accuracy Estimate	Standard Error	95% Confidence Interval	
			Low	High
Global	71.6	0.25	71.1	72.1
Africa	61.7	0.66	60.3	63.0
Australia & Insular Asia	71.9	2.93	66.1	77.8
Eurasia	67.8	0.40	67.0	68.6
North America	61.3	0.62	60.0	62.5
South America	75.4	0.46	74.4	76.3

(Analysis follows Card (1982) to correct for bias induced because contingency table marginal proportions don't match global proportions)

Per-Class Accuracies

Table 5. Global per-class accuracies, consistent-year land cover product (percent)

IGBP Land Cover Class	Producer's Accuracy				User's Accuracy				Areal Proportions			
	Est.	Std. Err.	CI -	CI +	Est.	Std. Err.	CI -	CI +	Est.	Std. Err.	CI -	CI +
1. Evergreen Needleleaf	60.0	1.0	58.0	62.0	75.8	1.0	73.8	77.9	4.9	0.1	4.7	5.1
2. Evergreen Broadleaf	90.3	0.5	89.2	91.4	92.7	0.3	92.0	93.4	9.8	0.1	9.7	10.0
3. Deciduous Needleleaf	57.7	2.8	52.2	63.3	42.3	3.2	36.0	48.7	0.9	0.1	0.8	1.1
4. Deciduous Broadleaf	34.0	1.5	31.0	37.1	43.3	1.7	40.0	46.7	1.4	0.1	1.3	1.5
5. Mixed Forest	61.5	1.1	59.4	63.6	58.7	1.0	56.7	60.7	4.3	0.1	4.1	4.4
6. Closed Shrubland	14.2	1.1	12.1	16.3	46.2	3.0	40.3	52.2	1.9	0.1	1.7	2.1
7. Open Shrubland	85.0	0.6	83.7	86.3	46.8	1.0	44.8	48.7	9.6	0.2	9.2	9.9
8. Woody Savanna	51.6	1.4	48.8	54.4	37.5	1.2	35.1	39.8	4.2	0.1	4.0	4.5
9. Savanna	52.4	1.4	49.6	55.1	39.1	1.4	36.4	41.8	4.6	0.1	4.3	4.8
10. Grasslands	66.2	1.2	63.7	68.7	55.3	1.2	53.0	57.6	5.6	0.1	5.4	5.9
11. Permanent Wetlands	37.9	2.7	32.6	43.2	62.6	2.9	56.8	68.4	0.5	0.0	0.4	0.6
12. Cropland	58.1	0.6	56.8	59.4	87.4	0.4	86.5	88.3	14.0	0.2	13.7	14.3
14. Cropland/Nat Veg Mosaic	42.5	1.1	40.2	44.8	53.5	1.6	50.4	56.6	3.9	0.1	3.7	4.1
15. Snow and Ice	96.6	0.4	95.9	97.4	99.9	0.1	99.8	100	10.8	0.0	10.7	10.9
16. Barren/Sparse	74.8	0.7	73.4	76.2	92.8	0.4	92.0	93.6	14.9	0.1	14.6	15.2
17. Water	98.3	0.2	97.9	98.8	99.4	0.1	99.2	99.6	8.0	0.0	8.0	8.1

Confidence Values by Land Cover Type

Table 6. Global confidence values by land cover class (percent)	
IGBP Land Cover Class	Average Confidence Value
1. Evergreen Needleleaf	68.3
2. Evergreen Broadleaf	89.3
3. Deciduous Needleleaf	66.7
4. Deciduous Broadleaf	65.9
5. Mixed Forest	65.4
6. Closed Shrubland	60.0
7. Open Shrubland	75.3
8. Woody Savanna	64.0
9. Savanna	67.8
10. Grasslands	70.6
11. Permanent Wetlands	52.3
12. Cropland	76.4
14. Cropland/Natural Veg	60.7
15. Snow and Ice	87.2
16. Barren	90.0
17. Water	(Not Available)
Average Value, All Classes	70.7
Area-Weighted Average (Table 5)	78.3

Confidence Values by Region

Table 7. Global confidence values by continental regions (percent).	
Region	**Average Confidence Value**
Global	76.3
Africa	79.4
Austr & Insular Asia	83.2
Eurasia	76.8
North America	71.9
South America	78.5

Validation Conclusions

- Training sites are NOT a random sample
  - Many (perhaps most) training sites are placed in equivocal areas where the classifier needs new and better examples
  - Thus, the training sites do not represent well the broad regions of core areas for land cover classes
  - This leads to the conclusion that actual accuracies are probably better than observed from the training sites
  - So we estimate that:

GLOBAL ACCURACY IS 75–80 PERCENT

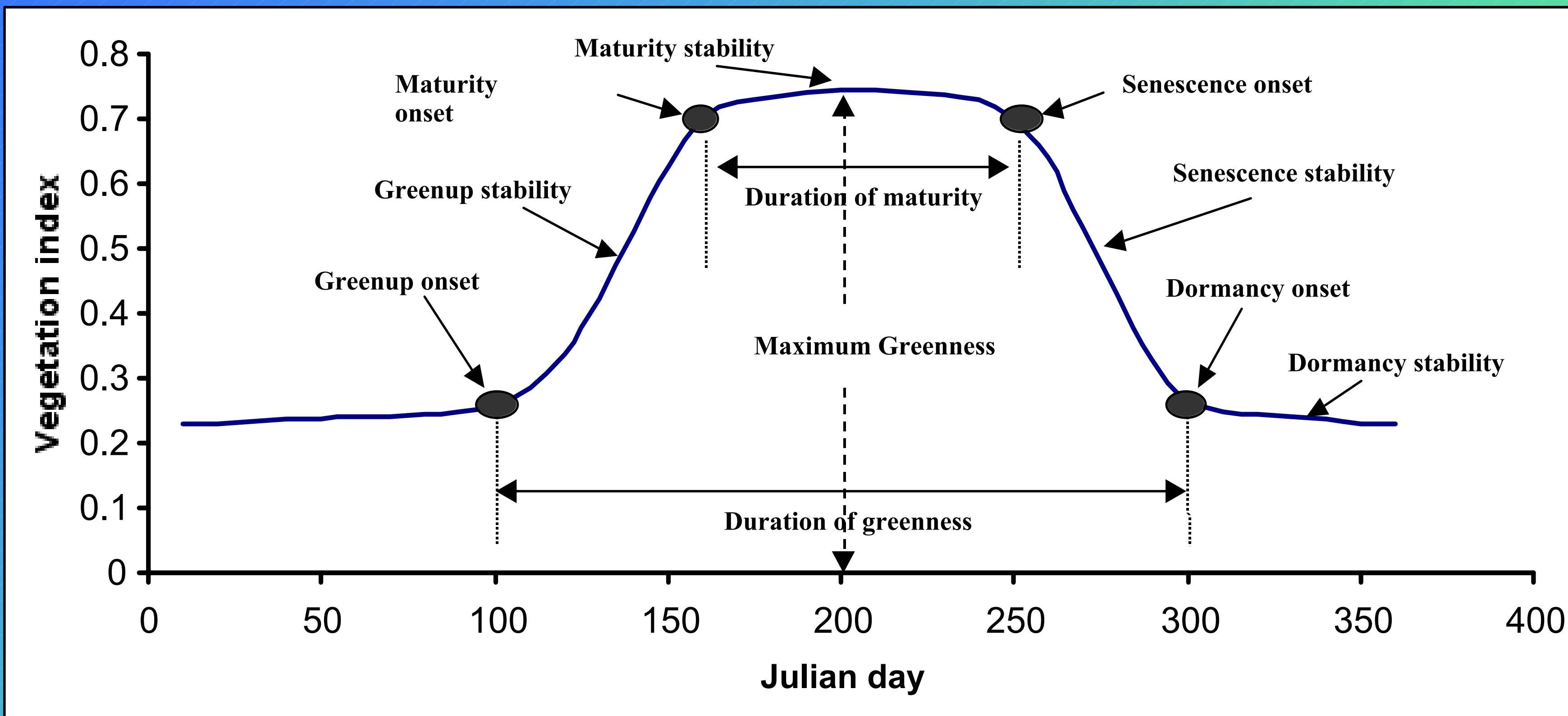
PER-CLASS ACCURACIES RANGE 60–90 PERCENT

CONTINENTAL REGION ACCURACIES RANGE 70–85 PERCENT

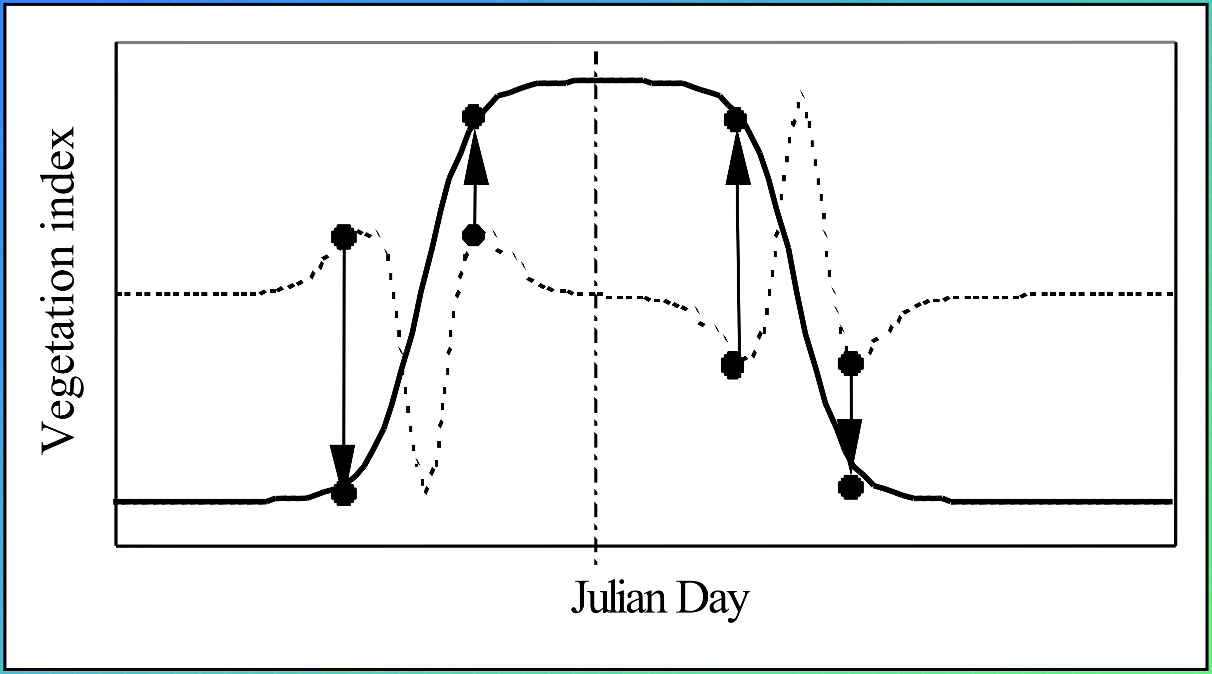
Land Cover Dynamics

- *Primary Objectives:*
  - Quantify interannual change
    - Uses change vectors comparing successive years
    - Identifies regions of short-term climate variation
    - Under development with Eric Lambin, Frederic Lupo at UCL, Belgium
  - Quantify phenology
    - Greenup, maturity, senescence, dormancy
    - Values of VI, EVI at greenup and peak, plus annual integrated values
    - Uses logistic functions fit to time trajectories of EVI

Land Cover Dynamics: Defining Phenological Attributes

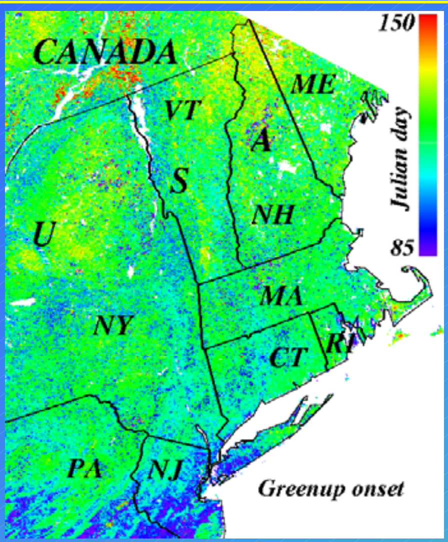


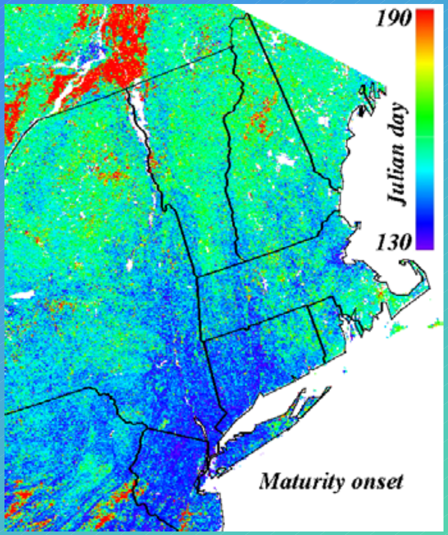
Quantifying Phenology

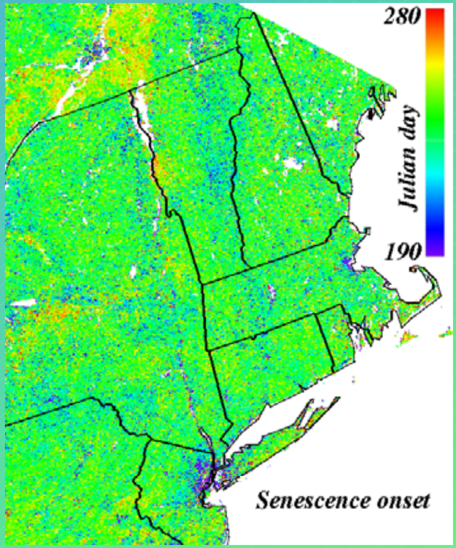


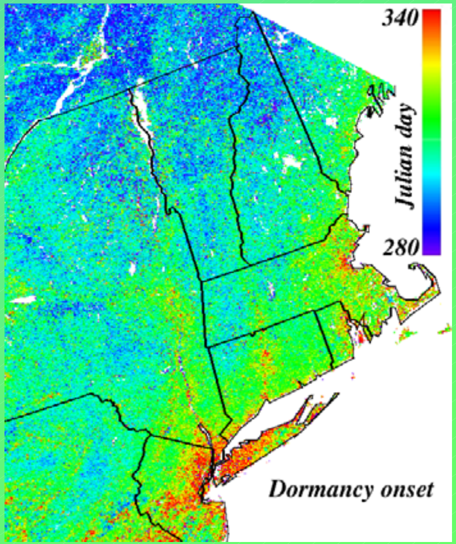
Use extreme points of the change rate of curvature to calculate phenology transition dates from NBAR-EVI

Northeast Phenology

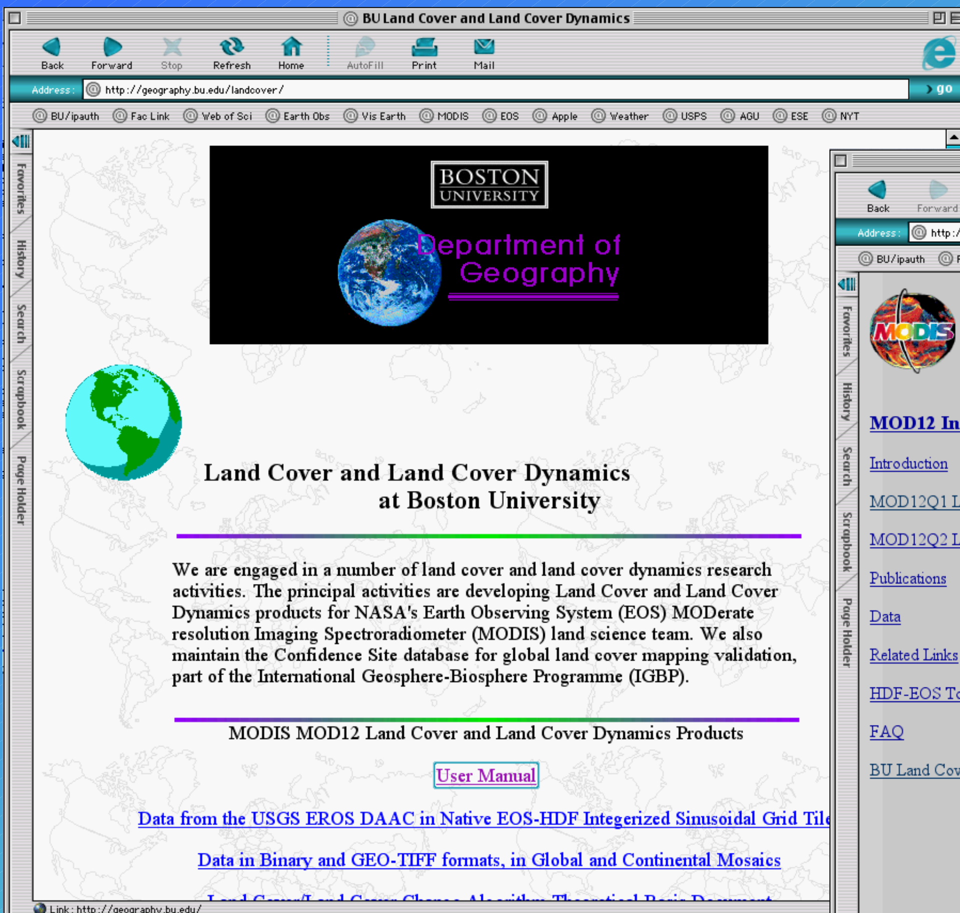


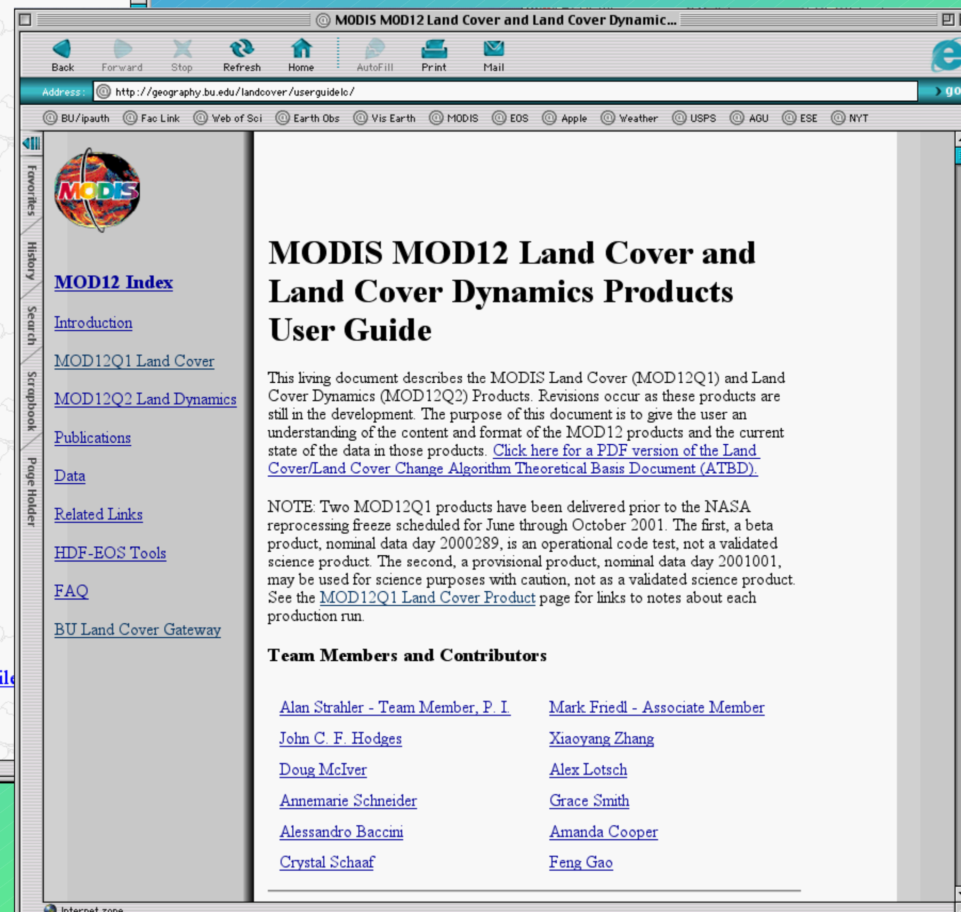






Web Site: http://geography.bu.edu/landcover





Introduction to the MODIS LAI and FPAR Algorithm

N. Shabanov, W. Yang, B. Tan, H. Dong,

R.B. Myneni, Y. Knyazikhin /Boston University

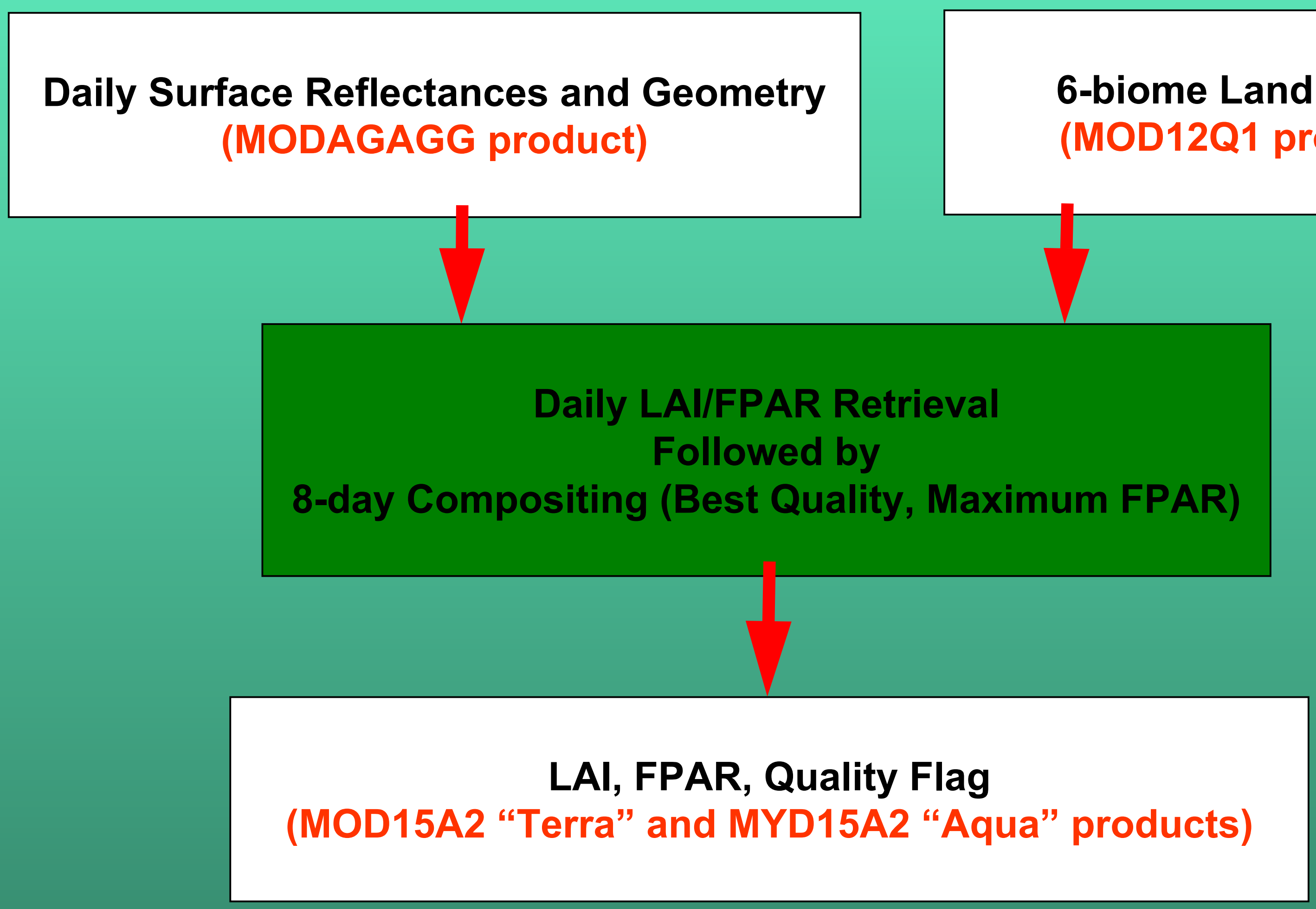
J.L. Privette, J.T. Morisette /NASA Goddard Space Flight Center

P. Votava, R. Nemani /NASA Ames Research Center

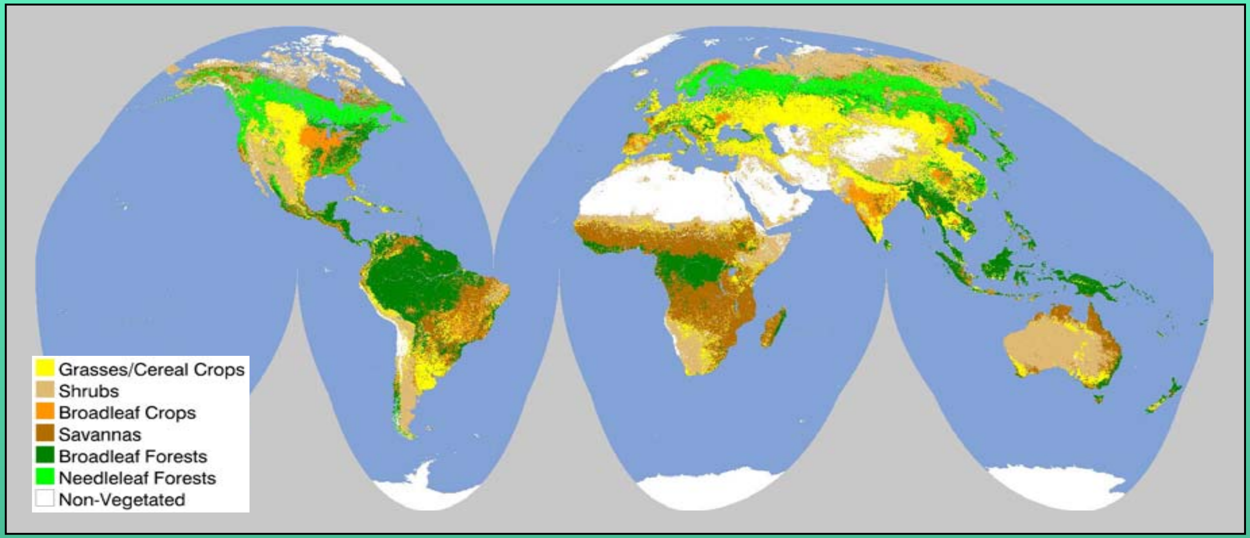
Boston University, Boston, MA,

September 16, 2003

MODIS LAI/FPAR Production Stream



MODIS 6-biome Land Cover Product

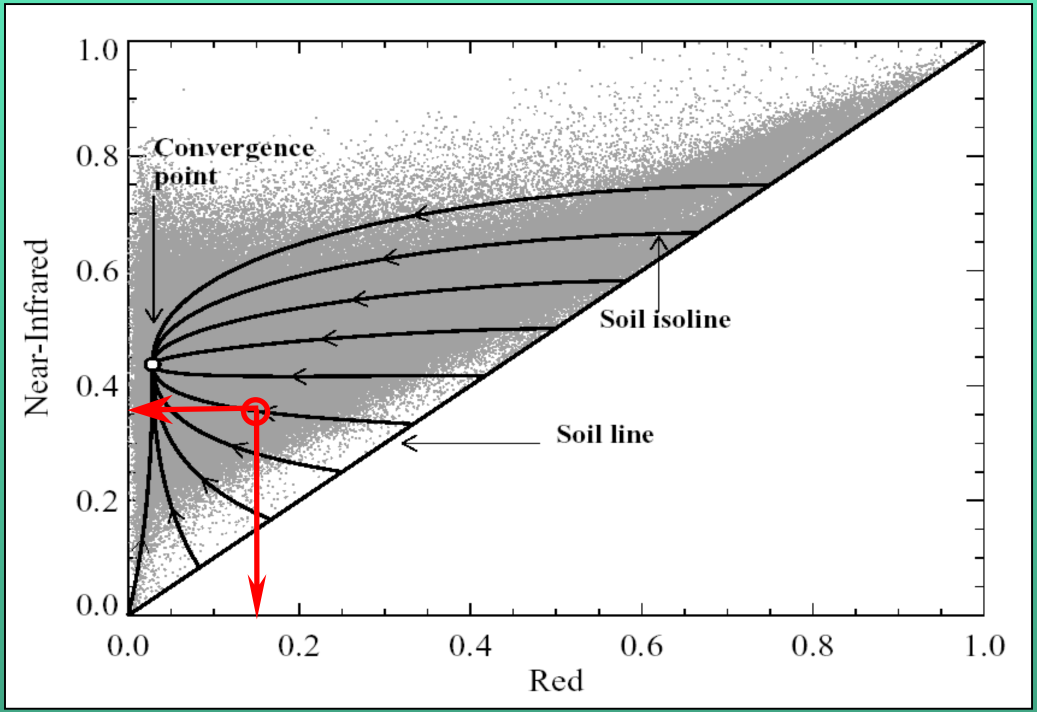


- **LAI and FPAR algorithm specify vegetation/soil properties as function of land cover type. Referencing land cover minimizes number of unknowns in inverse modeling.**

- **Classification with minimum distinguishable by Radiative Transfer theory number of classes (6 biomes) was selected for daily LAI retrievals.**

- **The latest version of land cover was derived based on one year of MODIS data.**

Retrieval Technique of the Main Algorithm



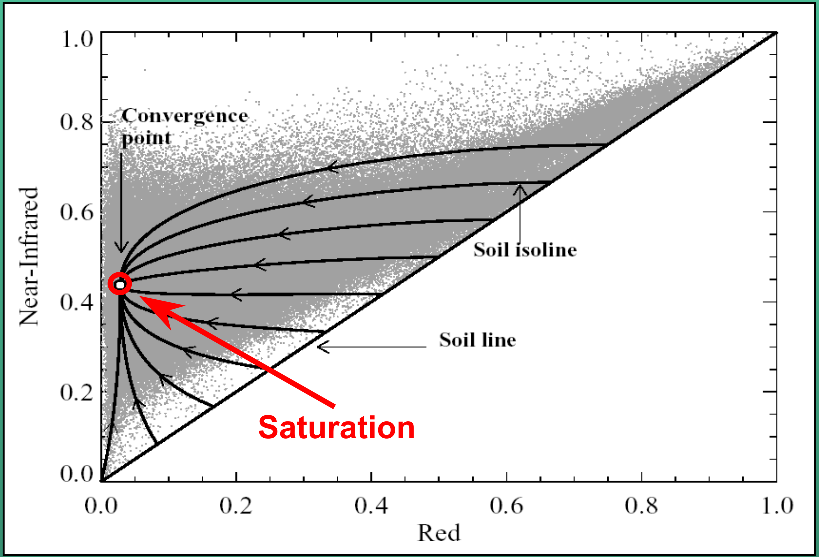
MODIS surface reflectances
RT Model simulations

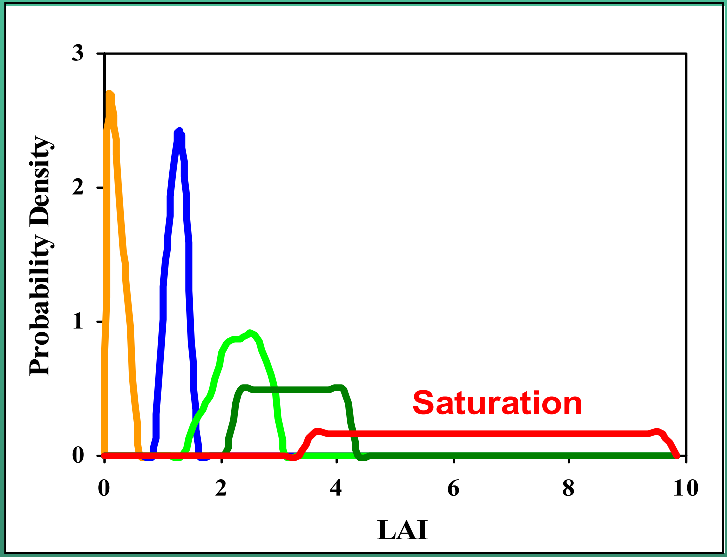
- To preserve information content of input data, MODIS channel data are used directly in LAI retrievals, instead of conversion to Vegetation Indexes.
- During retrievals, surface reflectances predicted by RT model are compared with MODIS data to identify LAI and FPAR.
- The algorithm is Look-Up Table based. LUTs parameterized with vegetation type, leaf optical properties, soil reflectance patterns, vegetation heterogeneity
- Retrieval technique takes into account uncertainties of model and input data, when searching LUTs for solution.

Statistical Nature of the Main Algorithm Retrievals

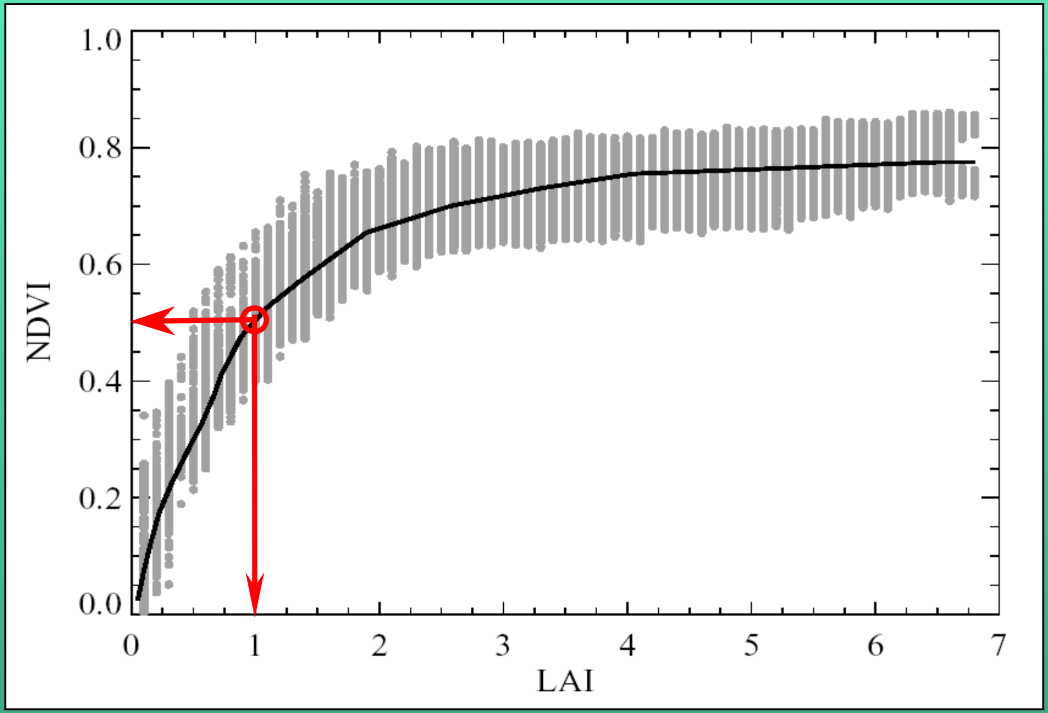
- Given surface reflectances and ancillary information, main algorithm generates distribution of solutions (LAI/FPAR) which fit to observed reflectances with specified uncertainties level and report averaged LAI/FPAR over this distribution.

- When LAI is high, surface reflectances have low sensitivity to LAI. This situation is called "saturation". Main algorithm performs saturation test and if saturation is found, set QA flag.





Retrievals Techniques of Back-Up Algorithm



MODIS surface reflectances
RT Model simulations

- When main algorithm fails, LAI is estimated with Back-Up algorithm.
- Back-Up algorithm retrievals are based on LAI-NDVI relationship, established from statistics of main algorithm retrievals.
- Generally retrievals with Back-Up algorithm should be considered as low quality.
- Typical examples of back-up algorithm retrieval includes retrievals under low sun/view zenith angles, under snow/cloudy conditions.

Reprocessing LAI and FPAR from MODIS Terra and Aqua Sensors

TERRA LAI/FPAR (MOD15A2 product):

- Collection 3-- Generation completed. Coverage: November 2000 – December 2002, Validation status = “Validated Stage 1”, QA status = “Inferred Passed”
- Collection 4-- Generation In progress. Released to public as of March 7, 2003. Coverage: March 2000 –present. Validation status = “Provisional”, QA status = “Inferred Passed”

AQUA LAI/FPAR (MYD15A2 product):

- Collection 3-- Generation In progress. Released to public as of September 12, 2003. Coverage: July 2003- present.
- Collection 4-- Reprocessing will start January 2004.

Boston University LAI and FPAR product:

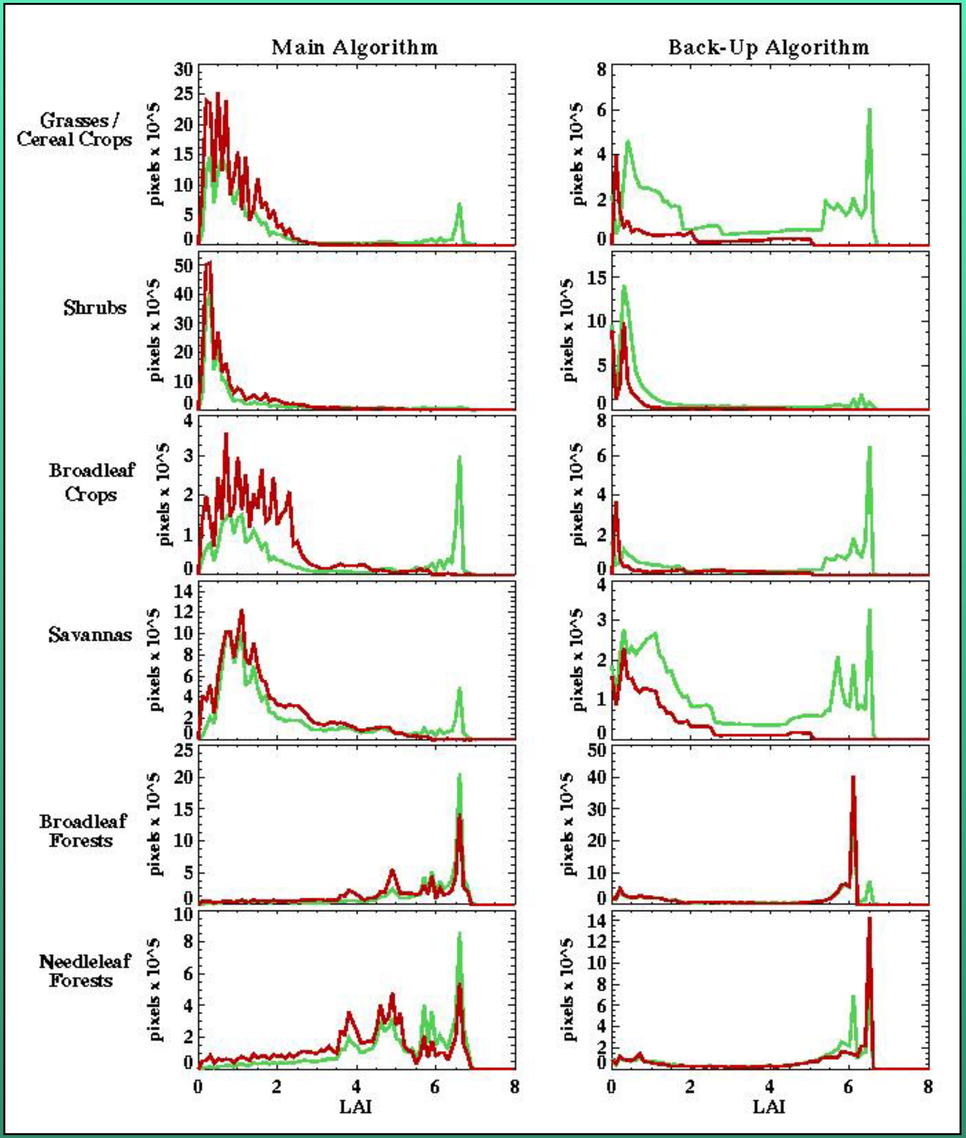
- Created on the basis of MOD15A2 product by best quality compositing. Monthly composites. 1- and 4- km spatial resolution. Generated from collection 3 and 4 MOD15A2 data. Available from PI website.

TERRA MODIS LAI/FPAR
Collection 4 Improvements

LUT Tuning for the Main and Back-up Algorithms

- Non-physical peaks at high LAI values for herbaceous vegetation (biome 1 - 4) were removed
- Validation feedback (BigFoot): improved agreement with field measurements (KONZA, grasses, ARGO, crops, etc.)
- Retrievals with main algorithm increased by 20% compared to collection 3 data

Collection 3 green
Collection 4 red



TERRA MODIS LAI/FPAR
Collection 4 Improvements (Cont.)

Achievements

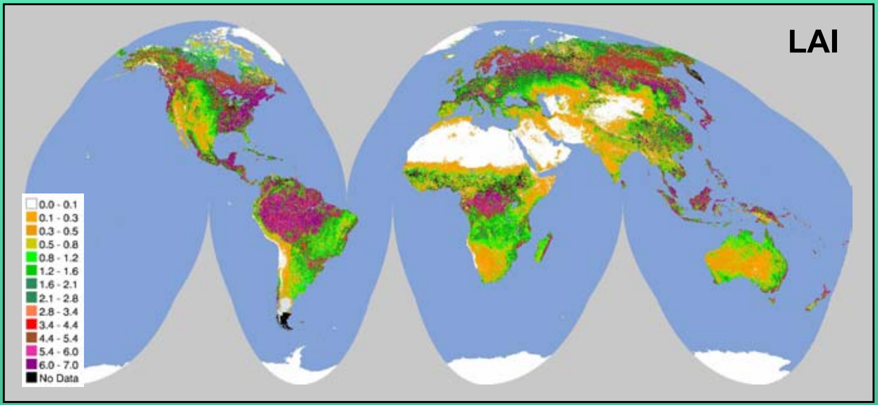
- **Spatial coverage of main algorithm increased by ~20% due to LUTs tuning and new compositing scheme:**

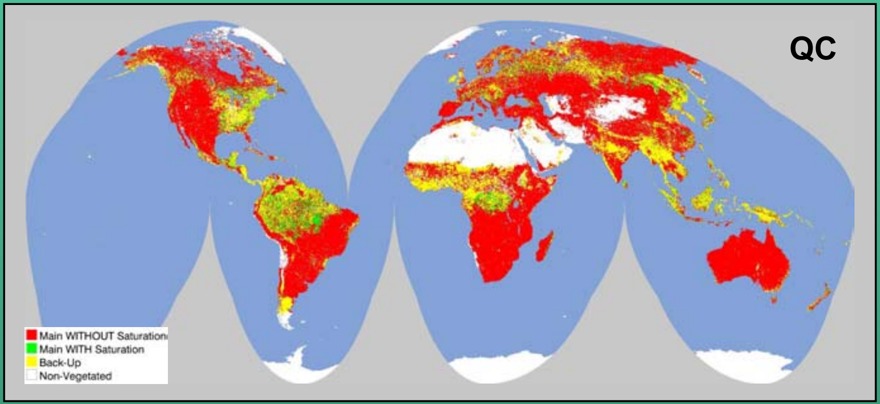
  Main=72%, Sat=5%, Bkup=23% (collection 4)
  Main+Sat=60%, Bkup=40% (collection 3)

- **Improved consistency with field observations over herbaceous vegetation**

Future Improvements (collection 5)…

- **Decrease dominance of back-up algorithm retrievals over woody vegetation (broadleaf and needleleaf forests)**
- **Further improve agreement with with field data**
- **Research on retrievals under snow condition (resolve needleleaf forests seasonality)**





MOD15A2, Collection 4 Data, July 20-27, 2001

Future Directions

Current MODIS LAI and FPAR contract will expire by the end of this year. Our future plans (if will be funded) are:

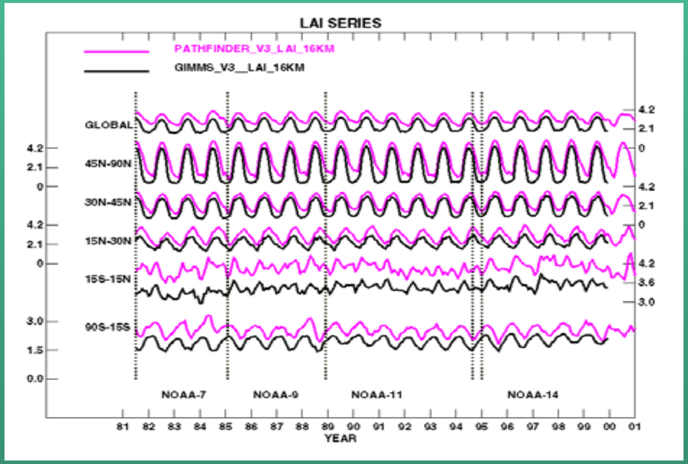
-- Continue tuning of the algorithm to achieve:

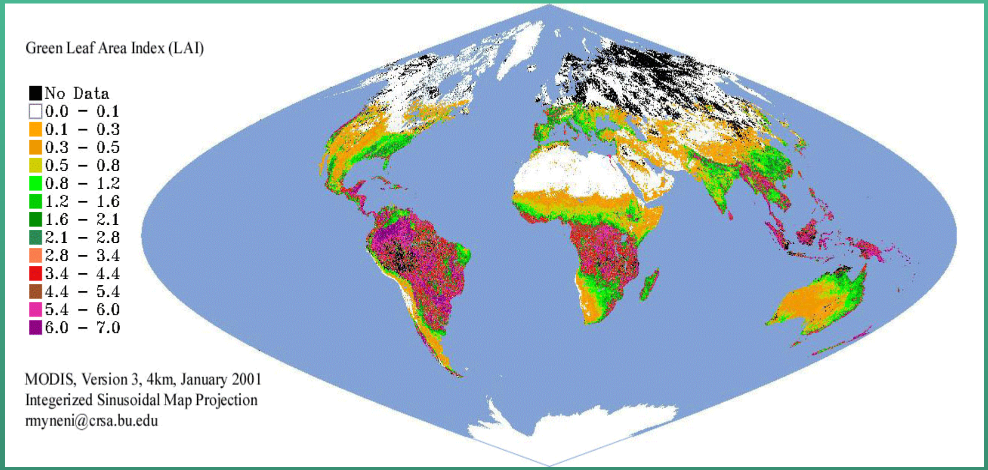
a) Increase in spatial coverage of the main algorithm retrievals by replacing back-up algorithm retrievals

b) Improve agreement with field data

-- Develop joint product, MODIS –Terra and –Aqua LAI and FPAR

-- Research on joining AVHRR and MODIS LAI and FPAR





Validation of the MODIS LAI and FPAR Product

N. Shabanov, W. Yang, B. Tan, H. Dong,

R.B. Myneni, Y. Knyazikhin /Boston University

J.L. Privette, J.T. Morisette /NASA Goddard Space Flight Center

P. Votava, R. Nemani /NASA Ames Research Center

Boston University, Boston, MA,

September 16, 2003

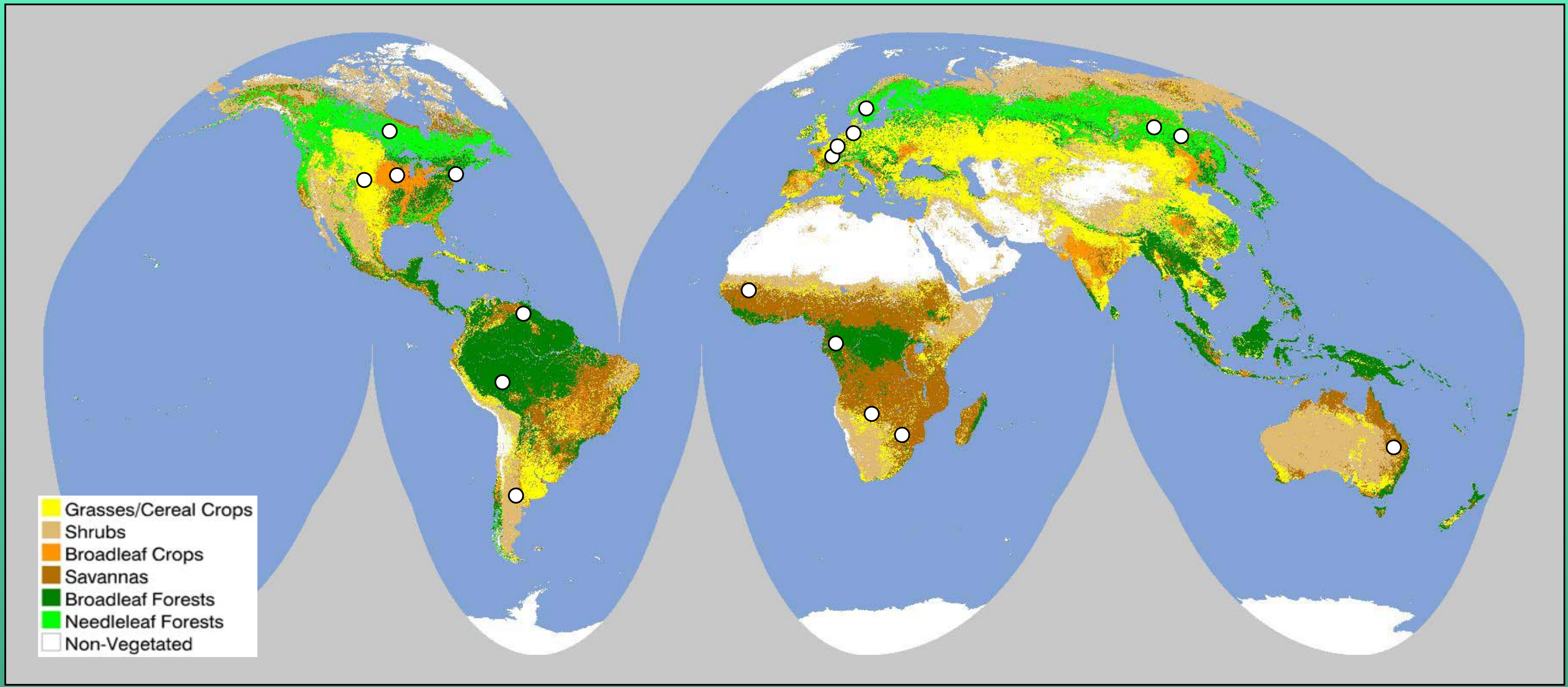
Definition of Validation

- Validation refers to assessing the uncertainty of satellite-derived products by analytical comparisons to reference data (e.g., in situ, aircraft and high-resolution satellite sensor data), which are presumed to represent the target values.

- Stage-1 Validation is defined by the MODIS as follows - product accuracy has been estimated using a small number of independent measurements from selected locations and time periods through ground-truth and validation efforts.

- <u>MODIS Terra LAI and FPAR product, collection 3 is Stage 1 validated</u>. The product have been validated with data from 9 field campaigns as of present moment:
  - Grasses: Konza, USA (Huang et al., 2003).
  - Savannas, Shrublands: South Africa (Privette et al., 2002, Tan et al., 2002, 2003).
  - Broadleaf Forests: Harvard Forest, USA (Shabanov et al., 2003).
  - Needleleaf Forest: Ruokolahti, Finland (Wang et al., 2003).

Field Campaigns Summary



- Validation performed by Boston University team in collaboration with NASA (Jeff Privette team), VALERI (Europe), BigFoot and others.
- Sampling field data over all 6 biomes. Preference was given to EOS core validation sites

Field Campaigns Summary (cont.)

	Biome: Validated by July, 2003	Transects: Planned
Grasses/ Cereal Crops	Konza, USA SAFARI 2000 wet season	Uardry, Australia Laprida, Argentina
Shrubs	Puechabon, France	Turco, Bolivia
Broadleaf Crops	Bondville, USA	Romille/Seine, France
Savannas	SAFARI 2000 wet season	Brazilia, Brazil
Broadleaf Forests	Harvard Forest, USA	Tapajos/Ji-Parana, Brazil Jaervselja, estonia
Needle Forests	Ruokolahti, Finland Flakaliden, Sweeden	Krasnoyarsk, Russia

Sources of Uncertainties in LAI and FPAR Product

Daily Surface Reflectances and Geometry
(MODAGAGG product)

6-biome Land Cover
(MOD12Q1 product)

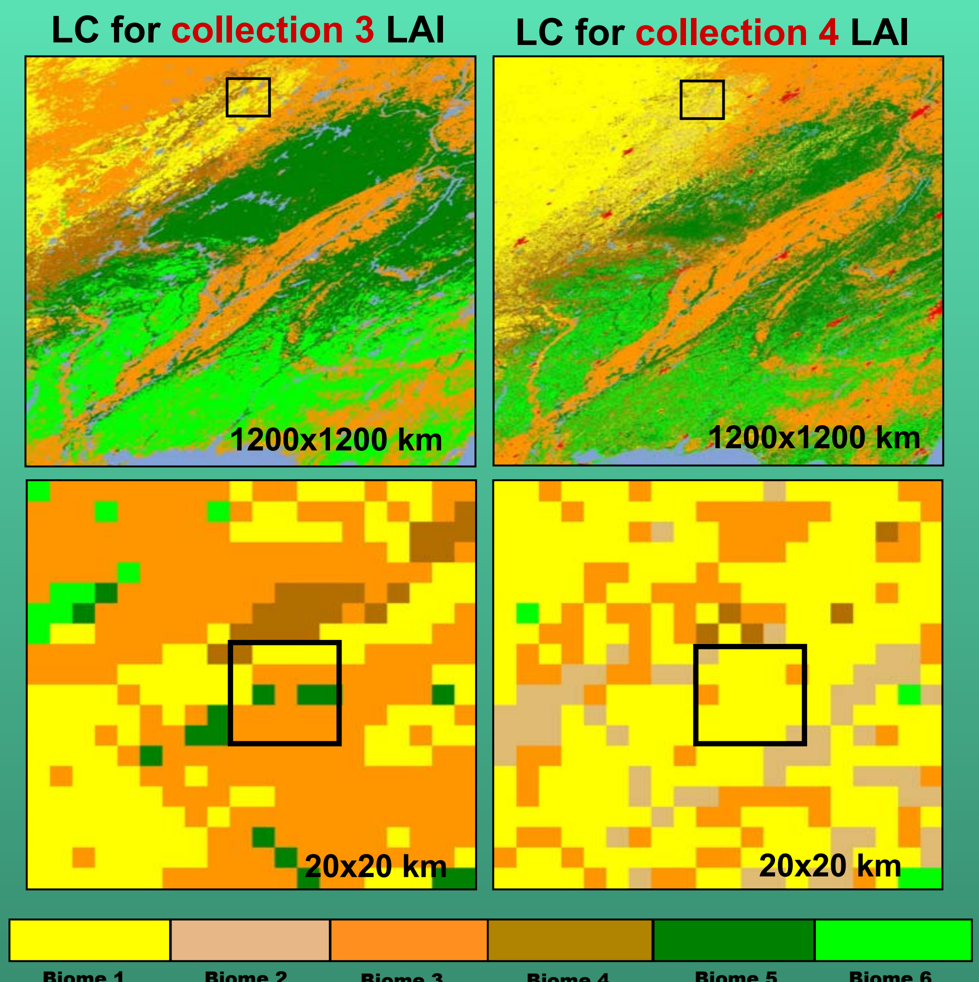
Daily LAI/FPAR Retrieval
Followed by
8-day Compositing (Best Quality, Maximum FPAR)

Uncertainties in LAI and FPAR product are due to two factors:

- **Uncertainties in input data streams, i.e.,**
  a) Surface reflectance data
  b) Land cover map

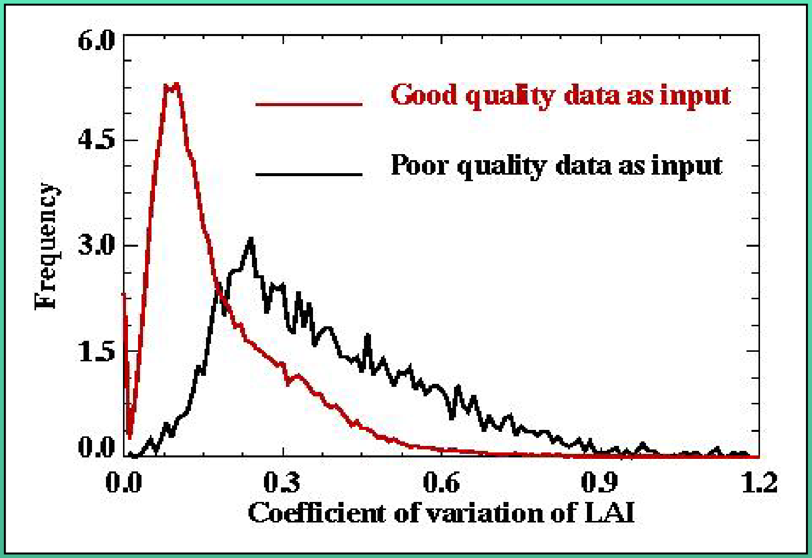
- **And uncertainties of LAI and FPAR algorithm, particularly Look-Up Tables**

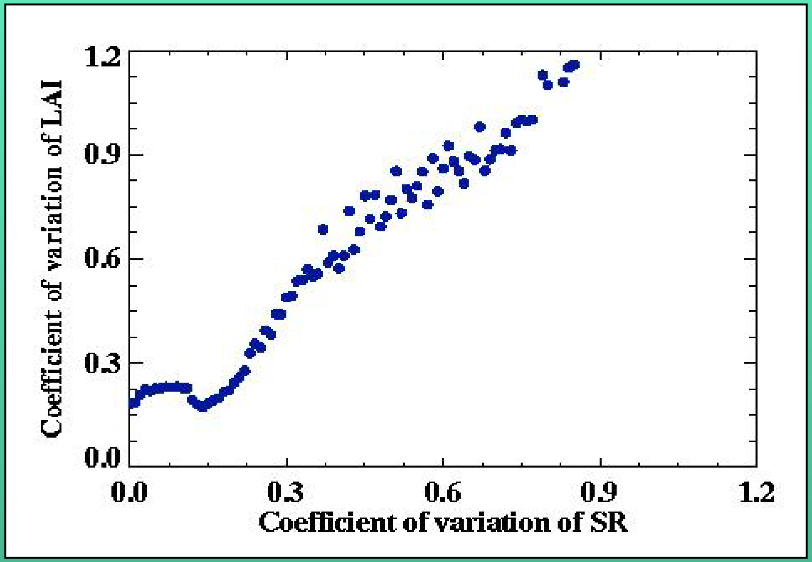
Impact of Biome Misclassification on LAI Retrievals



- MODIS LAI/FPAR algorithm references LC to select vegetation parameters from LUTs. Misclassification leads to errors in LAI estimation
- Collection 3 LAI used MODIS at-launch IGBP LC (AVHRR-based), cross-walked to 6 biome LC
- Collection 4 LAI referencing MODIS 6-biome LC product (based on one year MODIS data)
- Significant misclassification occur at local scale (5x5 km) for at-launch LC: this map predicts 24% of the pixels are grasses, while field measurements indicates that 64% of the pixels are grasses.

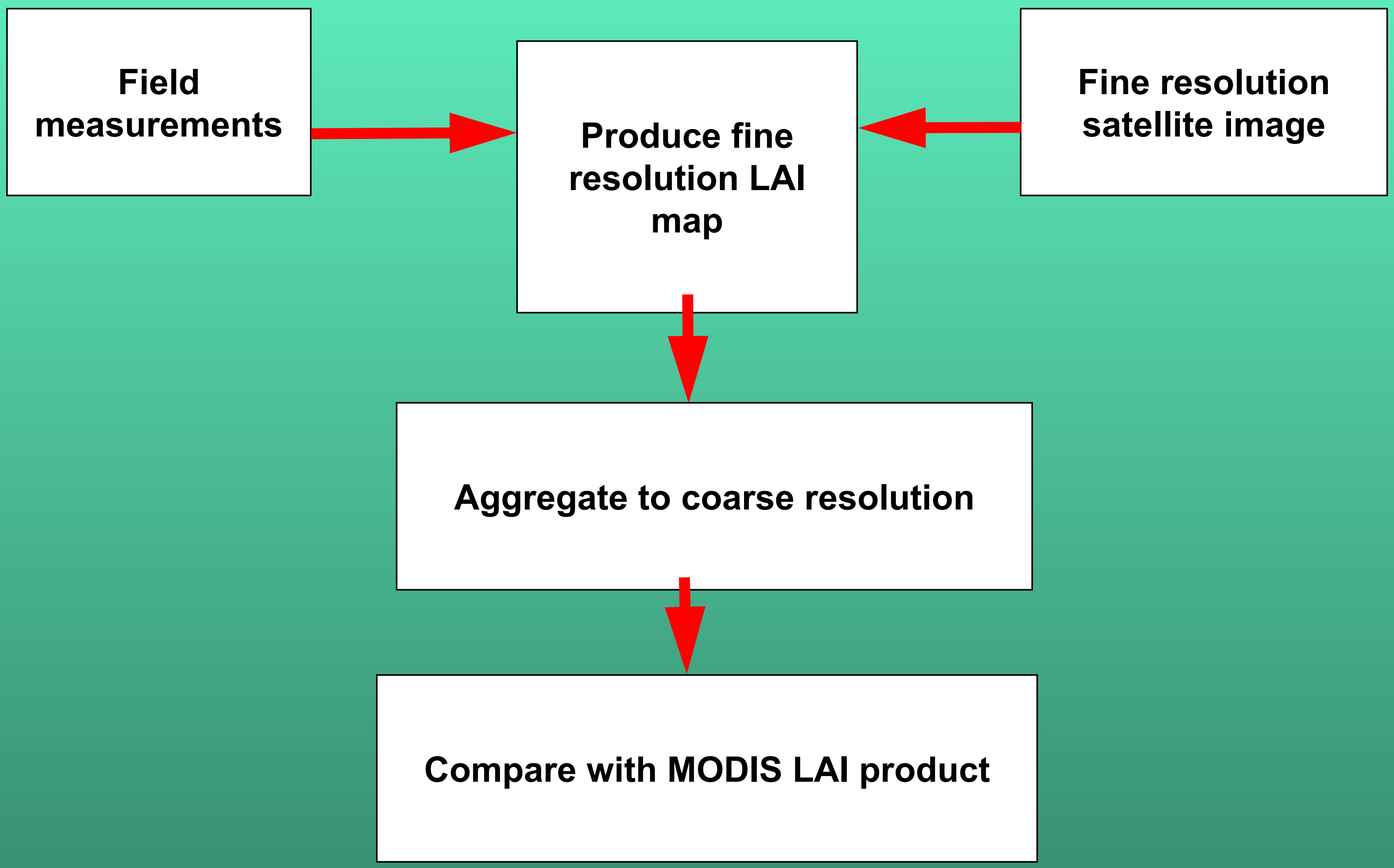
Impact of Uncertainties in Surface Reflectances on LAI Retrievals





- Definition: "Good data" are surface reflectance data with MODAGAGG QA = "Product produced at ideal quality" or "Product produces, less than ideal quality"
- Good quality data have lower uncertainty than poor quality data
- Uncertainties in LAI retrievals are proportional to uncertainties in surface reflectance variations

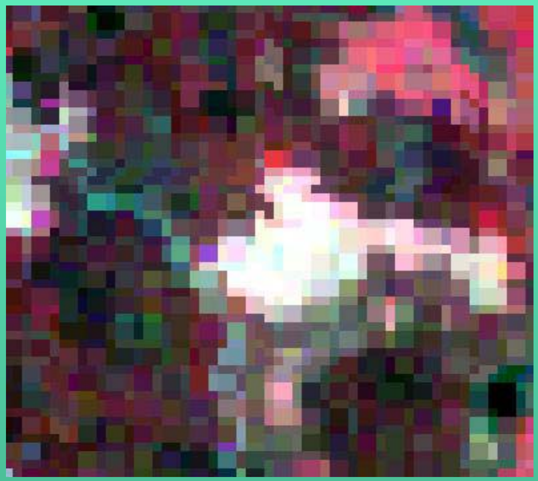
Sampling Strategy



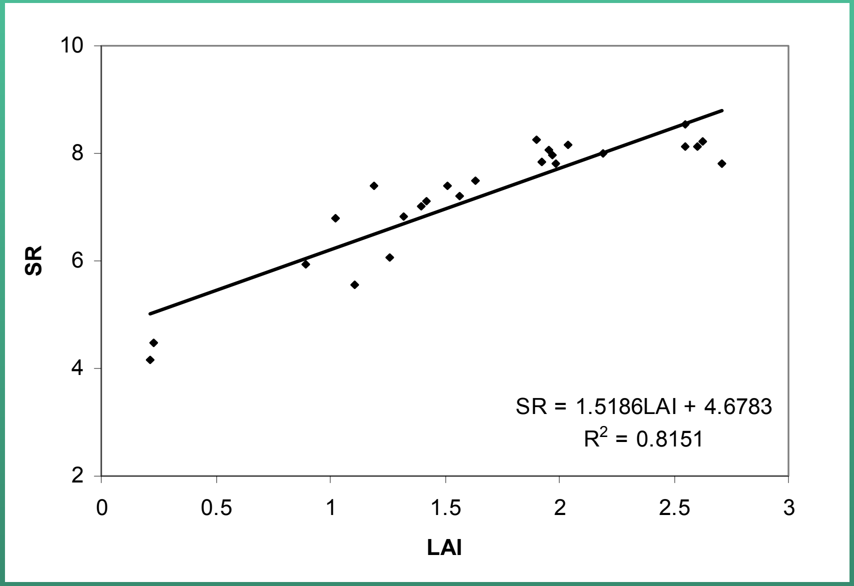
Ruokolahti, Finland

- June 14-21, 2000
- Needleleaf forest
- Establish scaling procedure to scale up from field measurements through fine-resolution (ETM+) LAI to moderate resolution (MODIS) LAI using Simple Ratio (NIR/Red)
- Patch-by-patch comparison was used

ETM+ reflectances

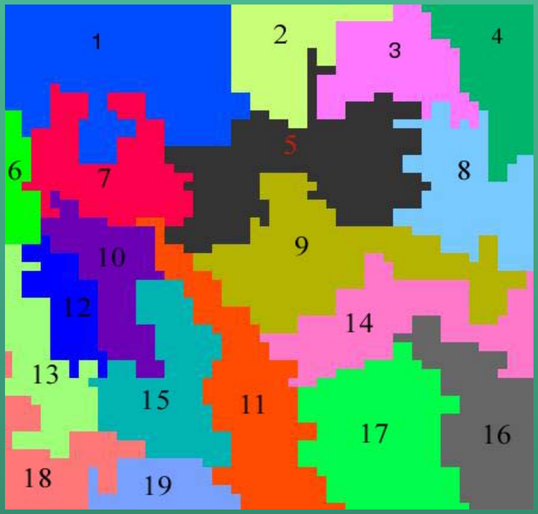


Estimation of regression curve: SR versus LAI

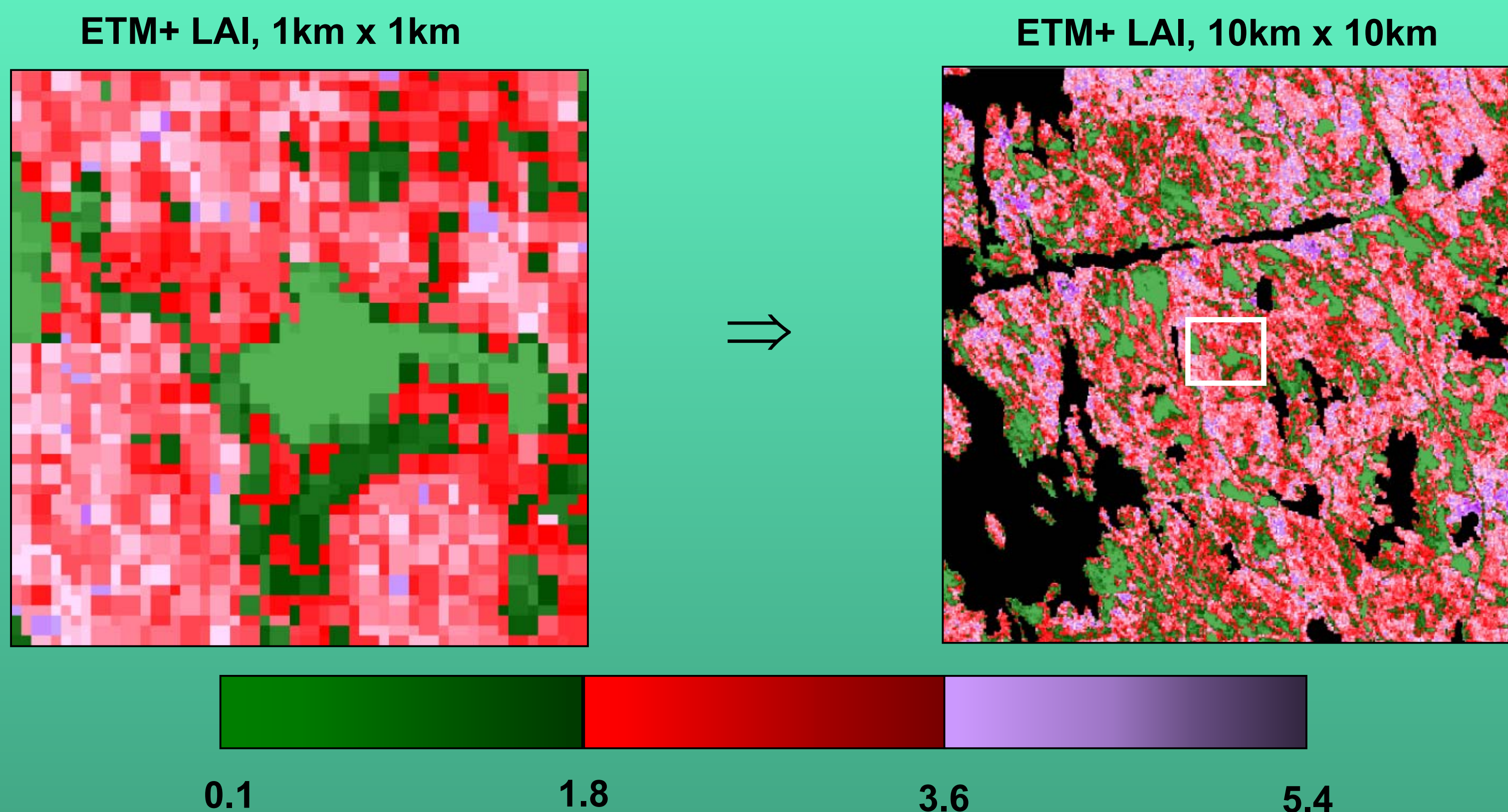


$\Leftarrow$

ETM+ segmentation



Ruokolahti, Finland (Cont.)

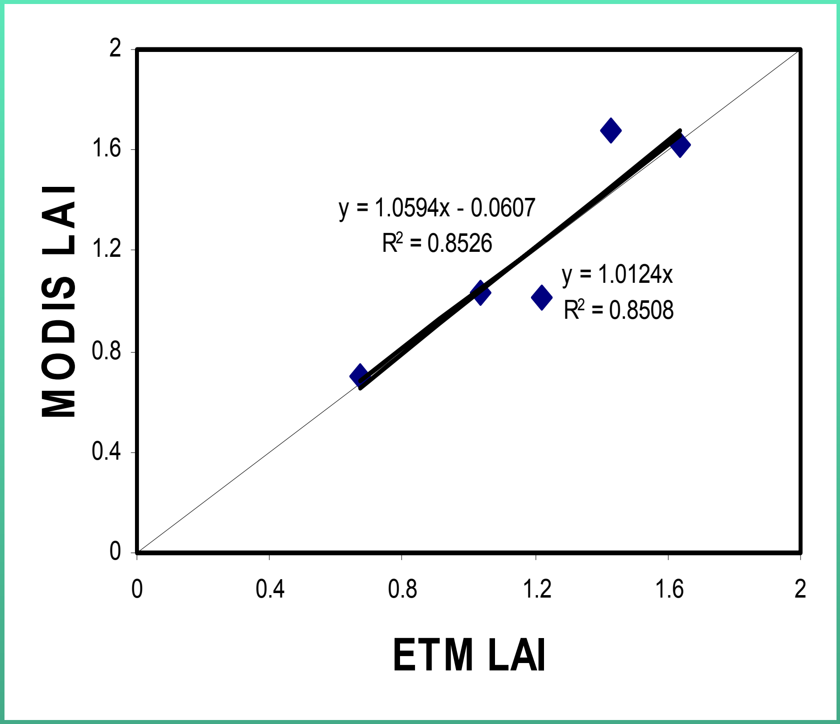


- Using relationship Simple Ratio to LAI estimated at 1-km, estimate LAI at 10-km from ETM+ data

Ruokolahti, Finland (Cont.)

Aggregated ETM+ LAI

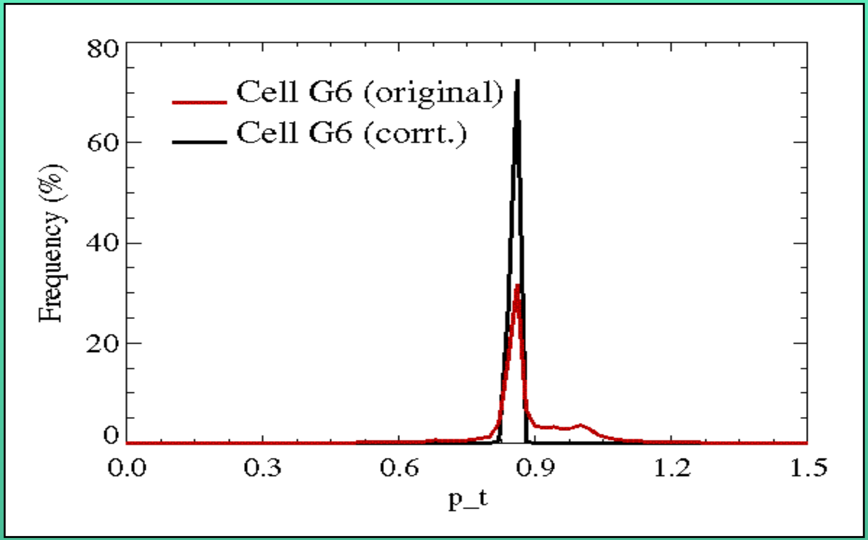
Patch-by-patch: LAI MODIS versus LAI ETM+

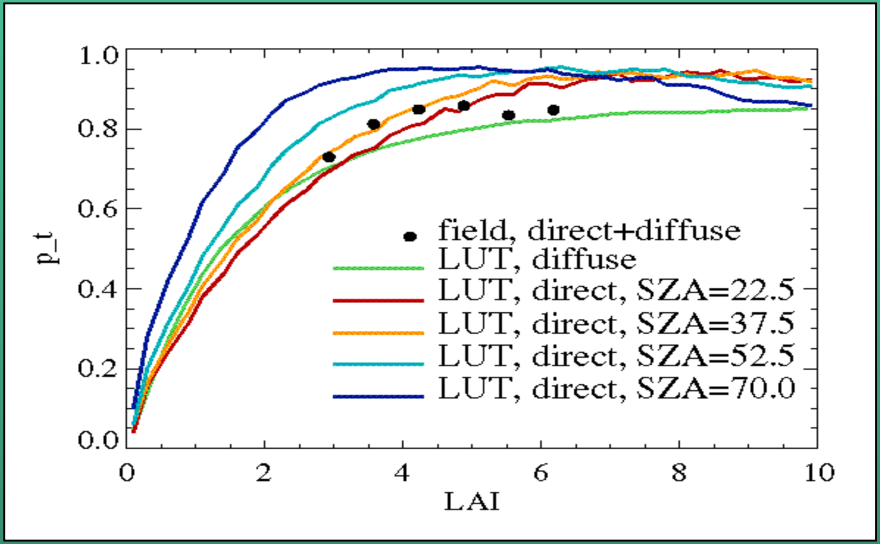


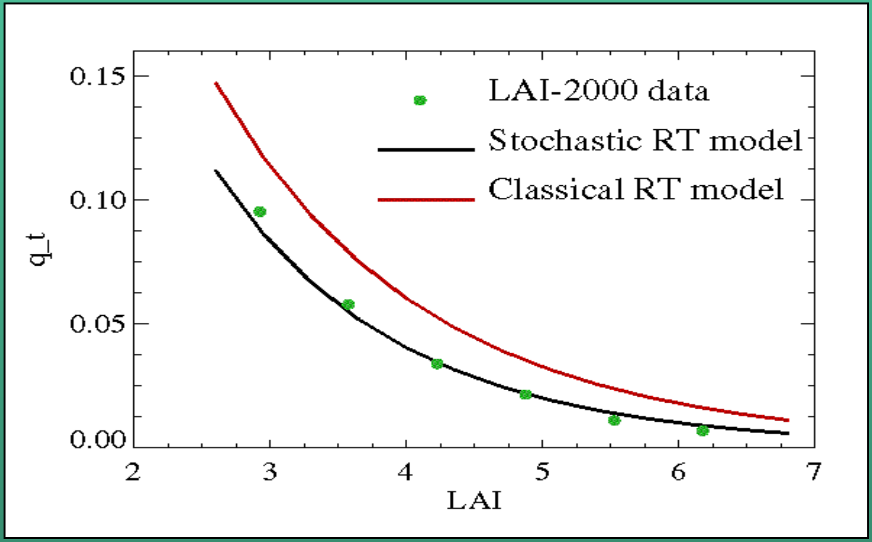
- **Finally degrade LAI at 30m from ETM+ to 1-km**
- **Patch-by-patch comparison: use patches of similar LAI to compare MODIS LAI with aggregated ETM+ LAI**

Harvard Forest, Massachusetts

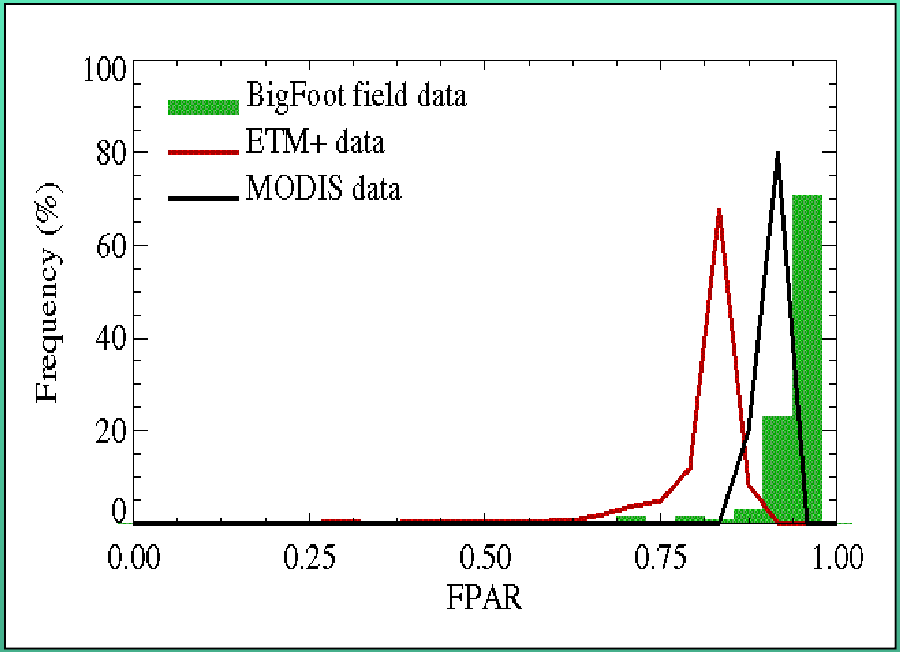
- **July 21-25, 2000**
- **Broadleaf forest**
- **Studied theoretical assumptions of LAI/FPAR algorithm to clarify physical interpretation of parameters**
- **Validated key parameters of Look-Up Table- ratio of uncollided to total radiation, p_t- parameter**

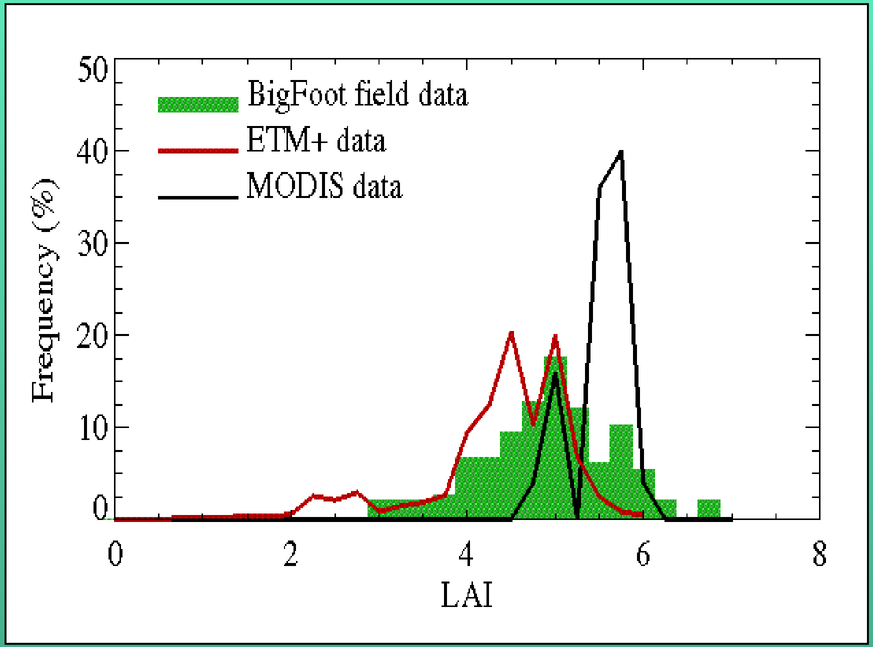






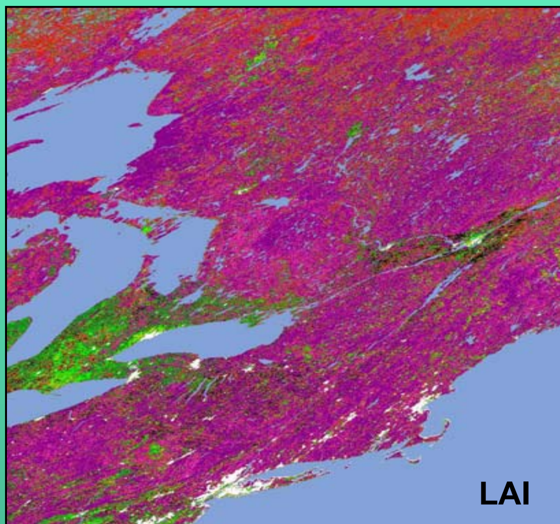
Harvard Forest , Massachusetts (Cont.)



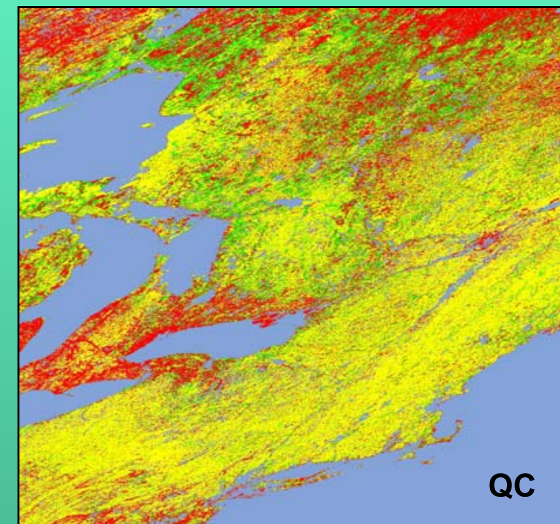


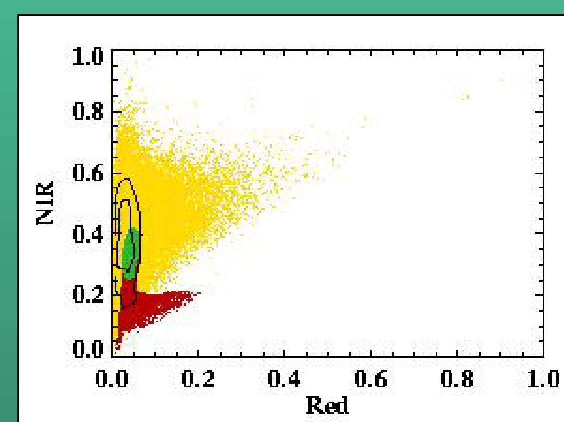
- Studied performance of the algorithm for high LAI values (saturation/low-sensitivity of surface reflectances to LAI).
- BigFoot field data of LAI and FPAR at 5x5 km area with efficient sampling were used.
- Field data match with LAI and FPAR product with uncertainties of 20%.

Harvard Forest , Massachusetts (Cont.)



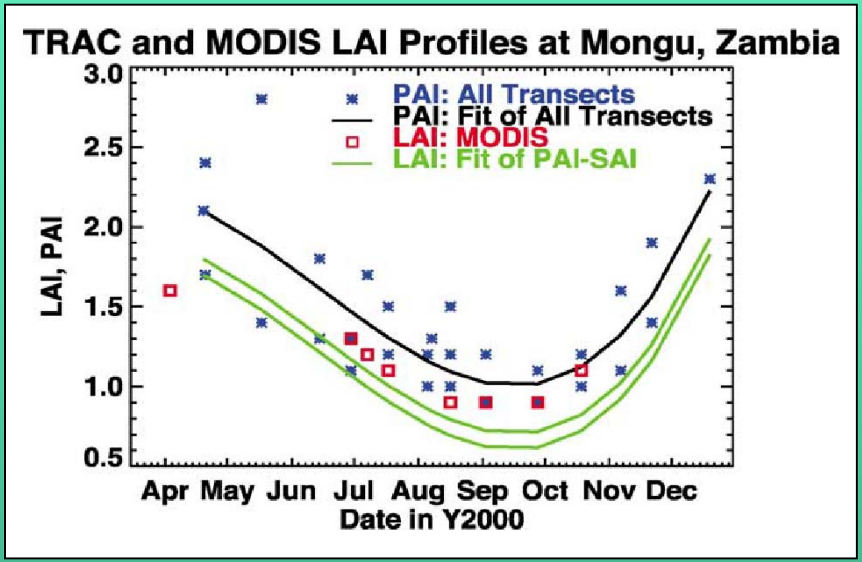
LAI~4-6

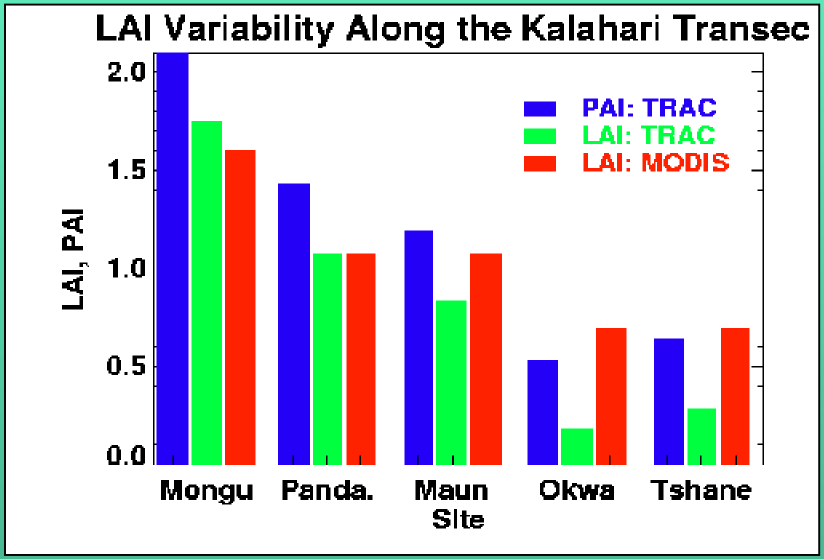




- Despite of agreement of MODIS LAI product with field measurements, QC indicates that retrievals are performed by back-up algorithm mostly, not main algorithm.
- This is due to mismatch of LUTs of main algorithm with MODIS surface reflectances. Currently research is performed to improve LUTs.

South Africa





- The first-year MODIS LAI algorithm correctly accommodates structural and phenological variability in semiarid woodlands and savannas, and is accurate to within the uncertainty of the validation approach used here.

Conclusions

- Validation and QA analysis performed within efforts of multiple teams resulted in identification of problems with product and developing strategy to refine product.
- Protocols (sampling strategy) for the validation of moderate resolution satellite products have been developed and fostered under CEOS-LPV and NASA sponsorship.
- MODIS Terra LAI and FPAR collection 3 product is stage 1 validated. We plan to extend validation results to collection 4 data (derived both from Terra and Aqua sensors)
- More field campaigns are planned to sample variability in LAI and FPAR as function of continental transects and different climatic conditions. This will allow to upgrade validation status of product to higher level of confidence.

Vegetation Structure from MODIS Bidirectional Reflectance

- *Feng Gao*

- *Boston University*
- *BRDF/Albedo MODIS Team*

Vegetation Structure from MODIS Bidirectional Reflectance

- Vegetation structure causes vegetated land surfaces to exhibit an anisotropic bidirectional reflectance distribution function (BRDF)
  - ***I.e., shadow hiding by three-dimensional objects causes a "hotspot" in the BRDF where illumination and viewing positions in the hemisphere coincide***
- The MODIS BRDF/Albedo product provides a three-parameter BRDF for each 1-km of land surface every 16-days (given at least one cloud-free MODIS look in that period)
  - ***The shape of the BRDF, as described by the BRDF model parameters, can be related to vegetation structure***

Kernel-Driven Semiempirical BRDF Models

- BRDF model—linear combination of two BRDF shapes and a constant (J. L. Roujean)
- BRDF shapes described by kernels
  - ***trigonometric functions of incidence and view angles***
  - ***derived from physical models for surface scattering***
- Analytical form

$$R = f_{iso} + f_{vol}k_{vol} + f_{geo}k_{geo}$$

- ***where f_{iso} is a constant for isotropic scattering; k_{geo}, k_{vol} are trigonometric functions providing shapes for geometric-optical and volume-scattering BRDFs; and f_{geo}, f_{vol} are constants that weight the two BRDFs***
- ***We use the Ross-Thick kernel for volume scattering and the Li-Sparse kernel for geometric scattering***

Ross-Thick Kernel

- Kernel formula

$$k_{thick} = \frac{(\pi/2 - \phi)\cos\phi + \sin\phi}{\cos\theta_i + \cos\theta_v} - \frac{\pi}{4}$$

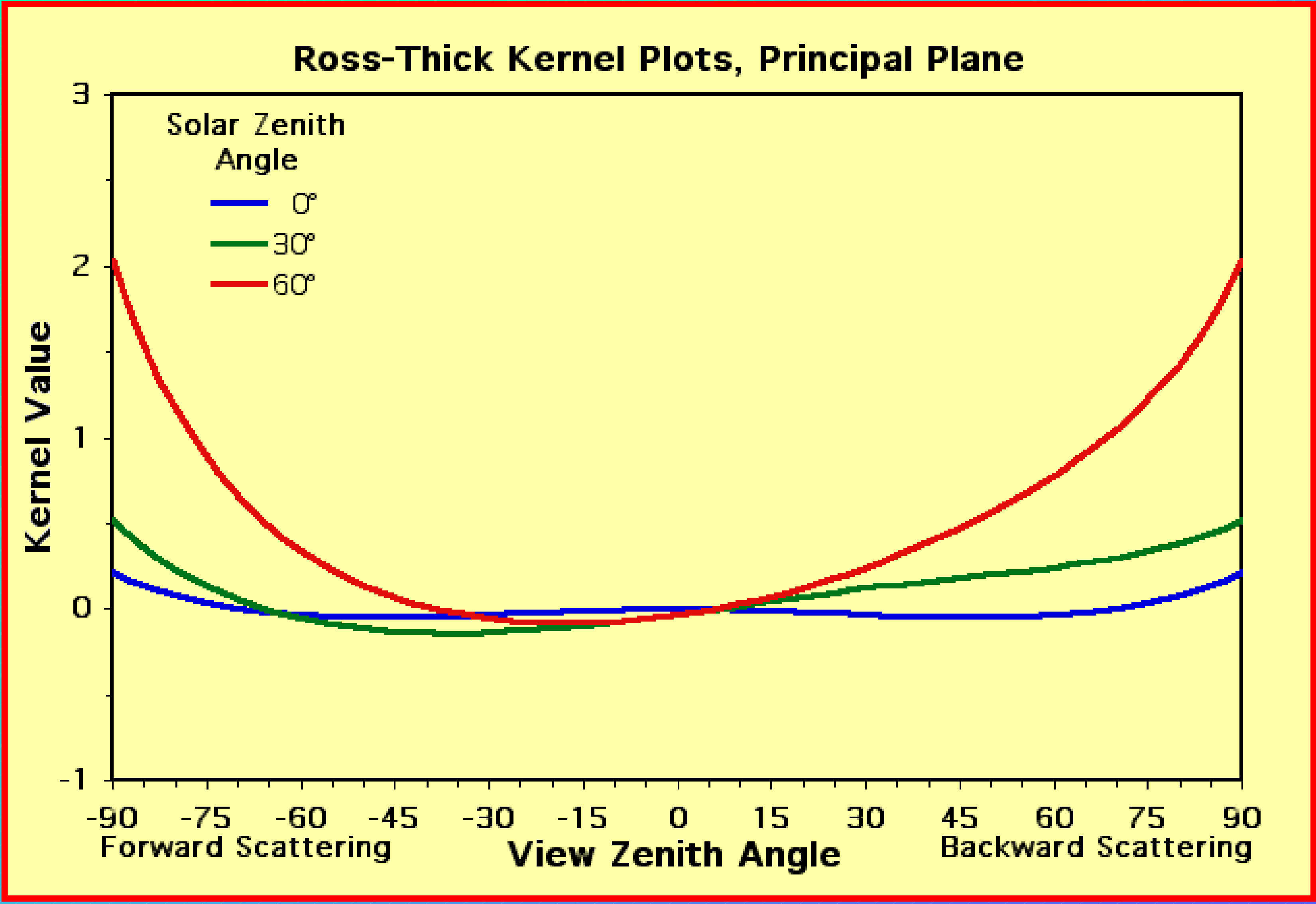
  - ***where ϕ is the phase angle between illumination and view positions in the hemisphere***

- Constants $R = c_1 k + c_2$

$$c_1 = \frac{4s}{3\pi}\left(1 - e^{-LAI\ B}\right) \qquad c_2 = \frac{s}{3} + e^{-LAI\ B}\left(\rho_s - \frac{s}{3}\right)$$

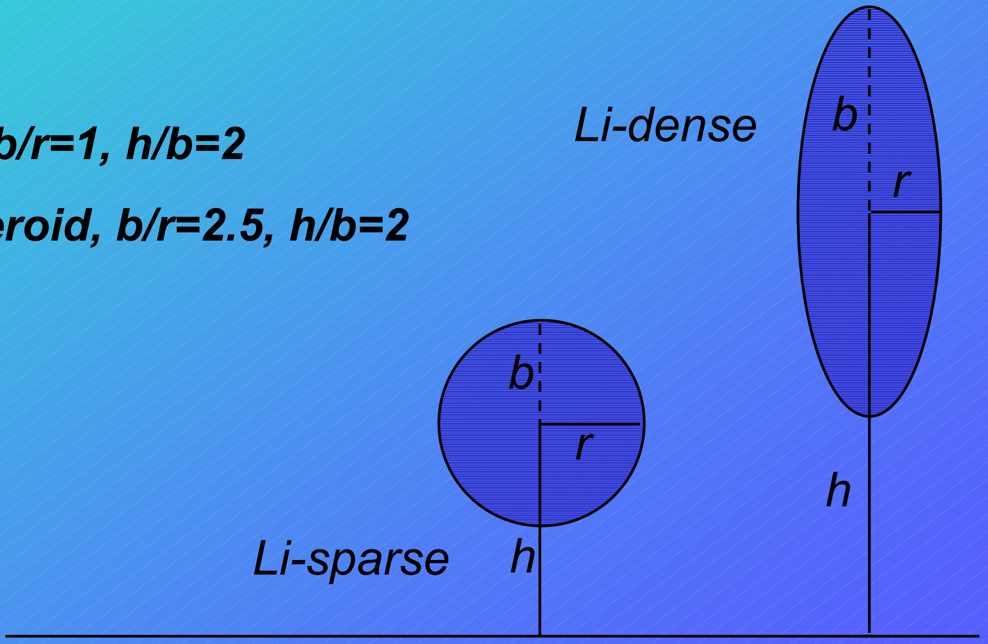
  - ***where s is leaf reflectance; ρ_s is the surface reflectance; LAI is the leaf area index; B is the average of secants of possible view and illumination angles (≈1.5)***

Ross-Thick Kernel Plots



Li Kernels

- Assumptions
  - ***Shadows are perfectly black***
  - ***Sunlit surfaces, whether object or background, are equally bright***
  - ***Some geometric approximations for the overlap of view and illumination shadows***
- Crown shape choices
  - ***Li-Sparse: low sphere, b/r=1, h/b=2***
  - ***Li-Dense: tall, thin spheroid, b/r=2.5, h/b=2***



Li-Sparse Kernel

- Kernel model

$$k_{sparse} = O(\theta_i, \theta_v, \phi) - \sec\theta_i' - \sec\theta_v' + \frac{1}{2}(1+\cos\phi')\sec\theta_v'$$

- ***where***

$$O = \frac{1}{\pi}(t - \sin t\cos t)(\sec\theta_i' + \sec\theta_v') \qquad \theta' = \tan^{-1}\left(\frac{b}{r}\tan\theta\right)$$

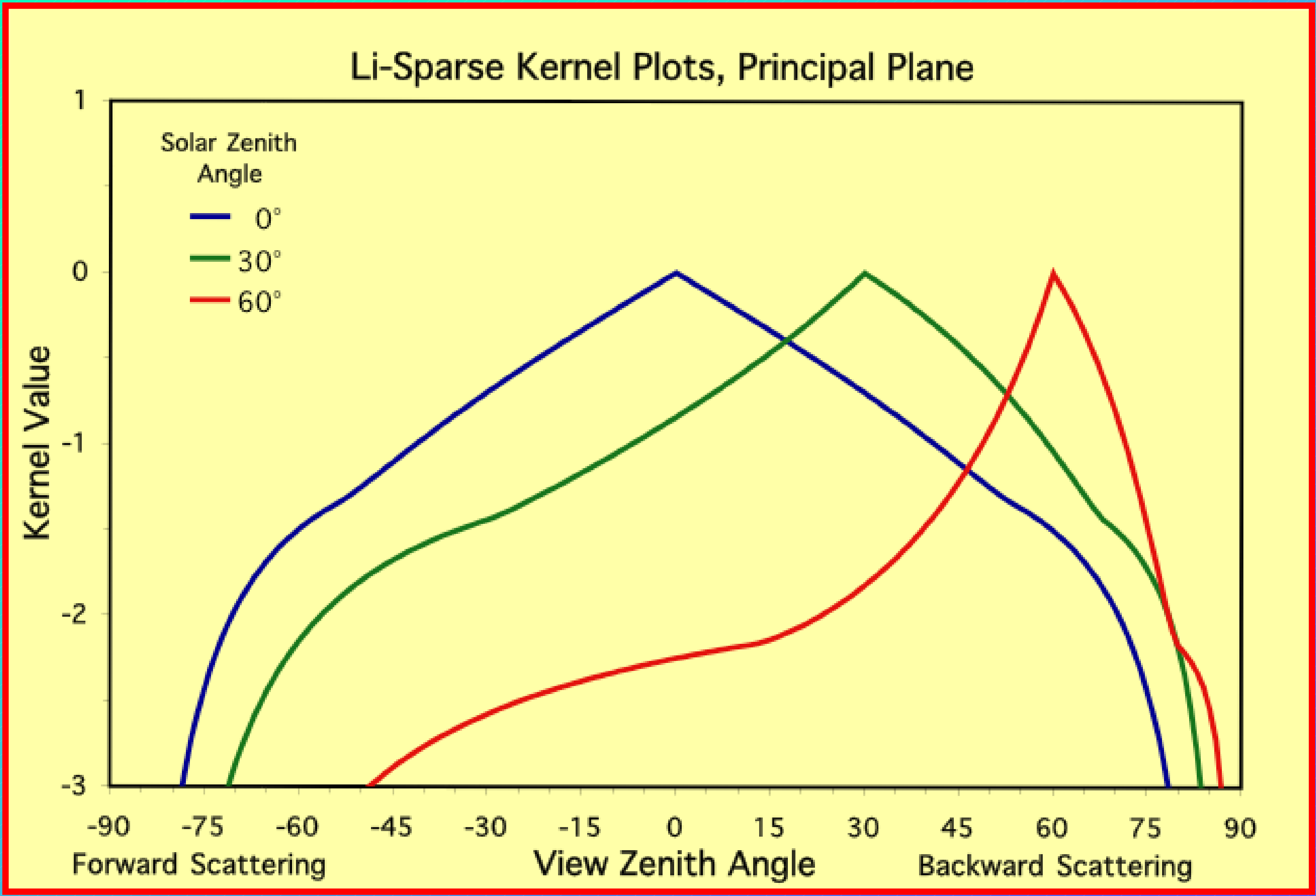
$$\cos t = \mathrm{Max}\left[\frac{h}{b}\frac{\sqrt{D^2 + (\tan\theta_i' \tan\theta_v' \sin\phi)^2}}{\sec\theta_i' + \sec\theta_v'}, 1\right]$$

$$D = \sqrt{\tan^2\theta_i' + \tan^2\theta_v' - 2\tan\theta_i' \tan\theta_v' \cos\phi}$$

$$\cos\phi' = \cos\theta_i' \cos\theta_v' + \sin\theta_i' \sin\theta_v' \cos\phi$$

- Constants $R = c_1 k + c_2 \qquad c_1 = C\lambda\pi r^2 \qquad c_2 = C$
  - ***where C is the brightness of sunlit surface; and λ is the count density of spheroids***

Li-Sparse Kernel Plots



Vegetation Structure Effects on BRDF (Gao, 2002)

Define Structural Scattering Index *SSI* as log ratio of weights for Ross-Thick kernel in NIR and Li-Sparse kernel in red:

$$SSI = \ln\left(\frac{f_{vol,NIR}}{f_{geo,red}}\right)$$

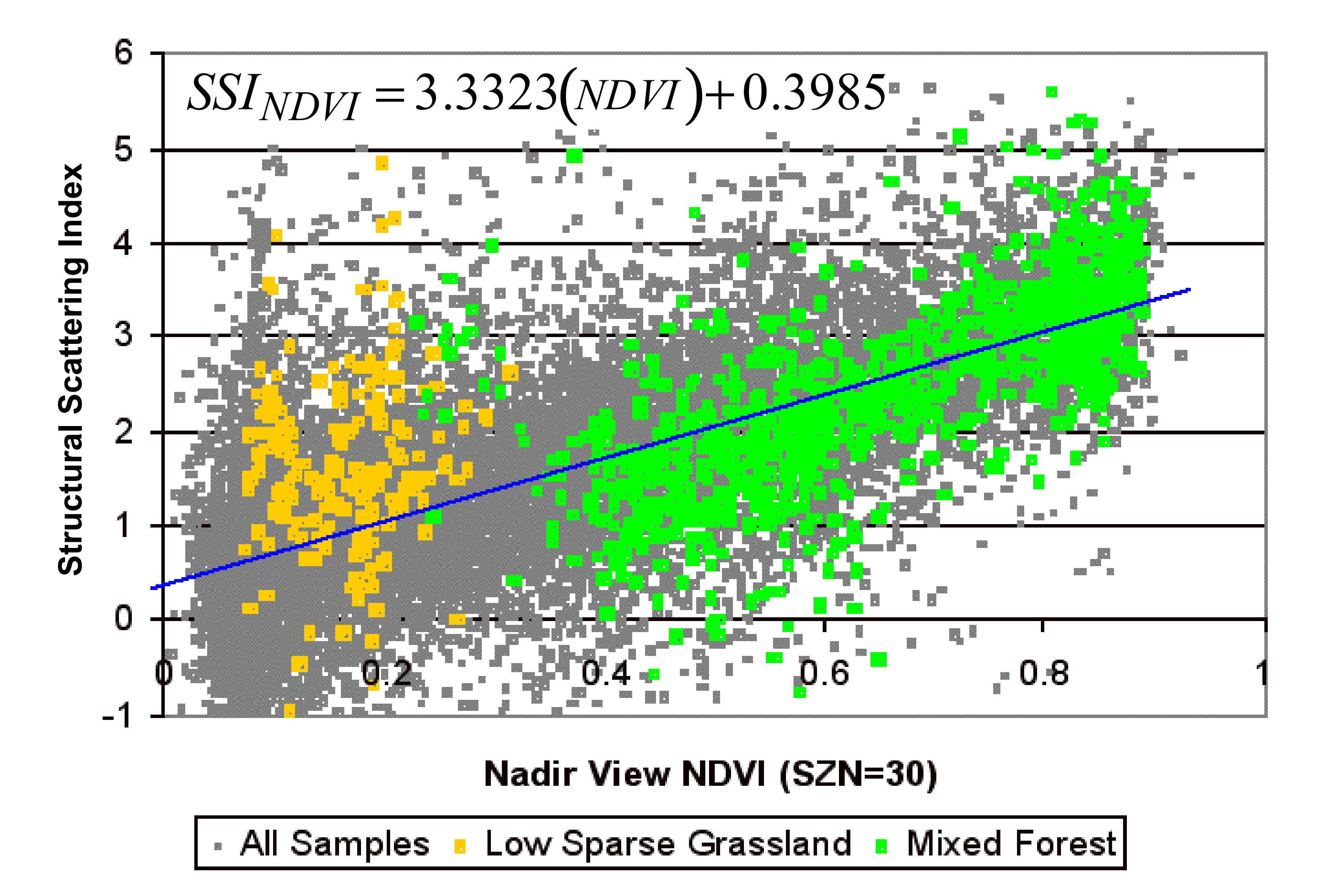
This index correlates well with NDVI as the importance of NIR volume scattering increases and red geometric scattering decreases with increasing leaf cover.

$$SSI_{NDVI} = 3.3323(NDVI) + 0.3985$$

But the deviation of *SSI* from SSI_{NDVI} is related to cover type, so we define the Relative Structural Scattering Index *RSSI* as:

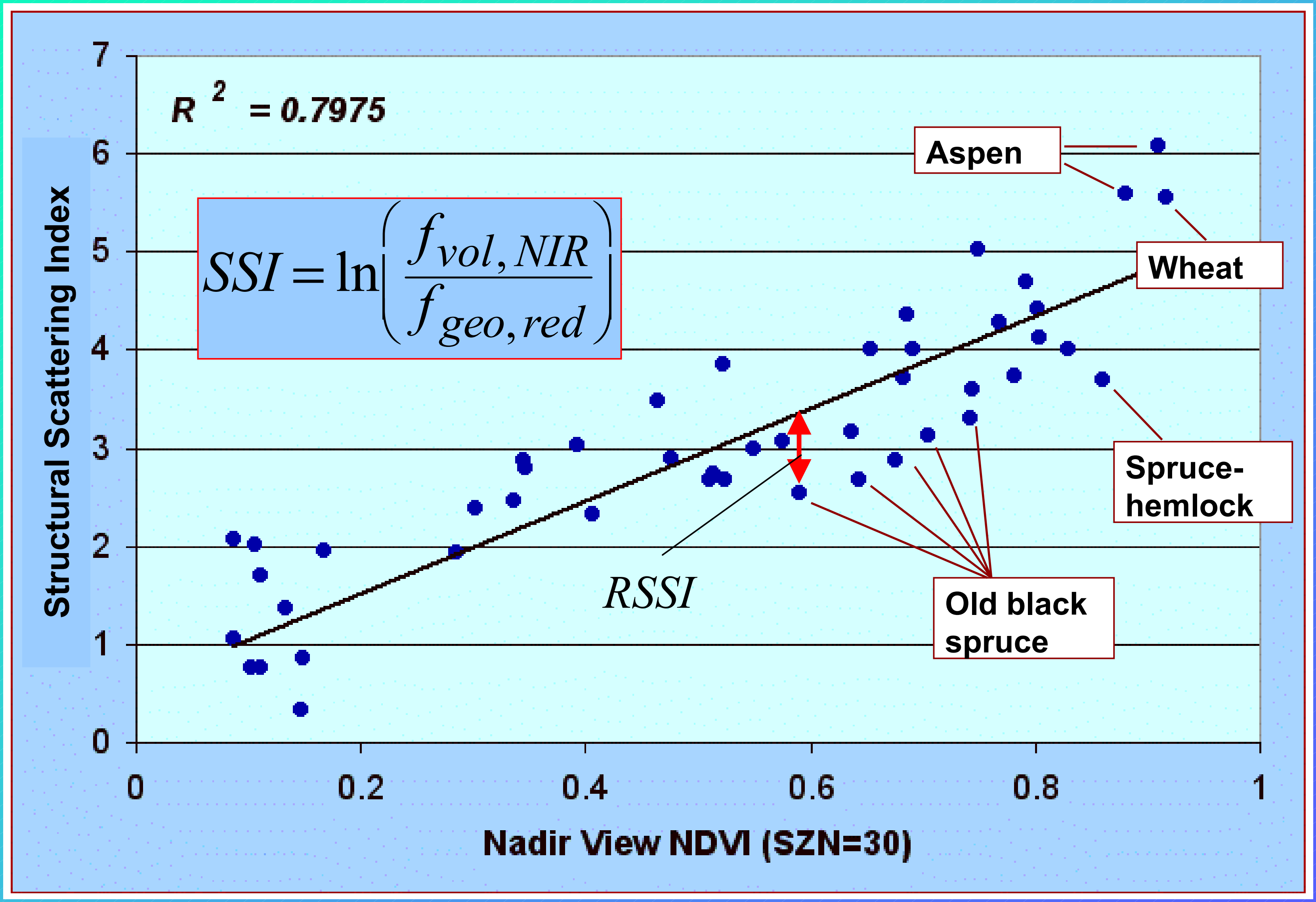
$$RSSI = SSI - SSI_{NDVI}$$

SSI vs. NDVI



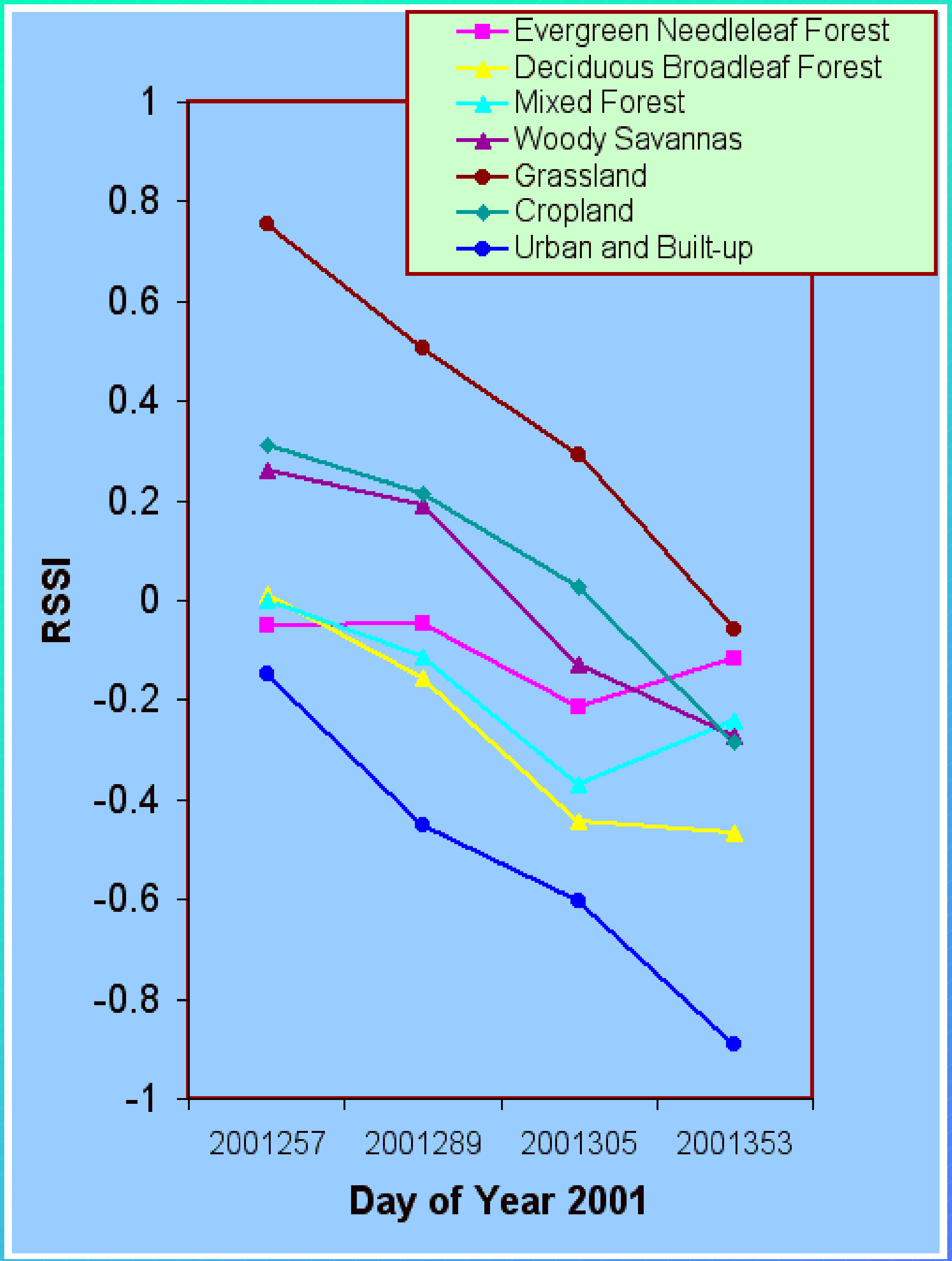
$SSI_{NDVI} = 3.3323(NDVI) + 0.3985$
Structural Scattering Index
Nadir View NDVI (SZN=30)
6
5
4
3
2
1
0
-1
0
0.2
0.4
0.6
0.8
1
All Samples
Low Sparse Grassland
Mixed Forest

RSSI

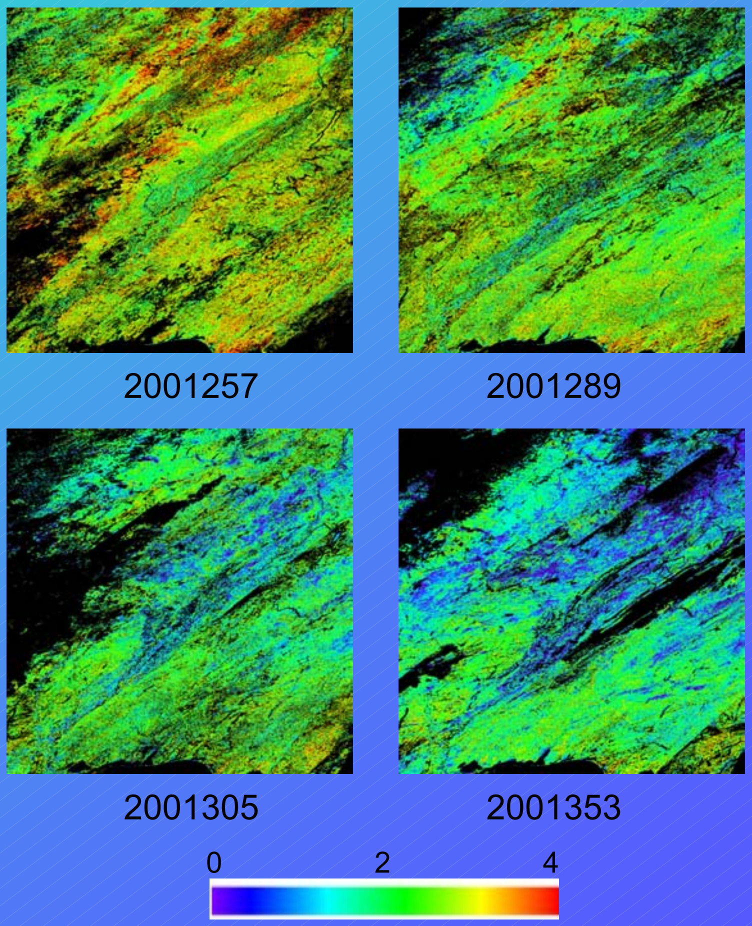


(Ground-measured BRDF Database)

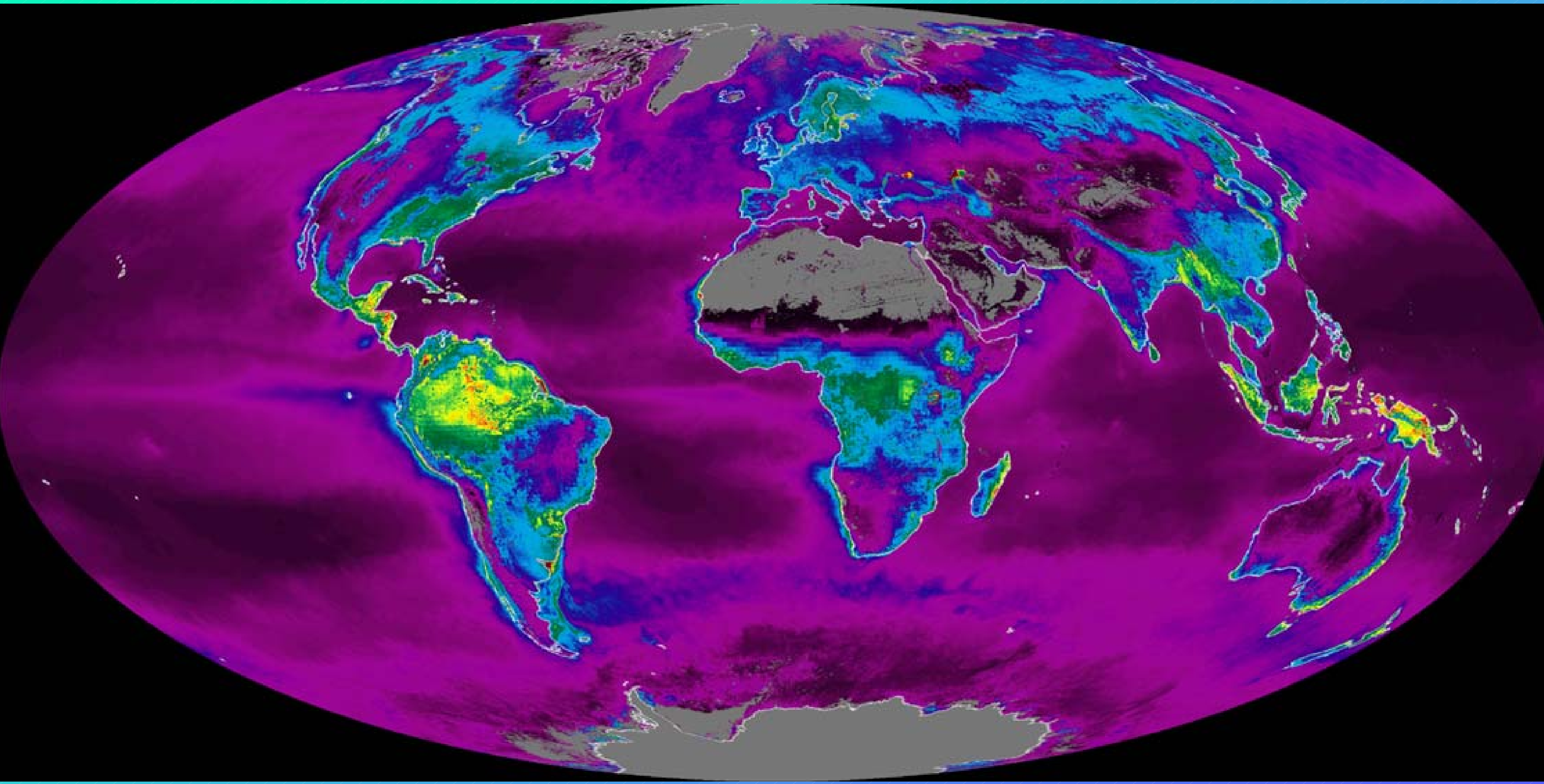
RSSI by IGBP Vegetation Type and Date



SSI maps of a MODIS land tile (h10v05, Southeast United States) for four 16 day periods



Ocean and Land Net Primary Productivity from MODIS



2002

Net Primary Productivity (kg C m-2 yr-1)

0 1 2 3

Wrap-Up

- MODIS products from Boston University for application to terrestrial carbon modelling include:
  - **Land Cover**
    - General biophysical parameterization based on vegetation cover and structure
  - **Phenology**
    - Timing of phenological cycle break points by region and land cover type for model parameterization
  - **Leaf-Area Index/Fraction of Photosynthetically-Active Radiation**
    - Annual development of green biomass for carbon fixation
  - **BRDF/Vegetation Structure**
    - Retrieval of vegetation structure for discrimination of land cover types

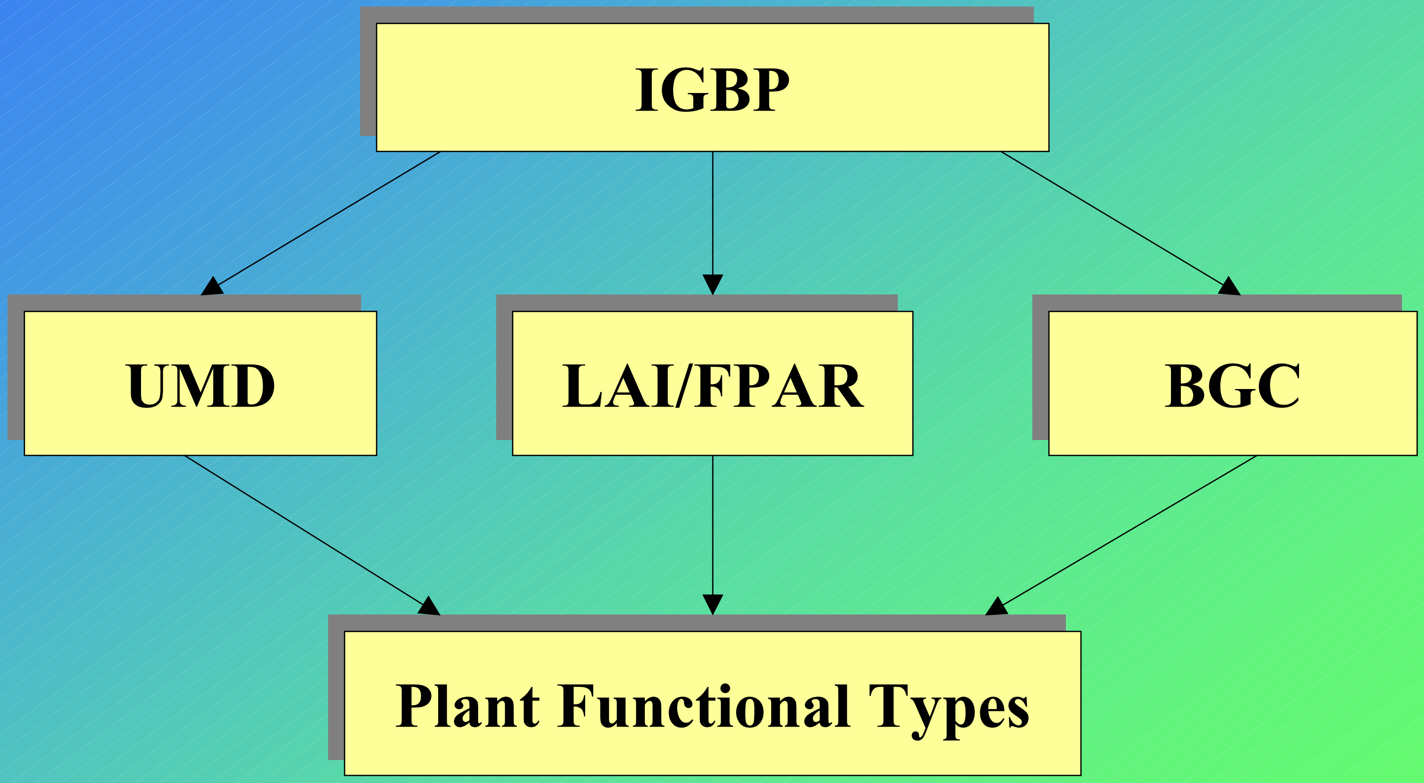
The MODIS BRDF/Albedo Product: Global Bidirectional Reflectance, Land Surface Albedo, and Nadir BRDF-Adjusted Reflectance

Crystal B. SCHAAF, Alan H. STRAHLER, Feng GAO, Yufang JIN, Wolfgang LUCHT, Xiaowen LI, Xiaoyang ZHANG, Elena TSVETSINSKAYA, Jan-Peter MULLER, Michael BARNSLEY, Philip LEWIS, Gareth ROBERTS, Christopher DOLL, Shunlin LIANG, David ROY, and Jeffrey PRIVETTE

Center for Remote Sensing, Boston University, 725 Commonwealth Avenue, Boston, MA 02215, USA
Tel. +617-358-0503, Fax +617-353-3200, e-mail:schaaf@bu.edu.
Potsdam-Institut für Klimafolgenforschung, Postfach 601203, D-14412 Potsdam, Germany
Research Center for Remote Sensing, Beijing Normal University, Beijing, China
Depart. of Geomatic Engineering, University College London, Gower St., London, WCIE 6BT, UK
Department of Geography, University College London, Gower St., London, WCIE 6BT, UK
Department of Geography, University of Wales Swansea, Singleton Park, Swansea, SA2 8PP, UK
Department of Geography, University of Maryland, College Park, MD 20742, USA
NASA/Goddard Space Flight Center, Code 923, Greenbelt, MD 20771, USA

MOD12Q1: What Is It?

- *Five Consistent Layers of Land Cover Class Labels*



The Land Cover Input Database

- ***242 Features From MODIS:***
  - *Temporal and spectral information; 16-day composites*
- ***Uses Surface Reflectance (NBAR)***
  - *View-angle corrected surface reflectance, 7 land bands*
- ***And Enhanced Vegetation Index (EVI)***
- ***Plus (in the future)….***
  - ***Spatial Texture from 250-m Band 2***
    - **Standard deviation-to-mean ratio in Band 2 (near-infrared)**
  - ***Snow Cover***
    - **MODIS Snow Cover Product, number of days with snow cover**
  - ***Land Surface Temperature***
    - **MODIS Land Surface Temperature, maximum value composite**
  - ***Directional Information***
    - **Bidirectional reflectance information from BRDF product**

Two-Stage Processing Strategy

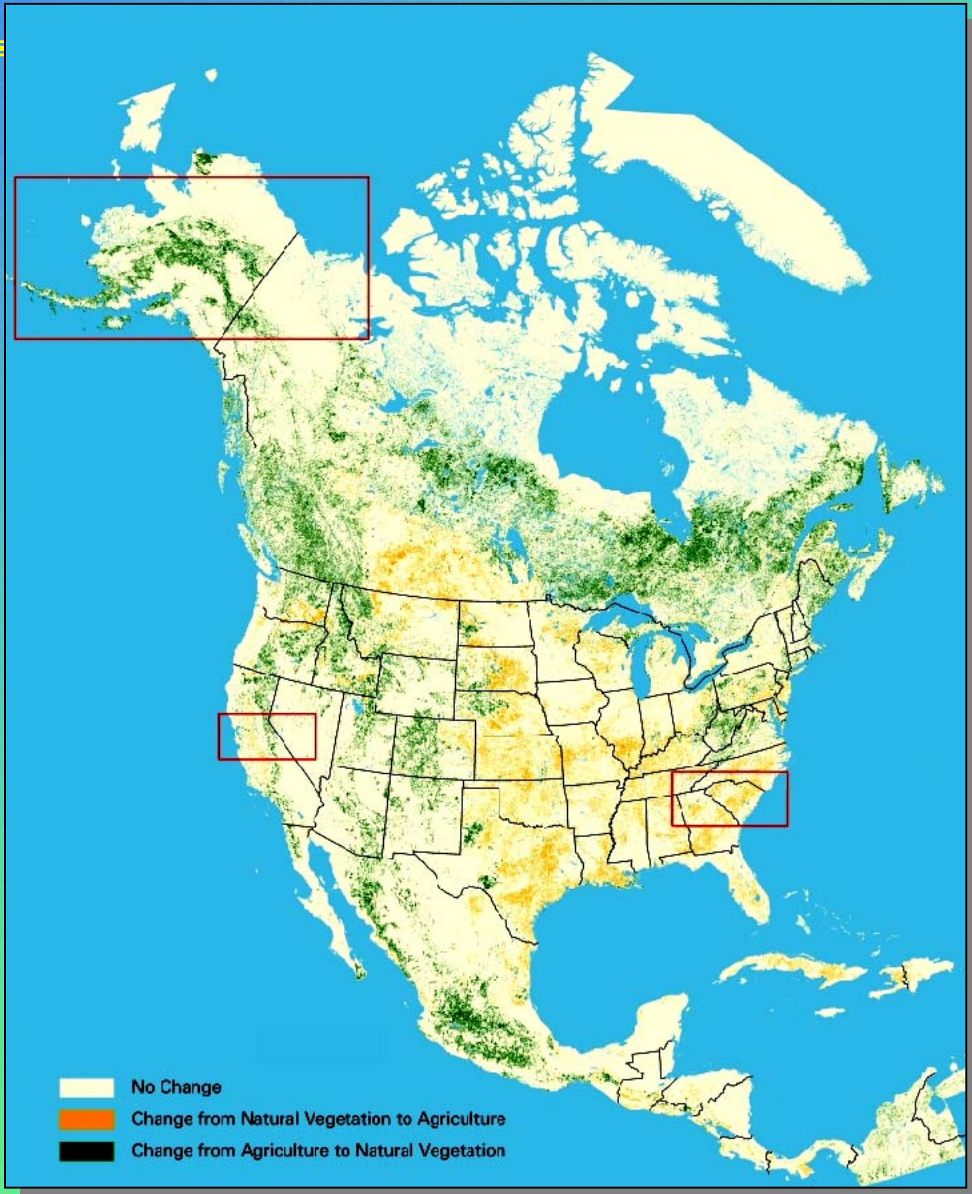
- *Monthly (32-day) Database Assembly Program*
  - Reads and processes spectral, spatial, directional inputs for each 32-day month
    - 32-day period = two 16-day, four 8-day cycles of MODIS input products
- *Quarterly (128-day) Classification with Annual Input (384-day)*
  - Reduces data volume by retaining only selected features
  - Classifies using advanced technology classifier (decision tree)
  - Runs every three months (96 days) to provide a quarterly updated product
- *Why Quarterly?*
  - Input data changes
  - Algorithm and processing improves

Advanced Technology Classifiers

- ***Supervised Mode***
  - **Use of supervised mode with training sites**
  - **Allows rapid reclassifications for tuning**
- ***Decision Trees*—C4.5 Univariate Decision Tree**
  - **Fast algorithm**
  - **Uses boosting to create multiple trees and improve accuracy, estimate confidence**
- ***Neural Networks*—Fuzzy ARTMAP**
  - **Uses Adaptive Resonance Theory in building network**
  - **Presently not in use. Too slow; does not handle missing data well.**

Priors: Agriculture

- **Problem:**
  - **Too much agriculture**
- **Solution:**
  - **Used available maps of agricultural intensity to parameterize likelihood of agriculture over natural vegetation**
- **Result:**
  - **6% of land area changed to natural vegetation**
  - **3% of land area changed to agriculture**

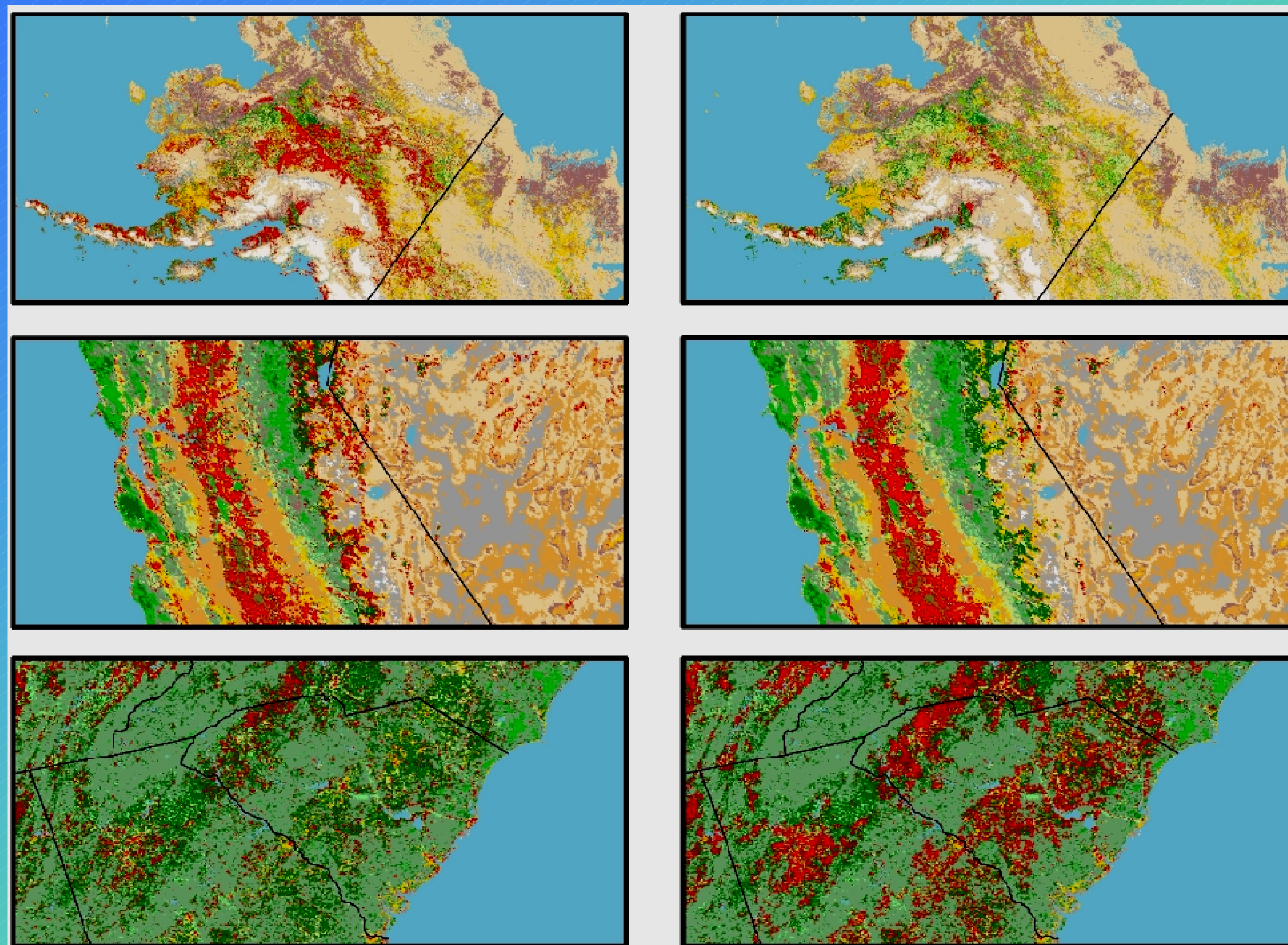


Regional Views

Alaska

California

South Carolina



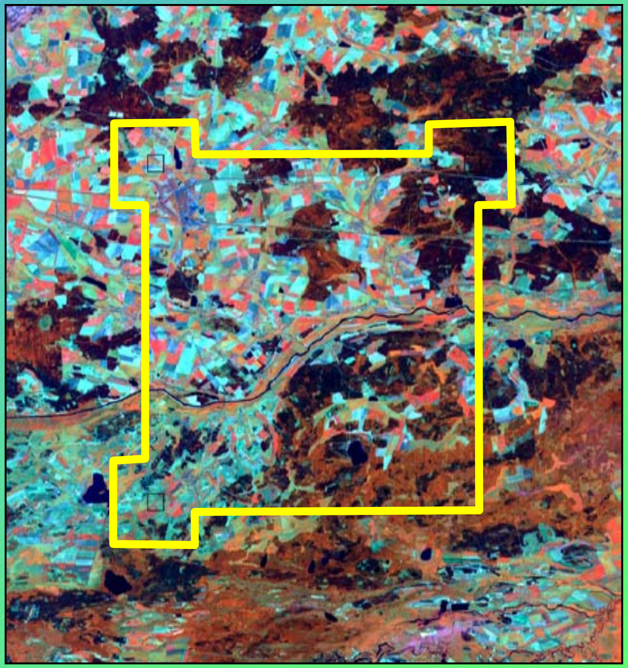
(Ag in red)

Before

After

Test Sites

- *IGBP-DISCover Core/Confidence Sites*
  - Random stratified sampling of classes on 1992-93 IGBP Global Land Cover Product
  - 425 sites identified; 413 SPOT and TM scenes acquired; 91% migrated to WWW by BU
- *BU STEP Database*
  - >2500 training sites from >700 TM scenes
  - About 1500 training sites in current use for supervised classification



A confidence site near Pinsk, Belarus (20 x 20 km)

Global Land Cover Products Released

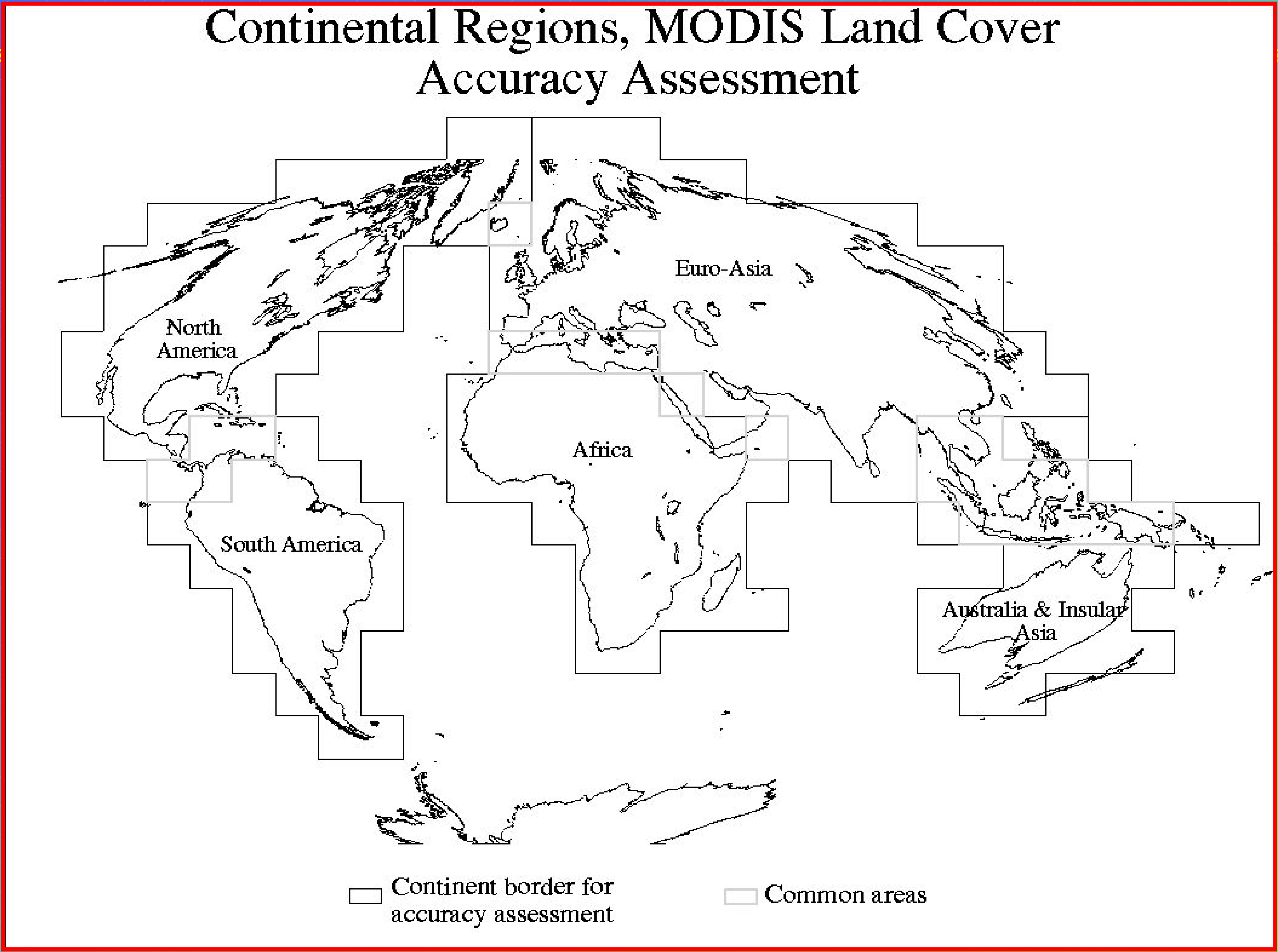
- *Beta Product, released April 15, 2001*
  - Based on 2 16-day periods of Nadir BRDF-Adjusted Reflectance (NBARs)
- *Provisional Product 2001001, released June 15, 2001*
  - Based on 9 16-day periods of NBARs within July 11– January 15, 2001
  - Used prior probabilities to help separate agriculture and natural vegetation
  - Included IGBP classification, secondary classes, confidence measures
  - Filled from at-launch product (EDC DISCover v.2) when only 0–2 views were available or when classification confidence was less than 40%

Global Site and Pixel Count

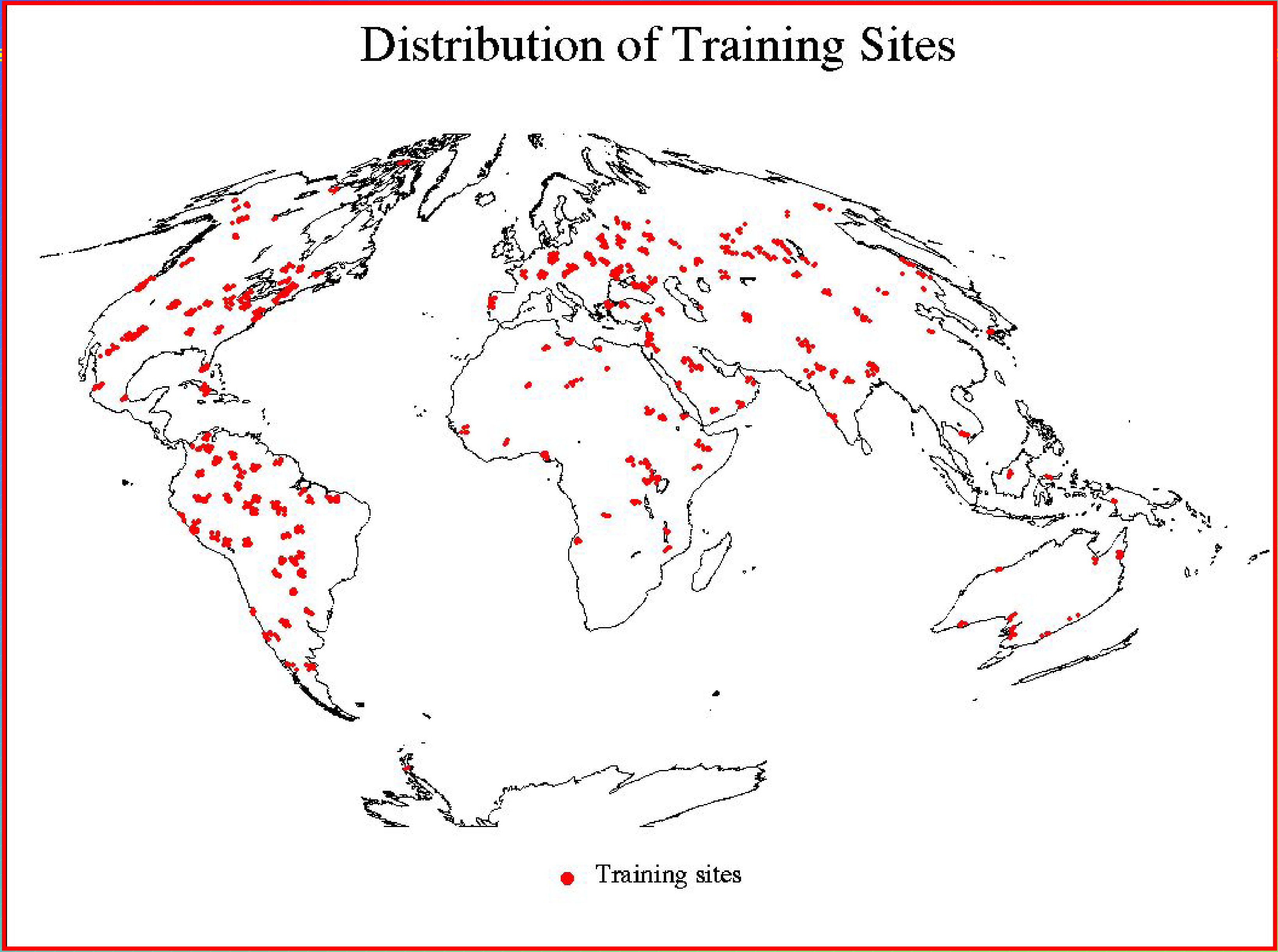
Table 1. Global counts of sites and pixels by land cover class.

IGBP Land Cover Class	Training Site Count	Training Pixel Count	Global Pixels Classified	Global Areal Percentage
1. Evergreen Needleleaf	131	2,056	7,100,847	3.9
2. Evergreen Broadleaf	204	5,409	17,583,346	9.7
3. Deciduous Needleleaf	15	261	2,374,908	1.3
4. Deciduous Broadleaf	57	758	2,016,765	1.1
5. Mixed Forest	96	2,077	8,209,766	4.5
6. Closed Shrubland	20	466	1,068,970	0.6
7. Open Shrubland	87	1,679	31,929,221	17.8
8. Woody Savanna	55	1,167	10,702,581	5.9
9. Savanna	44	1,098	11,218,832	6.2
10. Grasslands	87	1,474	12,363,432	6.8
11. Permanent Wetlands	13	289	559,675	0.3
12. Cropland	263	6,240	17,087,489	9.4
14. Cropland/Nat Veg Mosaic	72	1,447	5,660,478	3.1
15. Snow and Ice	10	1,346	16,501,715	9.1
16. Barren/Sparse	108	4,492	21,977,613	12.2
17. Water	63	9,213	14,575,749	8.1
Total	1,370	39,472	180,928,968	100.0

Continental Regions



Continental Regions



Site and Pixel Counts by Region

Table 2. Site and pixel counts by region.

IGBP Land Cover Class	Training Site Count	Training Pixel Count	Global Pixels Classified	Global Areal Percentage
Global	1,370	39,472	180,928,968	100.0
North America	368	13,731	30,918,663	17.1
South America	321	8,030	22,181,052	12.3
Eurasia	560	13,290	71,275,640	39.4
Africa	194	5,744	38,711,576	21.4
Australia-Insular Asia	46	1,766	18,046,575	10.0

Overall Accuracies

- *Proper accuracy statements require proper statistical sampling*
- *AVHRR state of the art has been 60–70 percent, depending on class and region*
- *MODIS accuracies are falling in 70–80 percent range*
- *Most “mistakes” are between similar classes*
- *Land cover change should NOT be inferred from comparing successive land cover maps*

References

Friedl, M.A., D. Muchoney, D.K. McIver, A.H. Strahler, and J.C.F. Hodges 2000: Characterization of North American land cover from AVHRR Data, Geophysical Research Letters*, vol. 27, no. 7, pp. 977-980.*

Friedl, M.A., C. Woodcock, S. Gopal, D. Muchoney, A.H.Strahler, and C. Barker-Schaaf 2000. A note on procedures used for accuracy assessment in land cover maps derived from AVHRR data, International Journal of Remote Sensing, *vol. 21, pp.1073-1077.*

Friedl, M.A., Brodley, C.E. and A.H. Strahler 1999: Maximizing land cover classification accuracies at continental to global scales, IEEE Transactions on Geoscience and Remote Sensing, *vol. 37, pp. 969-977*.

Friedl, M.A. and C.E. Brodley 1997: Decision tree classification of land cover from remotely sensed data, Remote Sensing of Environment, *vol. 61, pp. 399-409*.

Mciver, D.K. and M.A. Friedl 2002. Using prior probabilities in decision-tree classification of remotely sensed data, Remote Sensing of Environment, *Vol. 81, pp. 253-261.*

McIver, D.K. and M.A. Friedl 2001. Estimating pixel-scale land cover classification confidence using non-parametric machine learning methods, IEEE Transactions on Geoscience and Remote Sensing. *Vol 39(9), pp. 1959-1968.*

Muchoney, D., Borak, J, Chi, H., Friedl, M.A., Hodges, J. Morrow, N. and A.H. Strahler 1999: Application of the MODIS global supervised classification model to vegetation and land cover mapping of Central America, International Journal of Remote Sensing, *Vol 21, no 6 & 7, pp. 1115-1138.*

Muchoney, D. M., and Strahler, A. H., 2001, Pixel and site-based calibration and validation methods for evaluating supervised classification of remotely sensed data, Remote Sens. Environ., *in press.*

Kernel Models

- Kernel model derivation
  - ***To derive a kernel model, we simplify and manipulate a complete physical model until it reaches the form***

$$R = c_1 k + c_2$$

  - ***where k is a function only of view and illumination geometry (and fixed physical parameters)***
  - ***c_1 and c_2 are constants containing variable physical parameters***

Volume-Scattering Kernels

- Ross Kernels
  - ***Derived by Roujean and Li from theory by Ross***
  - ***Single-scattering approximation***
  - ***Homogeneous leaf canopy***
  - ***Uniform leaf angle distribution***
  - ***Leaf reflectance = leaf transmittance***
  - ***Lambertian background***
- Ross-thick (used in MODIS algorithm)
  - ***Approximation for high leaf area index, LAI >> 1***
- Ross-thin
  - ***Approximation for low leaf area index, LAI << 1***

Geometric-Optical Kernels

- Li Kernels
  - ***Based on geometric-optical BRDF models of Li et al.***
  - ***Spheroidal objects casting shadows that fall on a Lambertian background or on other objects***
  - ***Reflectance then depends on areal proportions of sunlit and shaded object and background as seen at varying view angles***
  - ***Driven by shape, height, and spacing of spheroids***
  - ***Random location of spheroids on the plane***
- Li-sparse (used in MODIS algorithm)
  - ***Spheroids are sparse and shadows fall on background***
- Li-dense
  - ***Spheroids are dense and mutually shadow each other***