globally leveraged integrated data explorer for research

Visualize and Analyze MODIS Imagery using GLIDER Tool

Dr. Rahul Ramachandran

Information Technology and Systems Center

University of Alabama Huntsville

rramachandran@itsc.uah.edu

http://miningsolutions.itsc.uah.edu/glider/

A data/tools workshop on GLIDER and HYDRA
June 2010,Taipei, Taiwan

UAHuntsville THE UNIVERSITY OF ALABAMA IN HUNTSVILLE

Goals

- Explore "information extraction" from remotely sensed data
  - Look at one of the most popular NASA datasets – MODIS
  - Learn about the GLIDER tool

UAHuntsville THE UNIVERSITY OF ALABAMA IN HUNTSVILLE

Remote Sensing Process

- Definition – systematic data collection and analysis procedures used for Earth Science application
- Focus on analysis – information extraction!

UAHuntsville
THE UNIVERSITY OF ALABAMA IN HUNTSVILLE

Remote Sensing Data: Basics

- A sensor is measuring electromagnetic radiance **L**

$$L = f\left(\lambda, s_{x,y,z}, t, \theta, P, \Omega\right)$$

- Wavelength or Frequency
- Location and size
- Temporal information – when and how often
- Angles that describe geometric relationships
- Polarization
- Radiometric resolution

MODIS – Level 1

MODIS Reflected Solar Bands

	Primary Use	Band No.	Bandwidth (nm)	Spectral Radiance	Required SNR
250 M	Land/Cloud Boundaries	1**	620-670	21.8	128
		2**	841-876	24.7	201
500 M	Land/Cloud Properties	3*	459-479	35.3	243
		4*	545-565	29.0	228
		5*	1230-1250	5.4	74
		6*	1628-1652	7.3	275
		7*	2105-2155	1.0	110
	Ocean Color/ Phytoplankton/ Biogeochemistry	8	405-420	44.9	880
		9	438-448	41.9	838
		10	483-493	32.1	802
		11	526-536	27.9	754
		12	546-556	21.0	750
		13	662-672	9.5	910
		14	673-683	8.7	1087
		15	743-753	10.2	586
		16	862-877	6.2	516
	Atmospheric Water Vapor	17	890-920	10.0	167
		18	931-941	3.6	57
		19	915-965	15.0	250

* 500m Spatial Resolution
** 250m Spatial Resolution

Spectral Radiance values are in W/m^2-um-sr
SNR = Signal-to-noise ratio

Radiances are converted to reflectances

MODIS – Level 1

MODIS Thermal Bands

Primary Use	Band	Bandwidth (μm)	Spectral Radiance	Required NEDT (K)
Surface/Cloud Temperature	20	3.660-3.840	0.45(300K)	0.05
	21	3.929-3.989	2.38(335K)	2.00
	22	3.929-3.989	0.67(300K)	0.07
	23	4.020-4.080	0.79(300K)	0.07
Atmospheric Temperature	24	4.433-4.498	0.17(250K)	0.25
	25	4.482-4.549	0.59(275K)	0.25
Cirrus Clouds	26	1.360-1.390	6.00	150 (SNR)
Water Vapor	27	6.535-6.895	1.16(240K)	0.25
	28	7.175-7.475	2.18(250K)	0.25
	29	8.400-8.700	9.58(300K)	0.05
Ozone	30	9.580-9.880	3.69(250K)	0.25
Surface/Cloud Temperature	31	10.780-11.280	9.55(300K)	0.05
	32	11.770-12.270	8.94(300K)	0.05
Cloud Top Altitude	33	13.185-13.485	4.52(260K)	0.25
	34	13.485-13.785	3.76(250K)	0.25
	35	13.785-14.085	3.11(240K)	0.25
	36	14.085-14.385	2.08(220K)	0.35

Spectral Radiance values are in W/m^2-um-sr
NEDT = Noise-equivalent temperature difference

Radiances are converted to temperature

Remote Sensing Data Analysis

- Radiometric Correction of Remote Sensor Data
  - Noise, error removal, calibration
- Geometric Correction of Remote Sensor Data
  - Map projections, Geographic Coordinate System
- Image Enhancement
  - Contrast stretching, Spatial/Frequency Filters, PCA
- Information Extraction
  - Parametric/Non parametric - classifiers
  - Heuristic based - indices

GLIDER - motivation

- Software tools that allow users to visualize, analyze and mine satellite imagery are currently limited.
- Available commercial packages are expensive.
- None of these packages provide all the GLIDER features

GLIDER Features

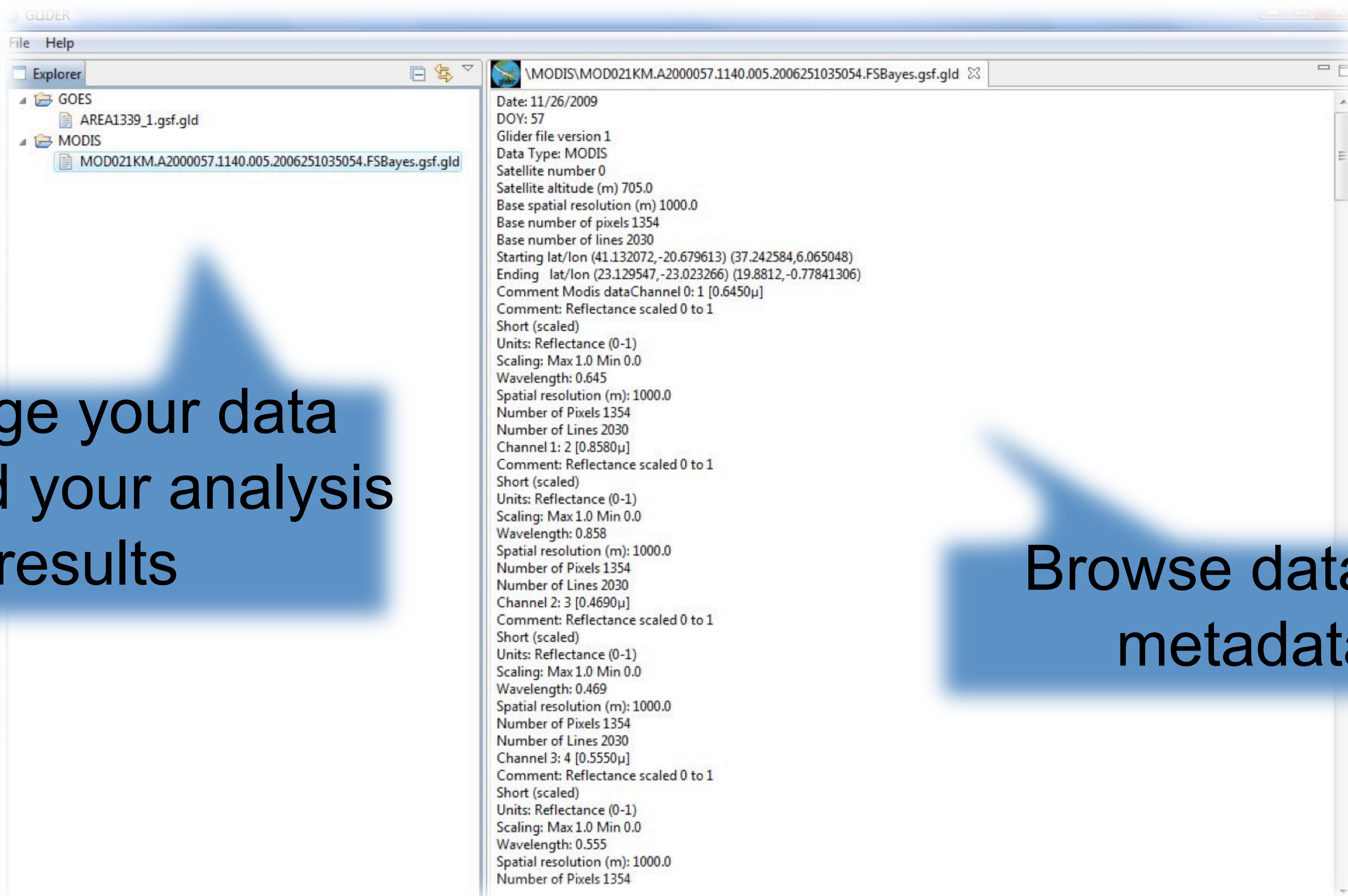
- *Visualize and analyze* satellite data in its native sensor view.
- Apply different *image processing algorithms* on the satellite data.
- Apply different *pattern recognition/data mining algorithms* on the satellite data.
- *Project* satellite data and analysis/mining results onto a globe and overlay additional layers.
- Provides *multiple views* to manage, visualize, and analyze satellite data.

UAHuntsville
THE UNIVERSITY OF ALABAMA IN HUNTSVILLE

GLIDER is using:

- ADaM
  - ADaM (Algorithm Development and Mining) toolkit
  - Contains 140+ image processing, pattern recognition and machine learning algorithms
- IVICS
  - Interactive Visualizer and Image Classifier for Satellites (IVICS)
  - Provides capability to visualize satellite imagery and select samples for supervised classification
- World Wind
  - Project satellite data and analysis/mining results onto a globe and overlay additional layers

GLIDER Views: Project Explorer



Manage your data files and your analysis results

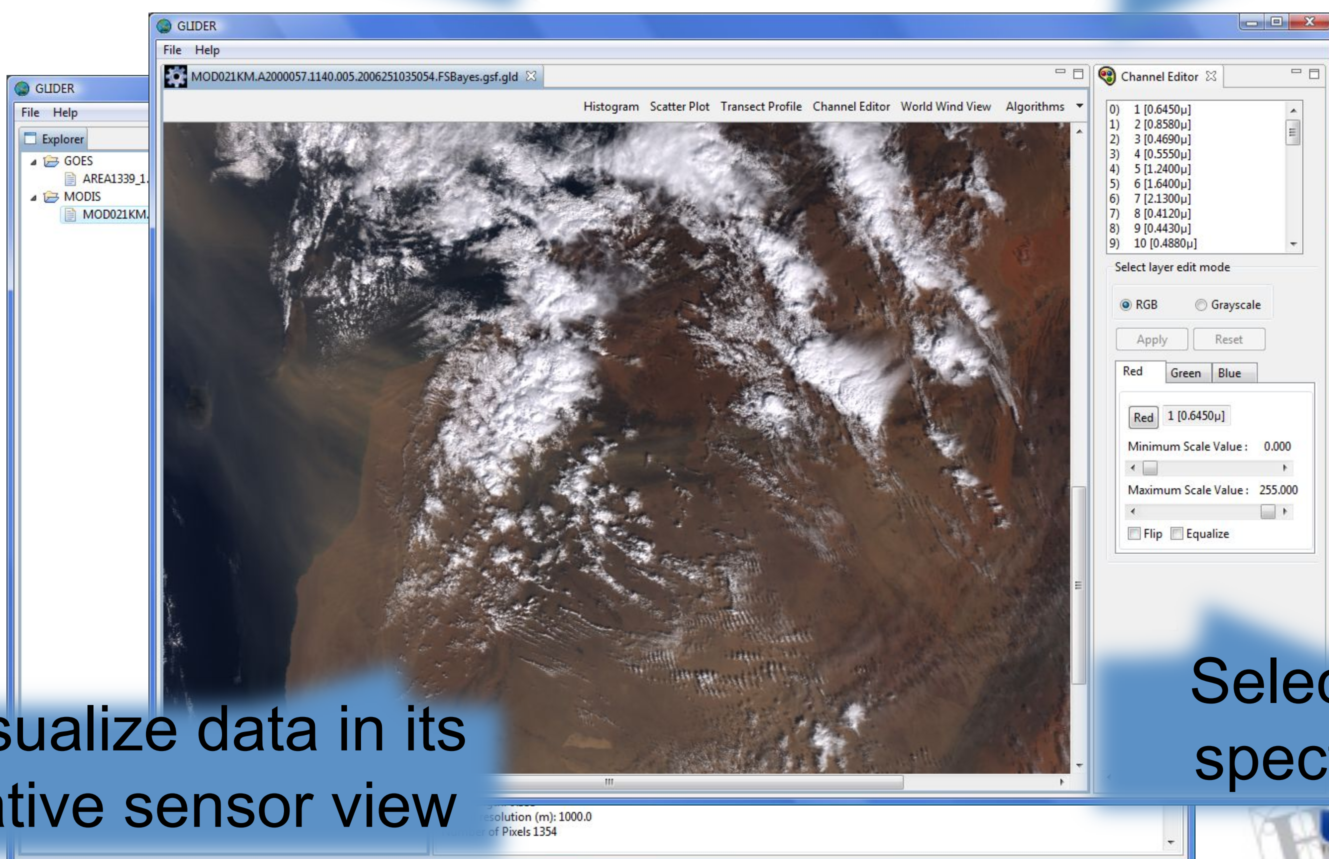
Browse data file metadata

UAHuntsville

GLIDER Views:
Image Analysis View

Analyze image using different features

Apply data mining algorithms

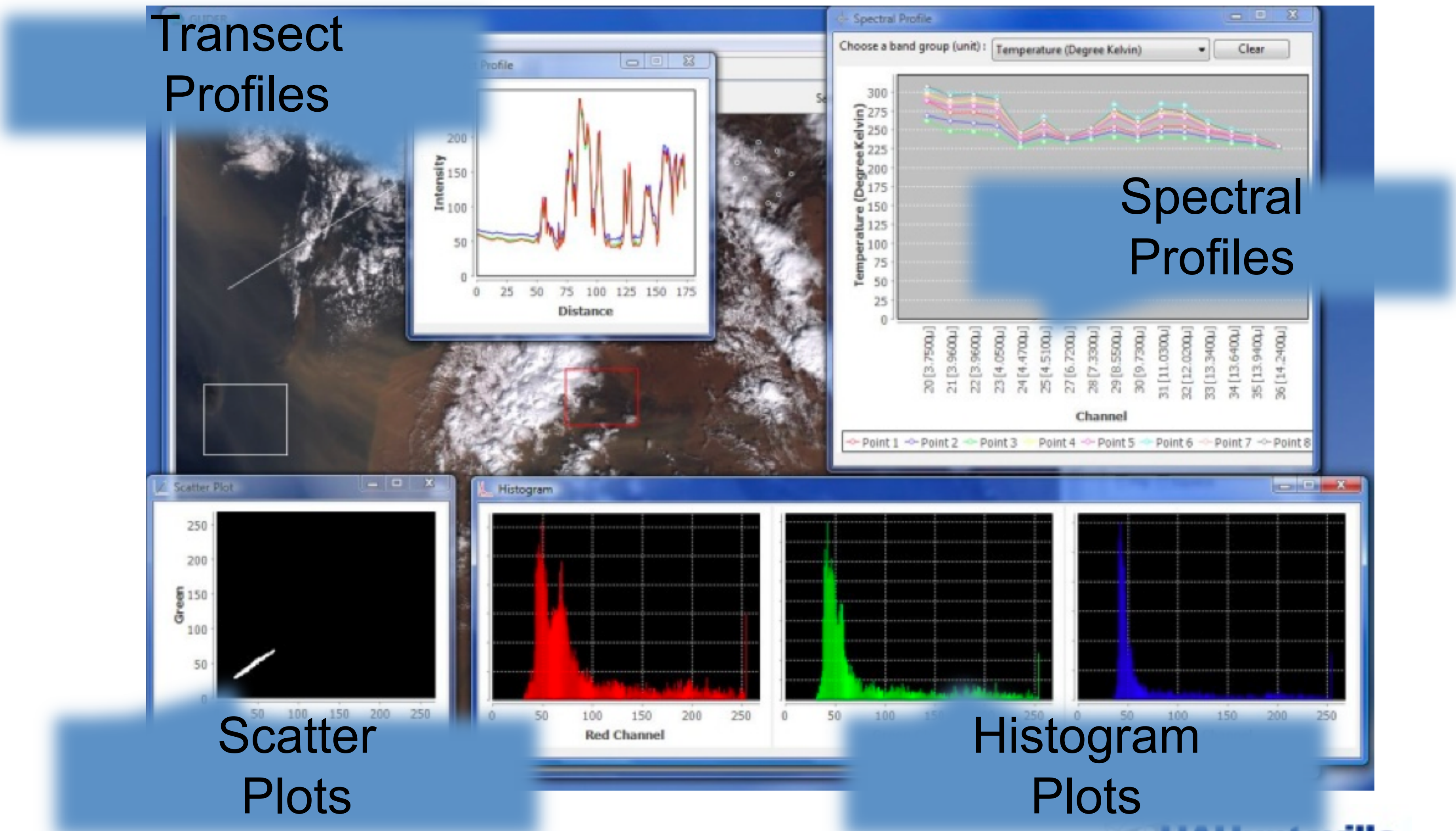


Visualize data in its native sensor view

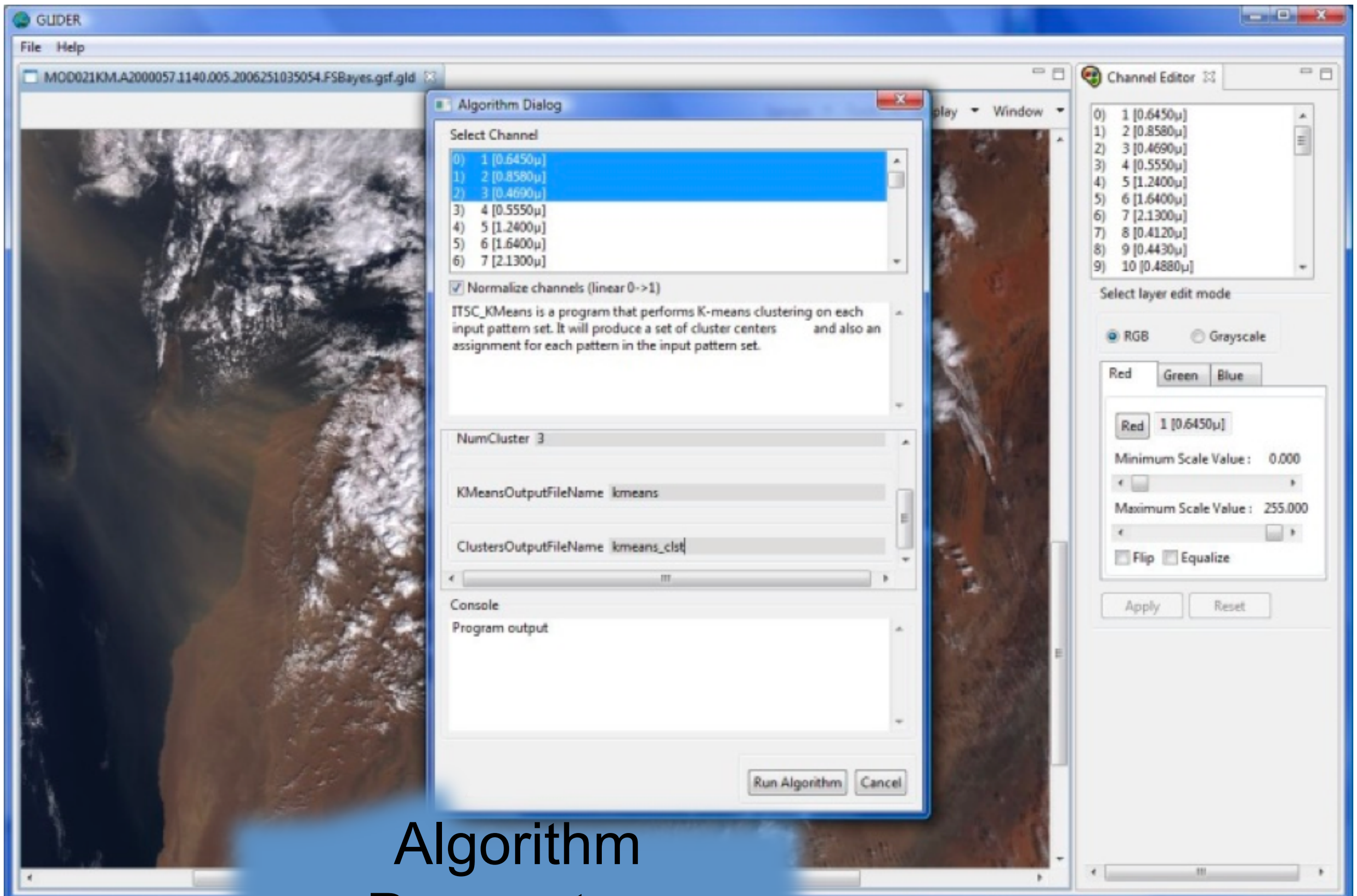
Select different spectral bands

UAHuntsville

Image Analysis Features



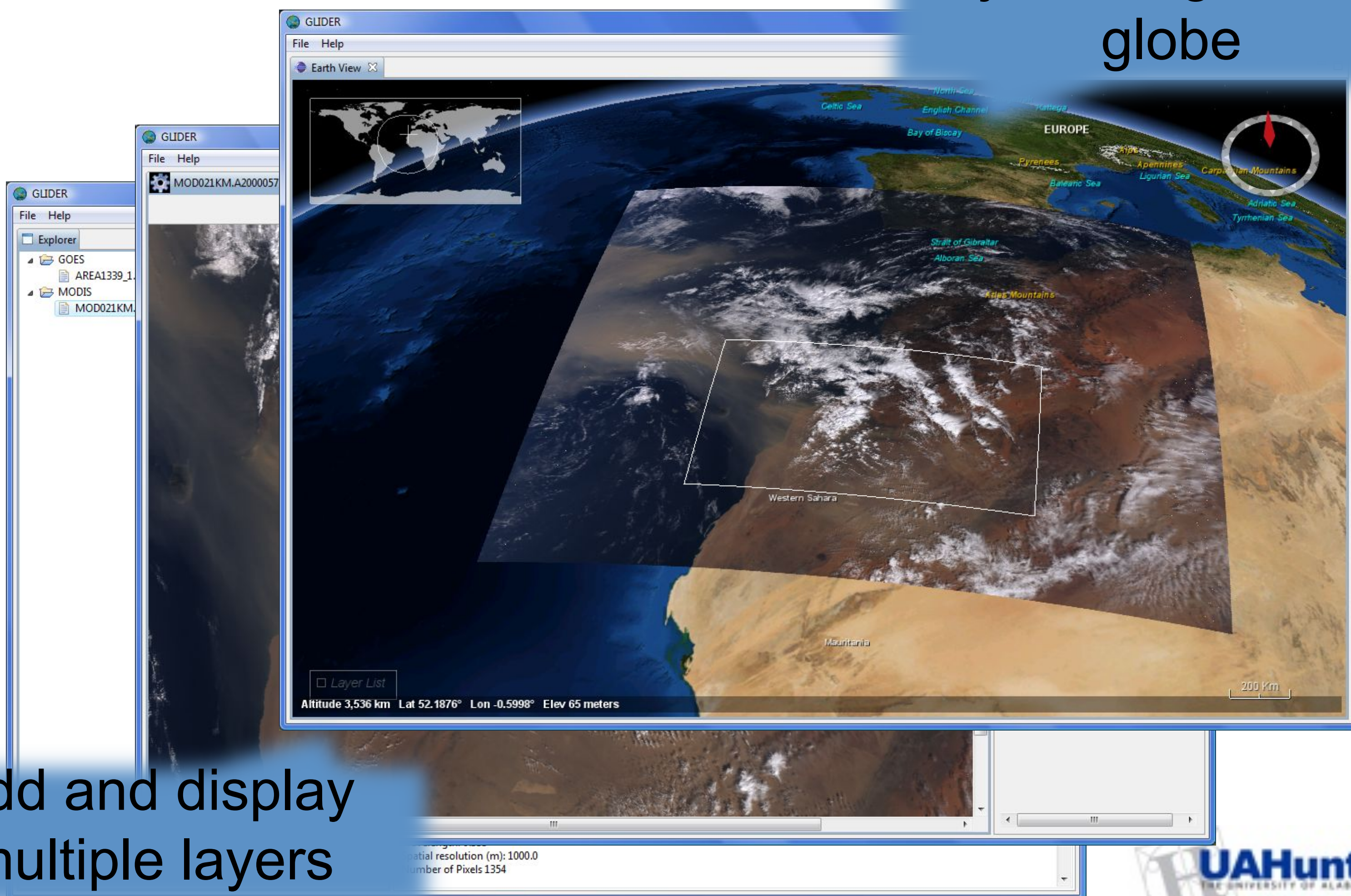
Clustering Algorithm Example



Algorithm Parameters

GLIDER Views:
Earth View

Project images on the globe

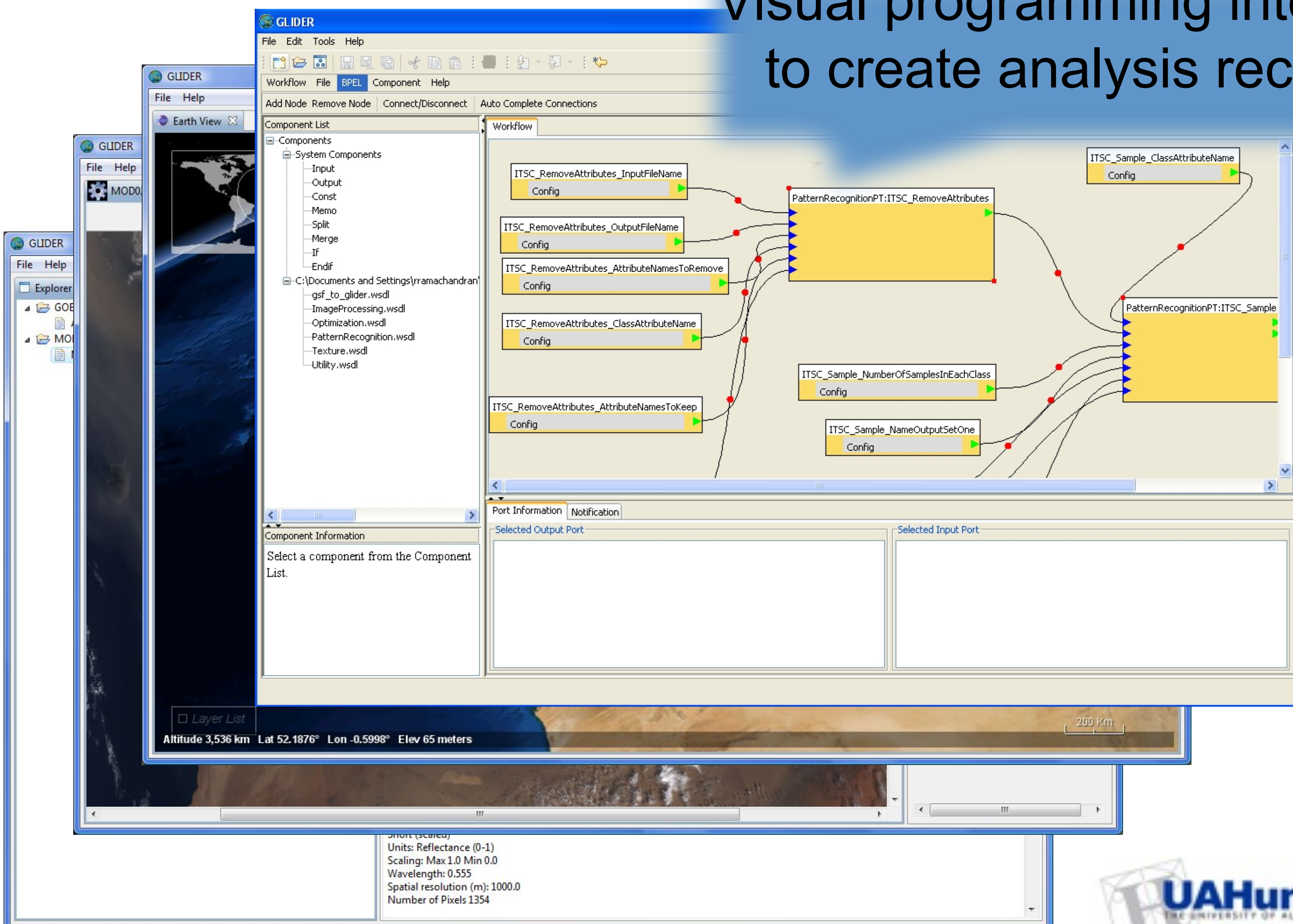


Add and display multiple layers

UAHuntsville
THE UNIVERSITY OF ALABAMA IN HUNTSVILLE

GLIDER Views: Workflow Composer

Visual programming interface to create analysis recipes



Learning Modules

- Module 1: “Midnight Oil”
  - Look at the Deepwater Horizon event
  - Learn basic GLIDER functionality while playing with several MODIS data files
- Module 2: “Smoke on the Water”
  - Learn how to create False color composites to visually separate features in MODIS data [Courtesy – Dr. Sundar Christopher, UAHuntsville]
- Module 3: “Dust in the wind”
  - Learn how to use band math feature in GLIDER to create indices
  - Use indices to detect dust
  - Browse a journal article and then apply the results from the paper to detect dust over china
- Module 4: “Ashes to Ashes” – Part 1
  - Look at Ash/Steam Plume event from Iceland's Eyjafjallajoekull Volcano
  - Learn how to subset imagery both spatially and spectrally
  - Apply clustering algorithm to generate classification maps
- Module 5: “Ashes to Ashes” – Part 2
  - Learn to construct a supervised classification process
  - Learn how take training samples
  - Create a *mining recipe/workflow* using visual programming

Pop Music Trivia – do you know what the titles of these modules refer to?

UAHuntsville
THE UNIVERSITY OF ALABAMA IN HUNTSVILLE

Learning Module 1

"Midnight Oil"

–Look at the Deepwater Horizon event

–Learn basic GLIDER functionality while playing with several MODIS data files

Deepwater Horizon

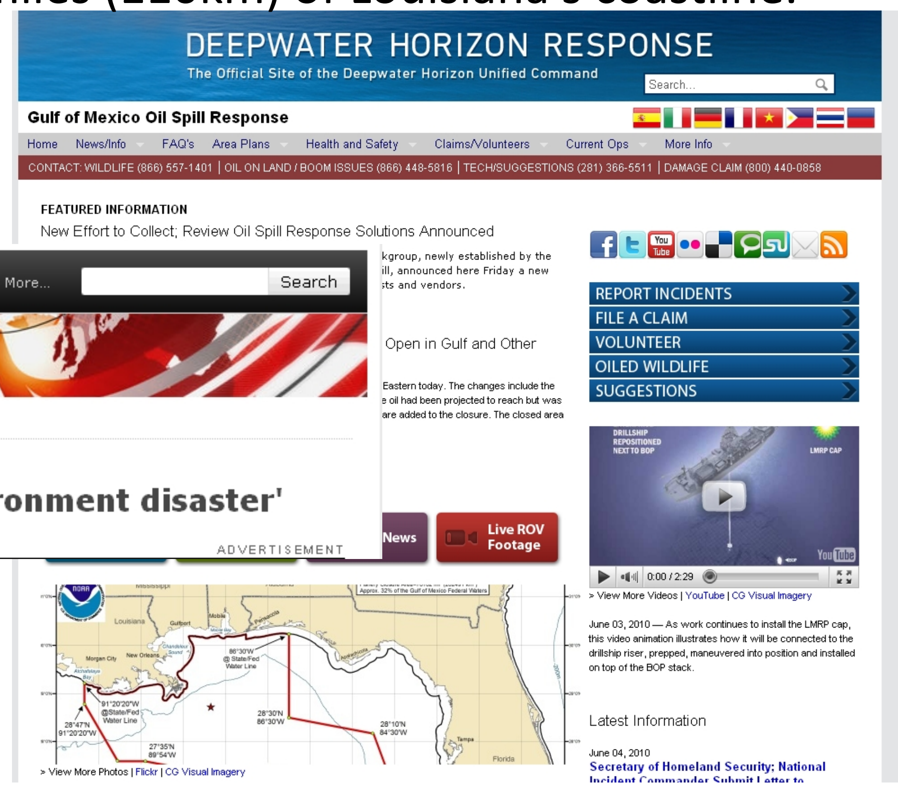
Deepwater Horizon

The DEEPWATER HORIZON is a Reading & Bates Falcon RBS8D design semi-submersible drilling unit capable of operating in harsh environments and water depths up to 8,000 ft (upgradeable to 10,000 ft) using 18¾in 15,000 psi BOP and 21in OD marine riser.

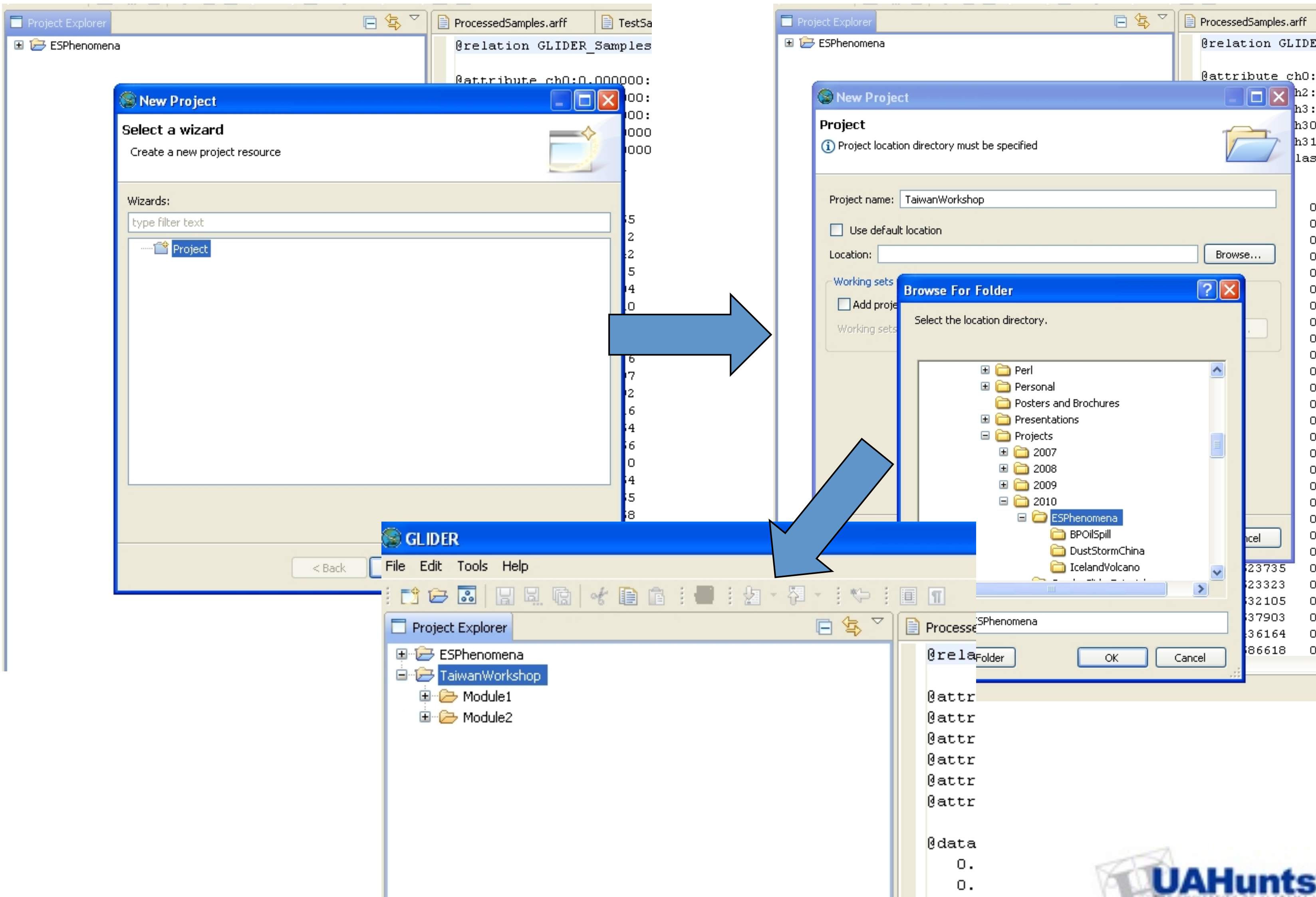
Rig Type	5th Generation Deepwater	
Design	Reading & Bates Falcon RBS-8D	
Builder	Hyundai Heavy Industries Shipyard, Ulsan, South Korea	
Year Built	2001	

Deepwater Horizon Event

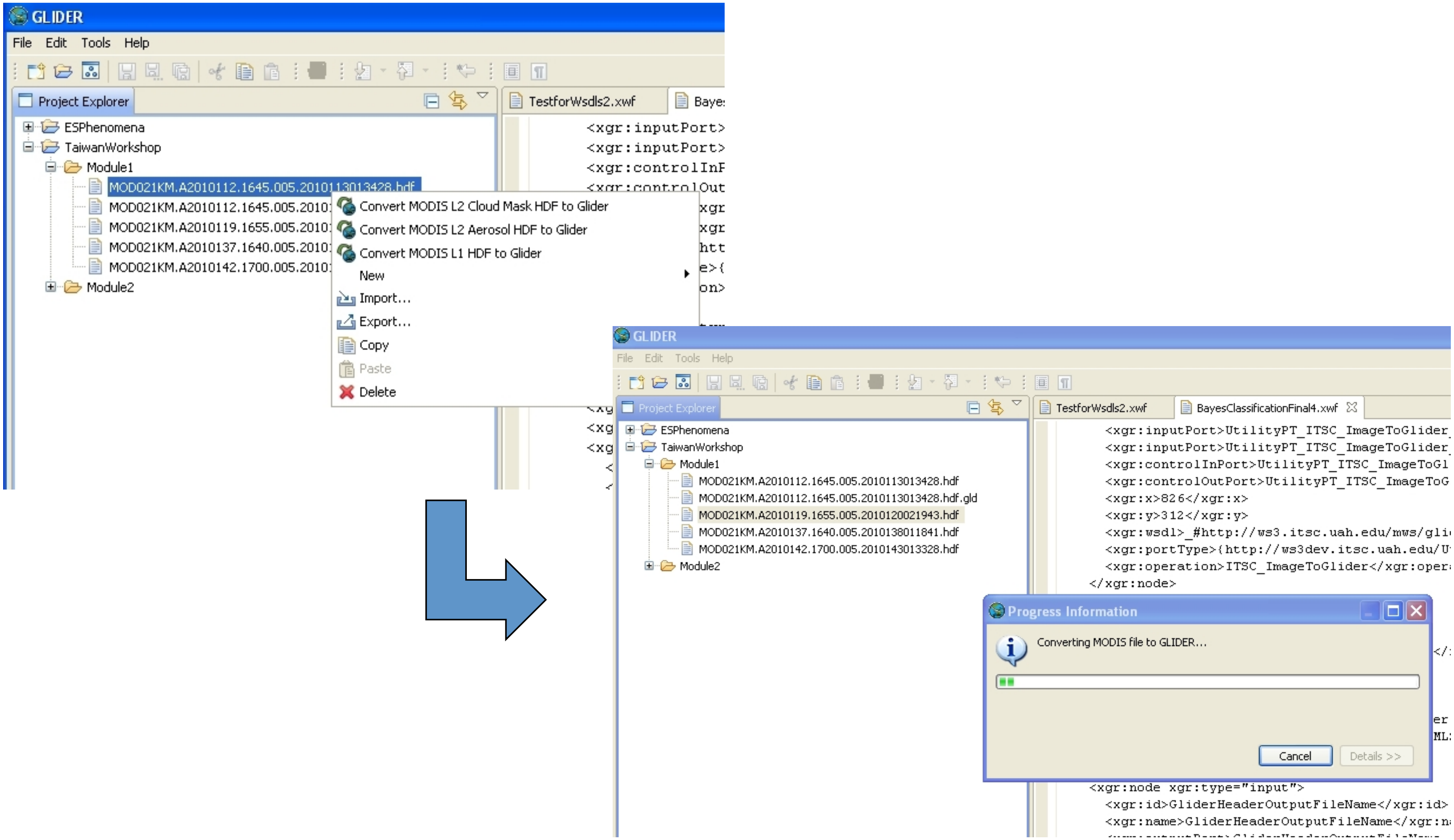
- April 20: At around 10 p.m. a fire is reported on the central time on the Deepwater Horizon rig. Eleven workers are killed
- At least 20 million gallons have now spilled into the Gulf of Mexico, affecting more than 70 miles (110km) of Louisiana's coastline.



Setting Up a Project in GLIDER



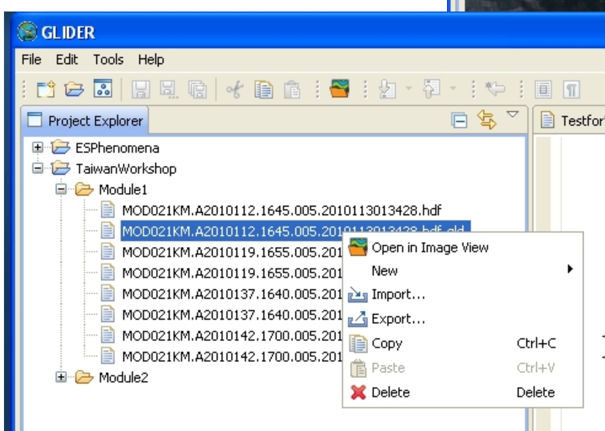
Convert L1 HDF to GLD File

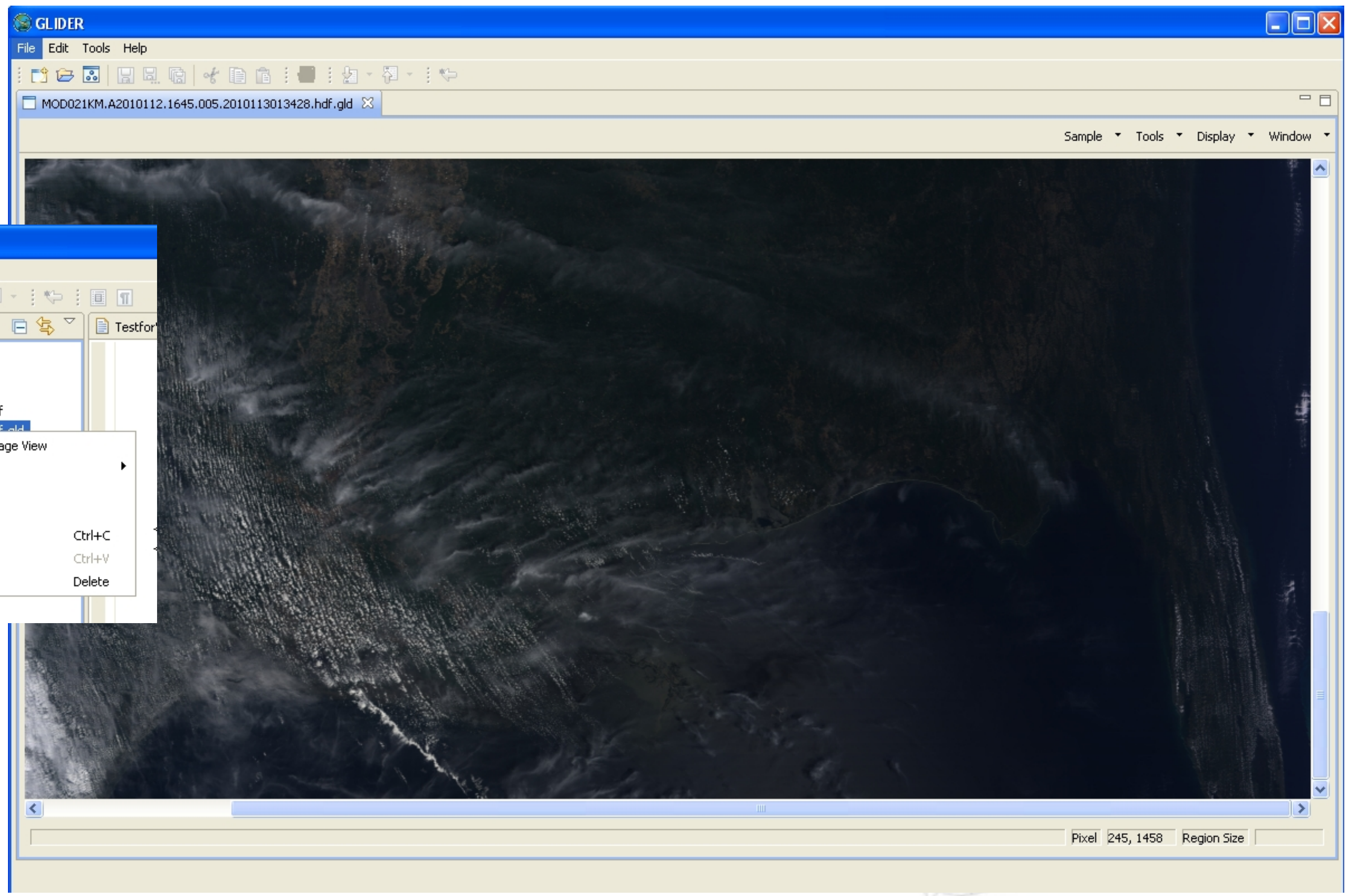


UAHuntsville
THE UNIVERSITY OF ALABAMA IN HUNTSVILLE

Image View

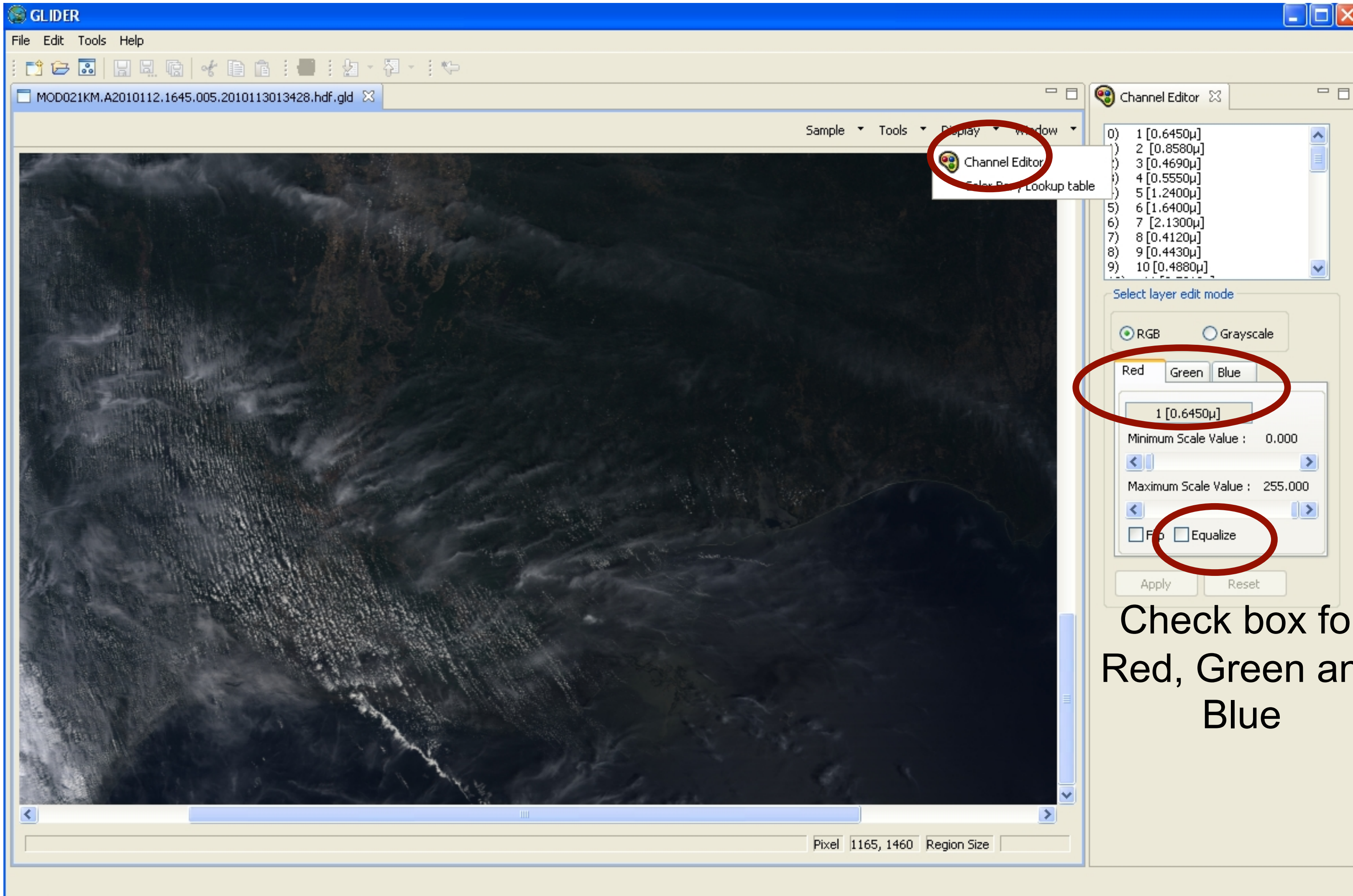
- Open MOD021KM.A2010112.1645.005.2010113013428.hdf.gld
- Smoke from the fire is clearly visible





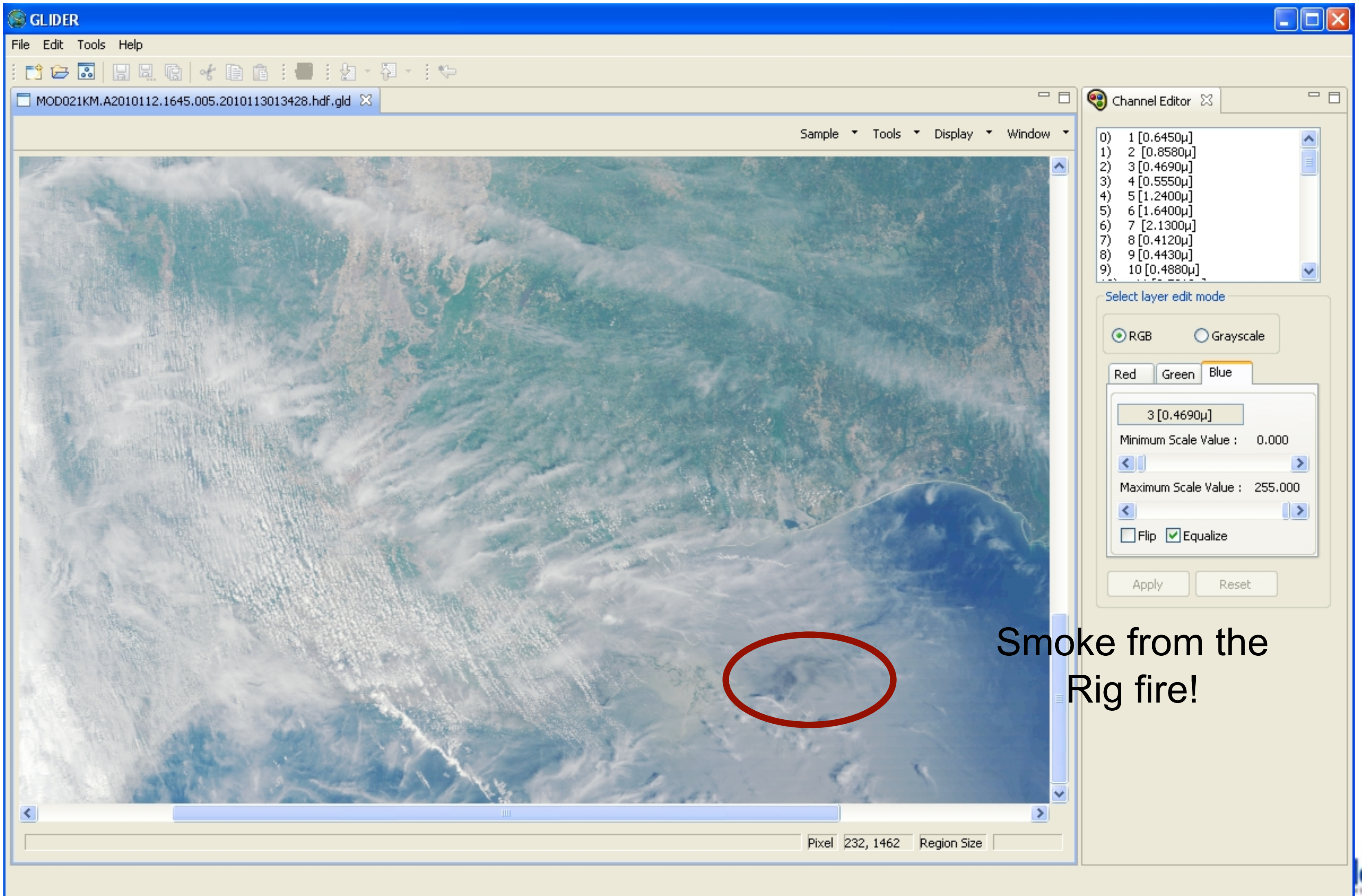
Displaying RGB True Color

Apply Histogram Equalization

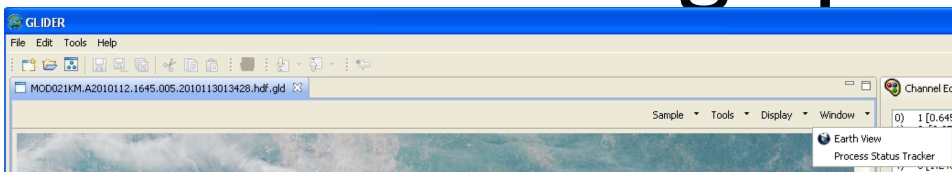


Check box for Red, Green and Blue

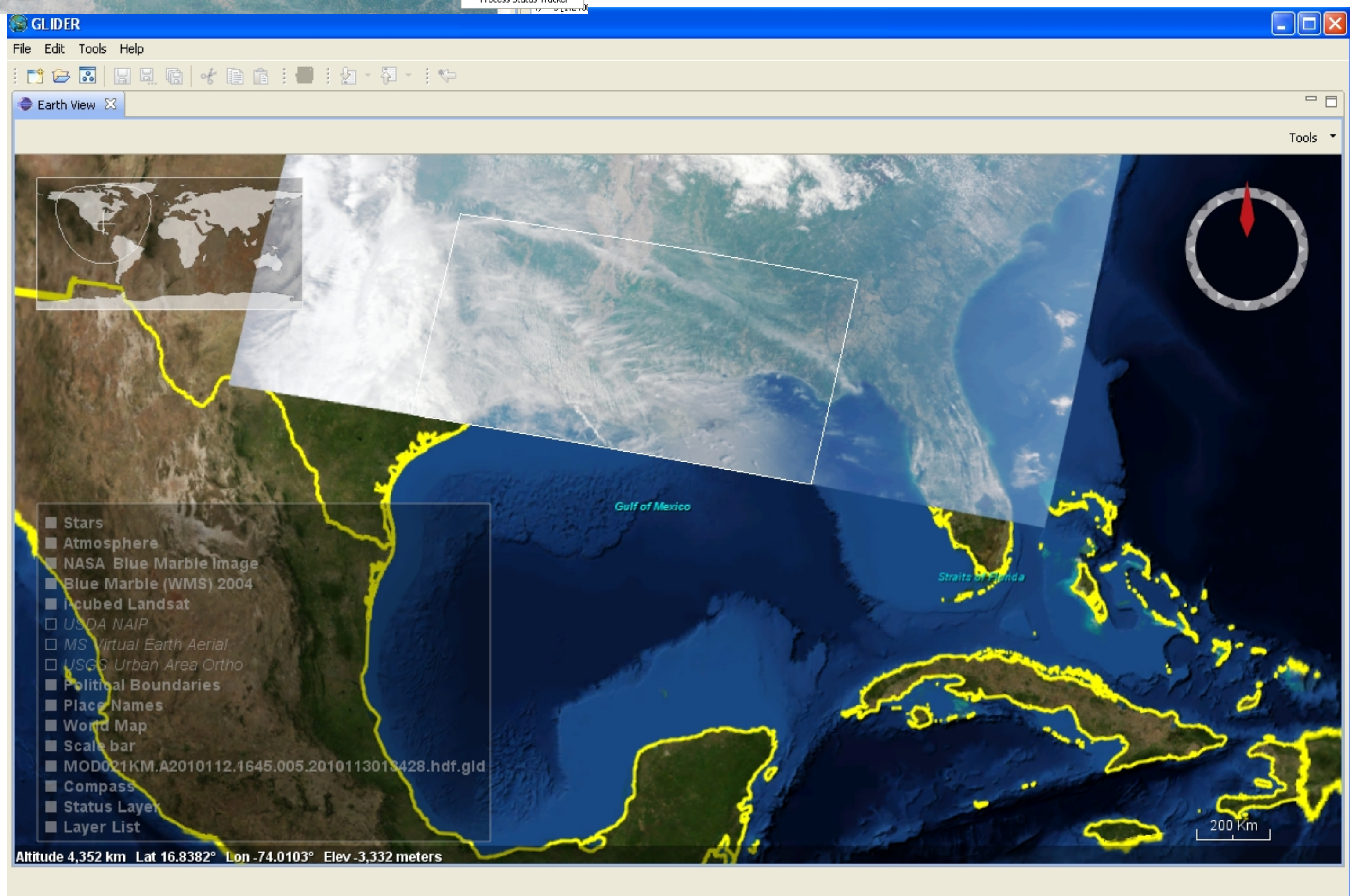
Result


GLIDER
File Edit Tools Help
MOD021KM.A2010112.1645.005.2010113013428.hdf.gld
Sample Tools Display Window
Channel Editor
0) 1 [0.6450μ]
1) 2 [0.8580μ]
2) 3 [0.4690μ]
3) 4 [0.5550μ]
4) 5 [1.2400μ]
5) 6 [1.6400μ]
6) 7 [2.1300μ]
7) 8 [0.4120μ]
8) 9 [0.4430μ]
9) 10 [0.4880μ]
Select layer edit mode
RGB
Grayscale
Red
Green
Blue
3 [0.4690μ]
Minimum Scale Value : 0.000
Maximum Scale Value : 255.000
Flip
Equalize
Apply
Reset
Pixel
232, 1462
Region Size
Smoke from the
Rig fire!

Proving Spatial Context

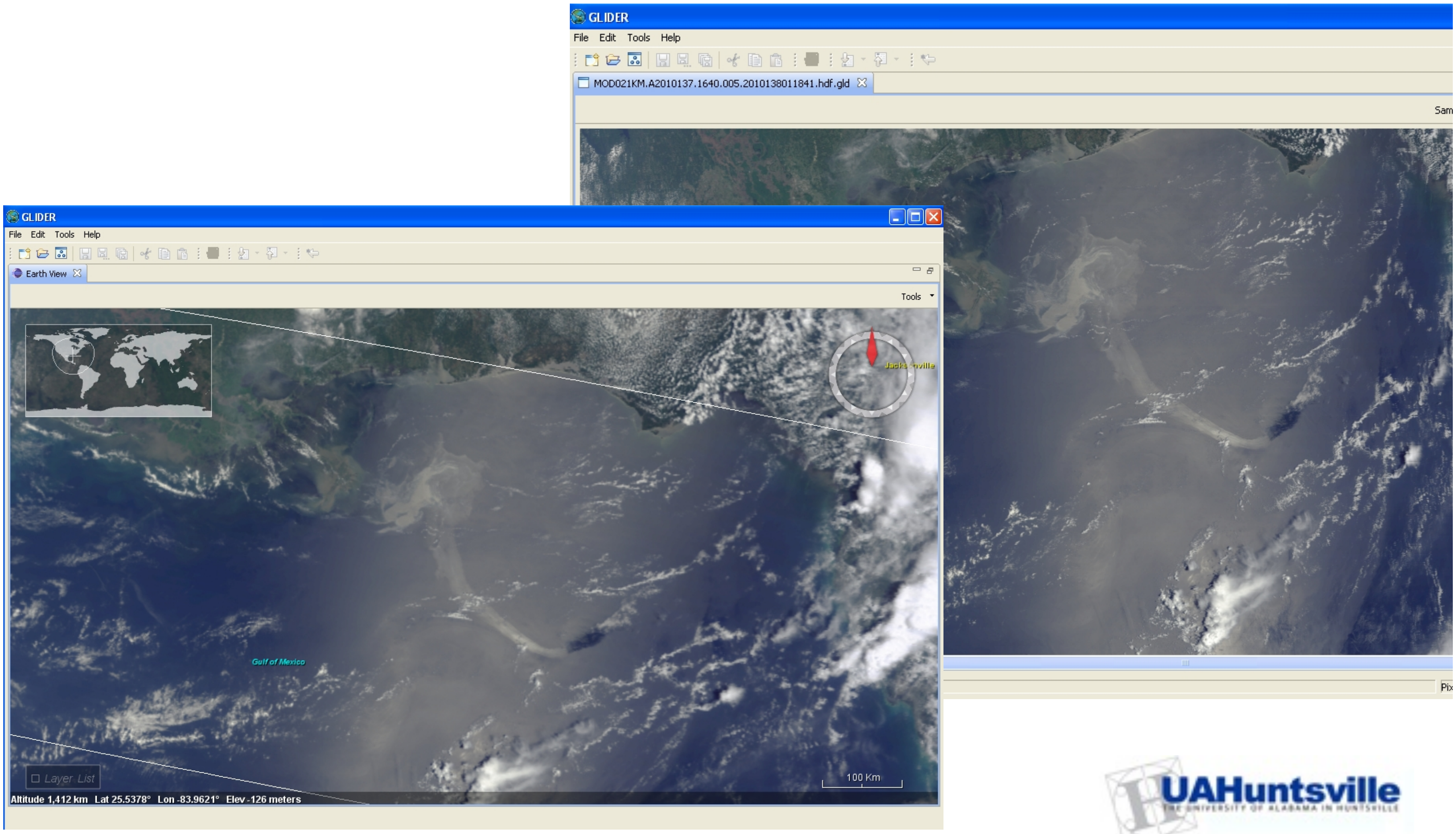


Select Earth View



Oil Spill Image

- Open MOD021KM.A2010137.1640.005.2010138011841.hdf.gld, apply histogram equalization
- Sun glint in this image makes the oil easily visible



DIY

- Open and visualize the other two gld files
- Change the order of the layers on Earth View

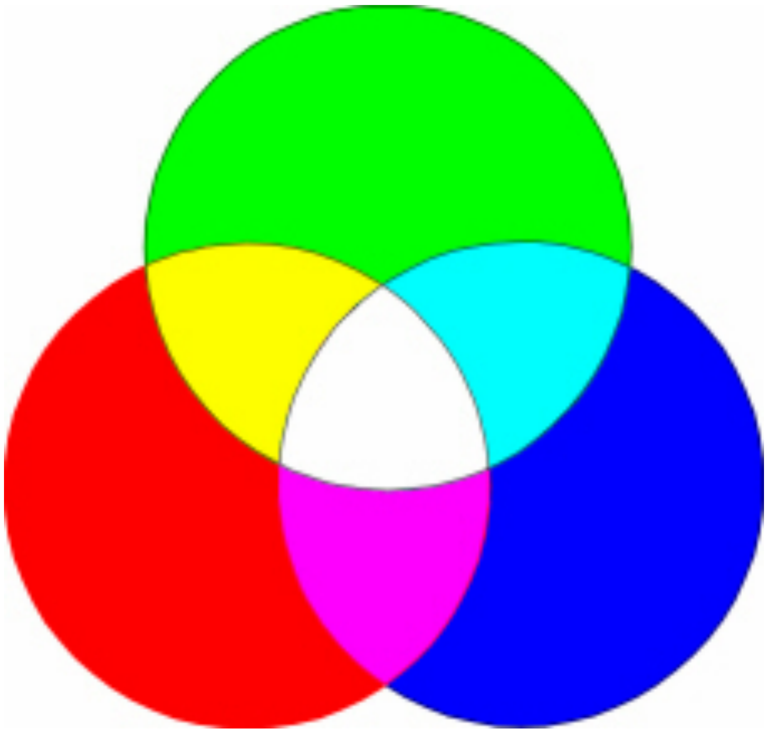
Learning Module 2

"Smoke on the Water"

–Learn how to create False color composites to visually separate features in MODIS data

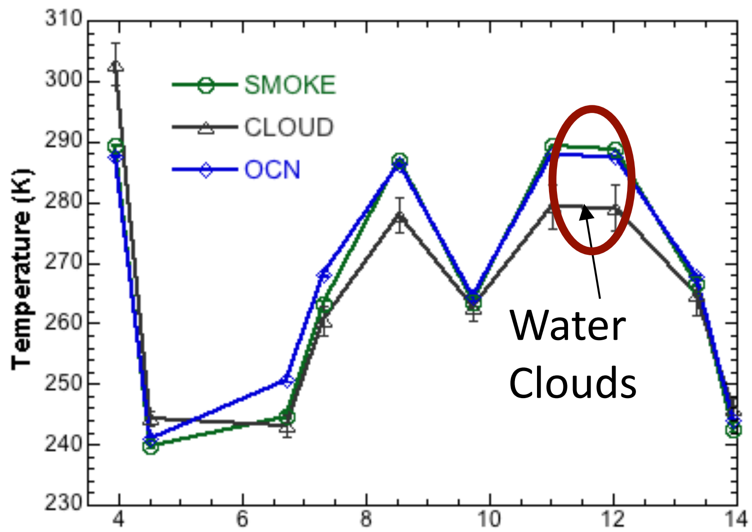
[Courtesy of Dr. Sundar Christopher, UAHuntsville]

Color Composite

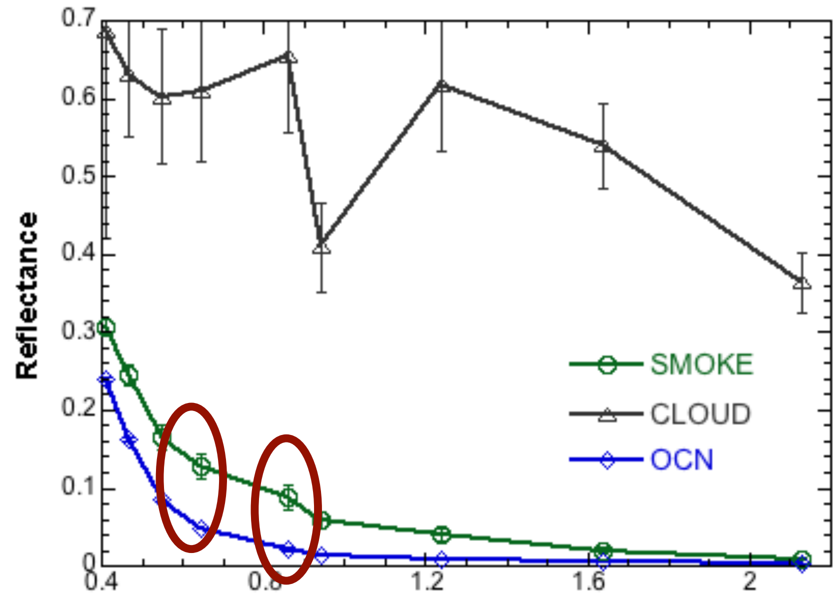


Yellow is a mix of red and green; orange is a mix of more red and some green; white is an equal mix of all three primaries, and black is simply the absence of any colored light of any wavelength

Spectral Signatures of Aerosols and Clouds



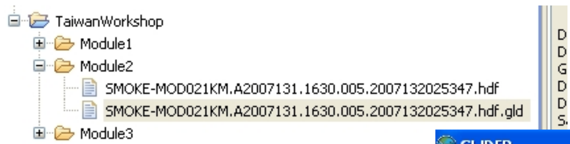
MODIS Wavelength (um)



MODIS Wavelength (um)

Wavelengths of interest: 0.645um, 0.858 um,11.03 um

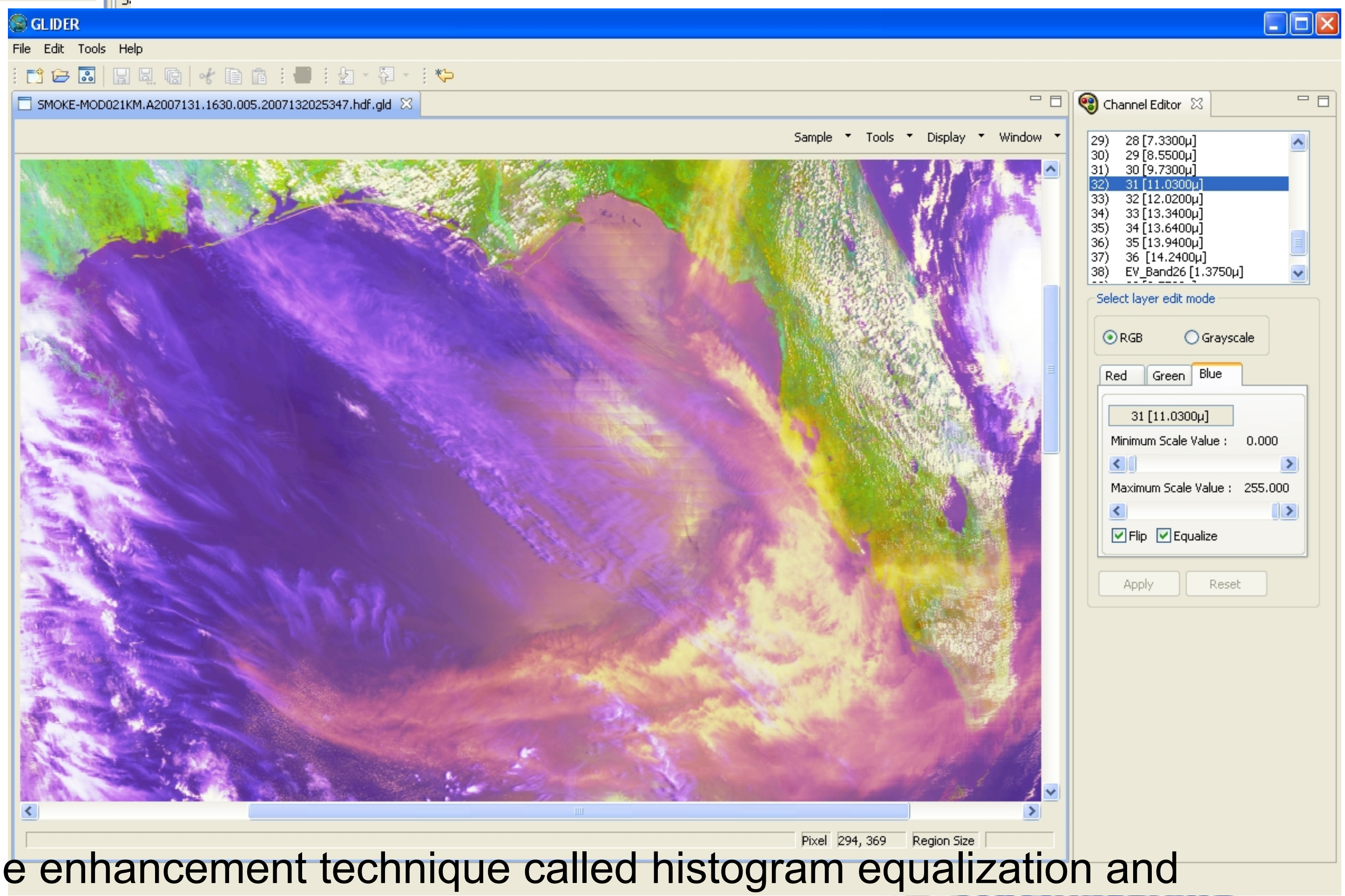
Image View



Set Red : .645um (B1), Mark,Equalize

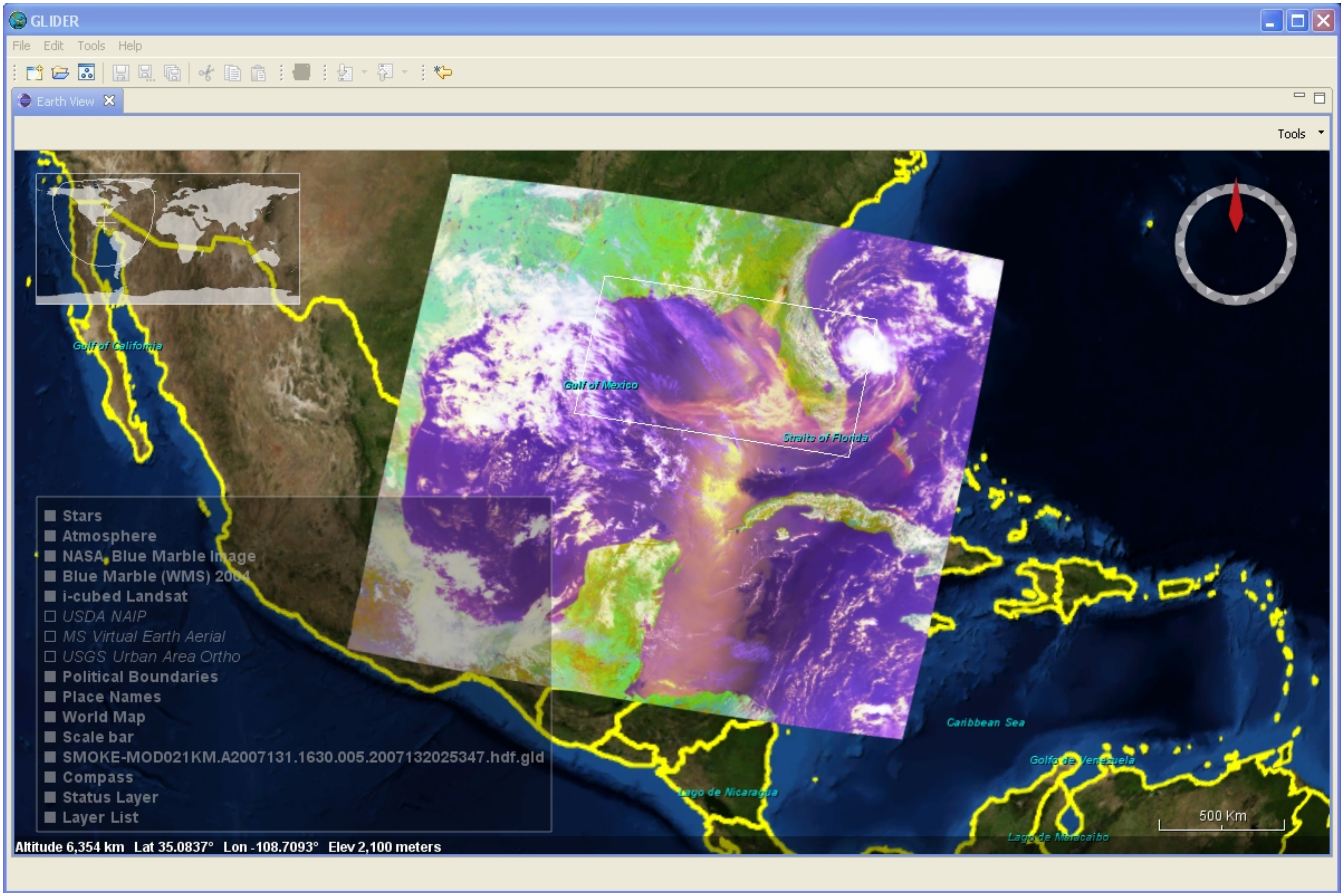
Set Green: 0.858 um (B2), Equalize

Set Blue : 11.03 um (B31), mark Equalize and Flip



Equalize is an image enhancement technique called histogram equalization and flip inverts the infrared channel to make clouds look brighter than the surface

Earth View


GLIDER
File Edit Tools Help
Earth View
Tools
Gulf of California
Gulf of Mexico
Straits of Florida
Caribbean Sea
Golfo de Venezuela
Lago de Nicaragua
Lago de Maracaibo
500 Km
Stars
Atmosphere
NASA Blue Marble Image
Blue Marble (WMS) 2004
i-cubed Landsat
USDA NAIP
MS Virtual Earth Aerial
USGS Urban Area Ortho
Political Boundaries
Place Names
World Map
Scale bar
SMOKE-MOD021KM.A2007131.1630.005.2007132025347.hdf.gld
Compass
Status Layer
Layer List
Altitude 6,354 km Lat 35.0837° Lon -108.7093° Elev 2,100 meters

Learning Module 3

"Dust in the Wind"

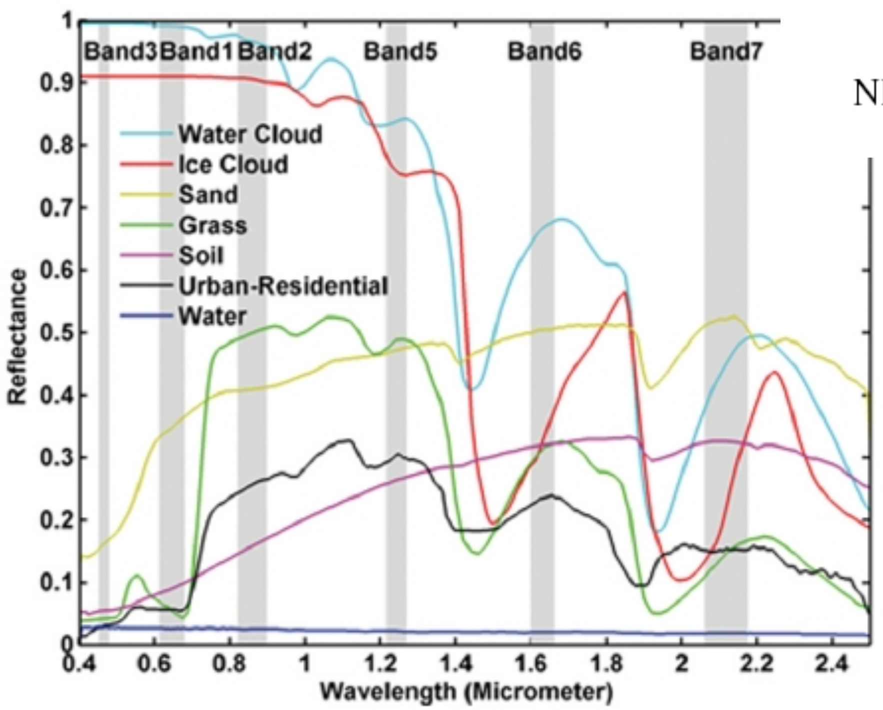
– Learn how to use band math feature in GLIDER to create indices
– Use indices to detect airborne dust
– Browse a journal article and then apply the results from the paper to detect dust over china

NDDI - Qu et al, 2006

484 IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, VOL. 3, NO. 4, OCTOBER 2006

Asian Dust Storm Monitoring Combining Terra and Aqua MODIS SRB Measurements

John J. Qu, *Member, IEEE*, Xianjun Hao, *Member, IEEE*, Menas Kafatos, *Member, IEEE*, and Lingli Wang



$$\mathrm{NDDI} = (\rho_{2.13\,\mu\mathrm{m}} - \rho_{0.469\,\mu\mathrm{m}})/(\rho_{2.13\,\mu\mathrm{m}} + \rho_{0.469\,\mu\mathrm{m}}) \quad (1)$$

The spectral characteristic of sand suggests that strong SDS signals can be obtained using the difference between the 2.13-μm band signal, which is high, and the 0.469-μm band, where the signal is relatively much lower. This difference distinguishes rather well between SDS and water or ice clouds.

SRB – Solar Reflectance Band
NDDI – Normalized Difference Dust Index
SDS – Sand and Dust Storms

NDDI – Qu et al, 2006

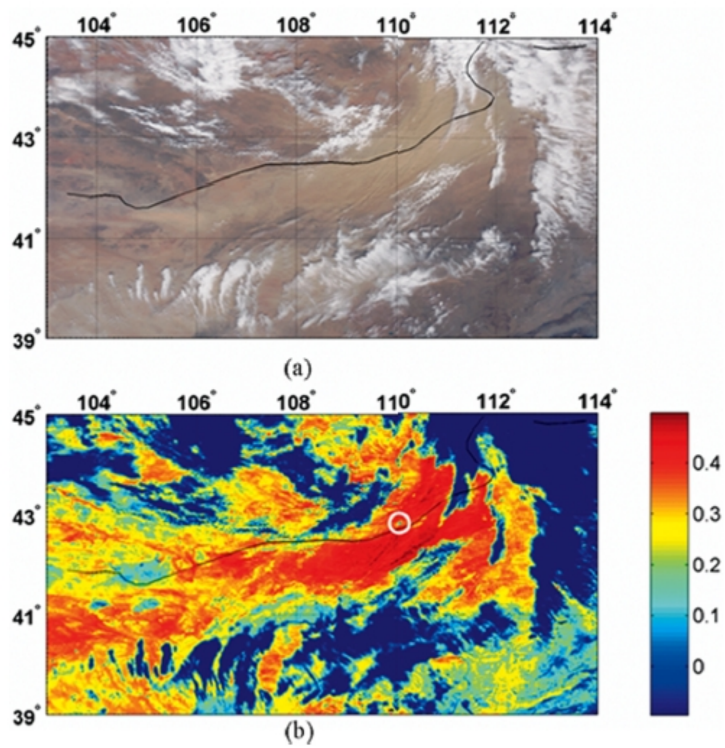
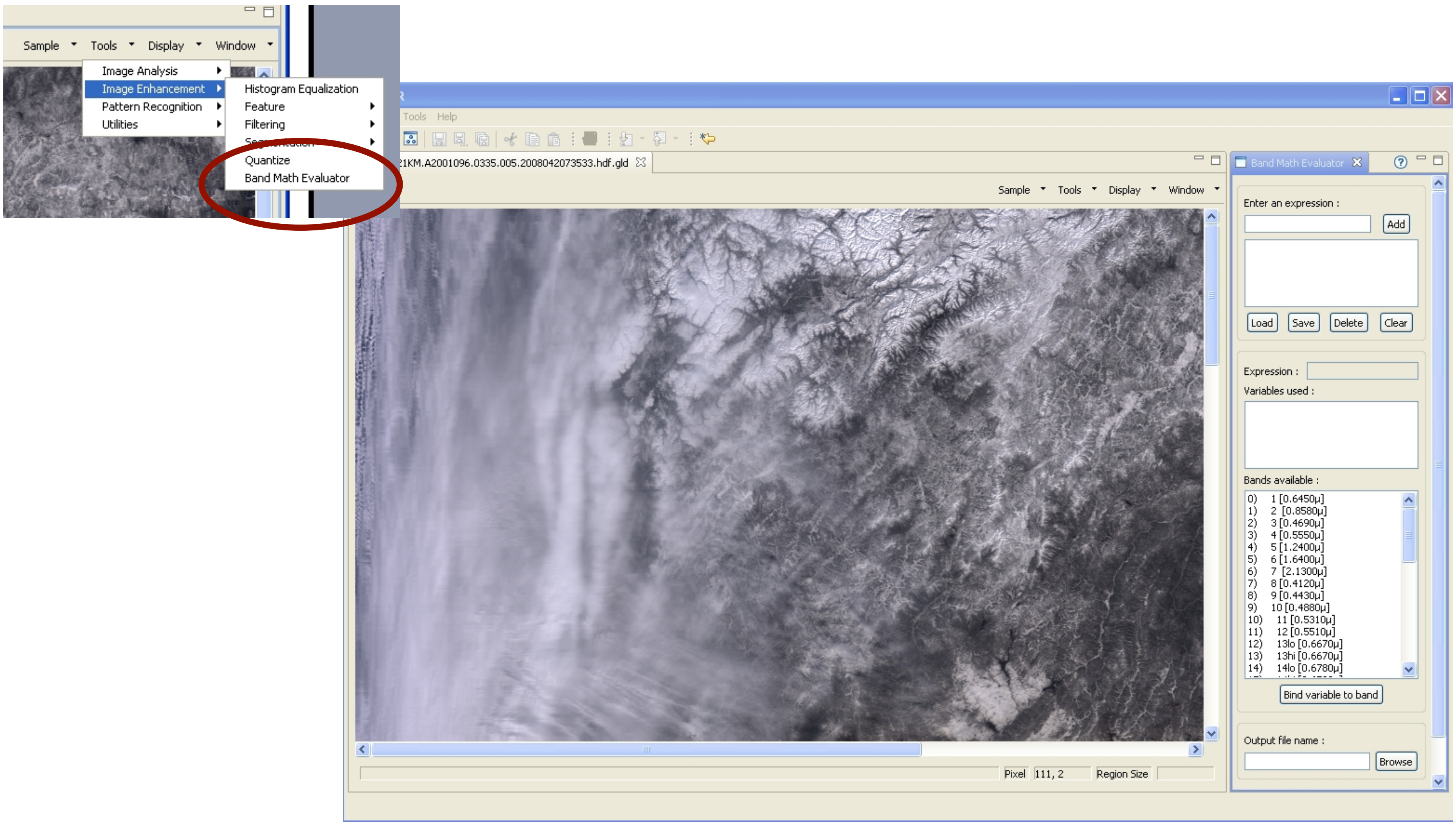


Fig. 3. (a) Terra MODIS true-color image (3:40 UTC, March 27, 2004) shows Asian dust storm over Northern China and Southern Mongolian regions. (b) Terra MODIS NDDI image shows the clouds and dust storms. The cloud and dust storm can be easily identified (for cloud NDDI < 0.0 and for dust storm NDDI > 0.28).

Thresholds for detection:
CLOUDS: NDDI < 0.0
SURFACE FEATURES: NDDI < 0.28
DUST: NDDI > 0.28

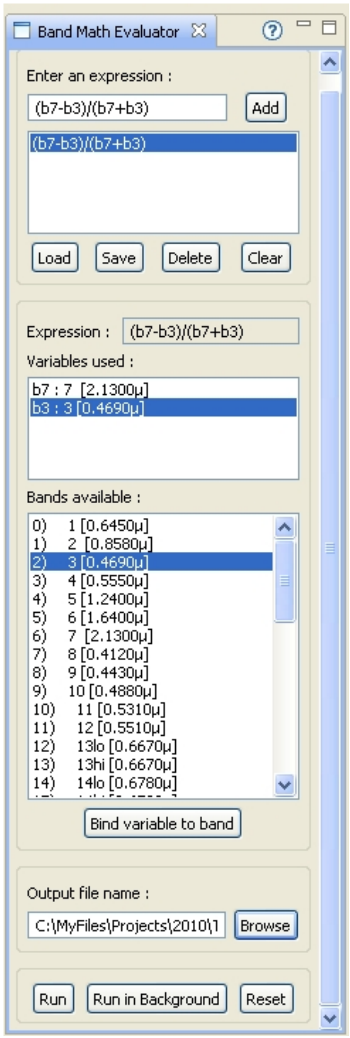
Lets Try it Out!

Open MOD021KM.A2001096.0335.005.2008042073533.hdf.gld in Image View

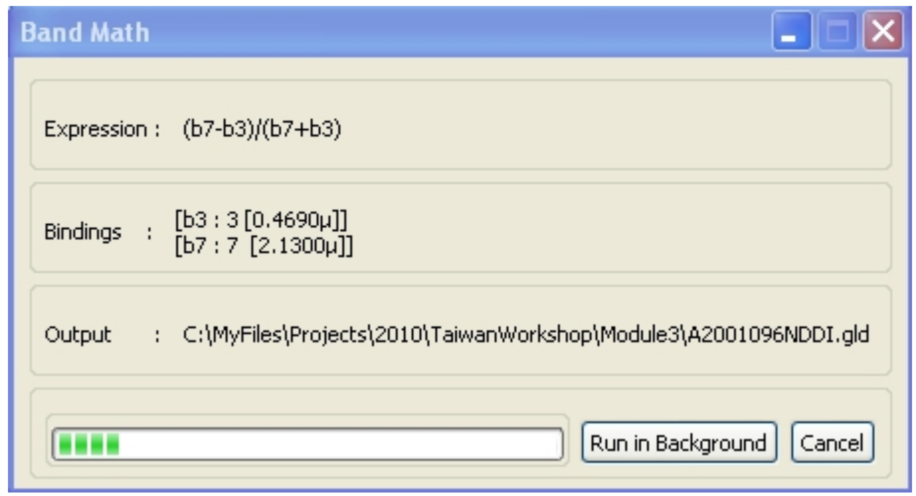


Enter the NDDI formula

Enter a Mathematical expression

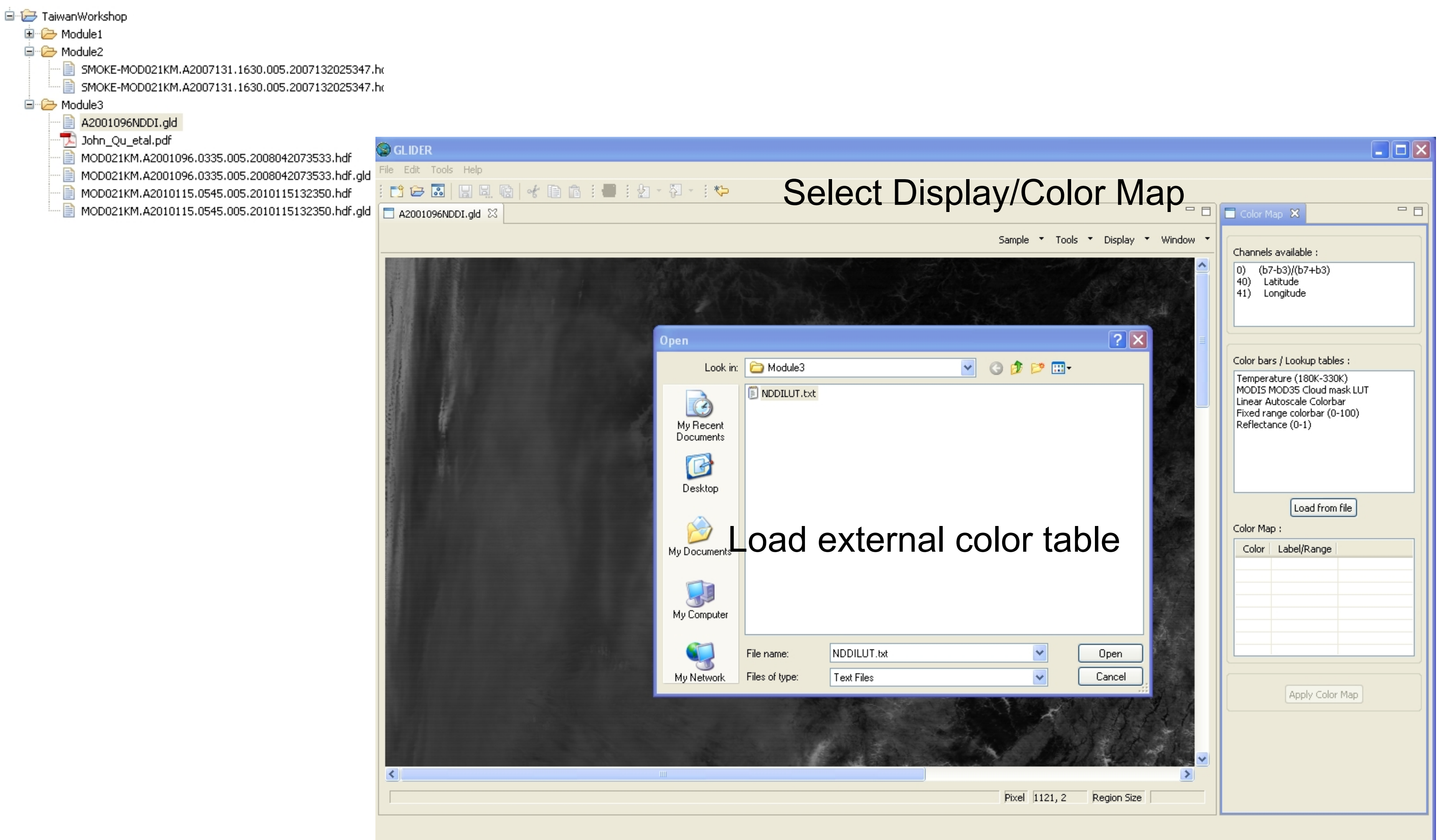


Bind bands to the variables

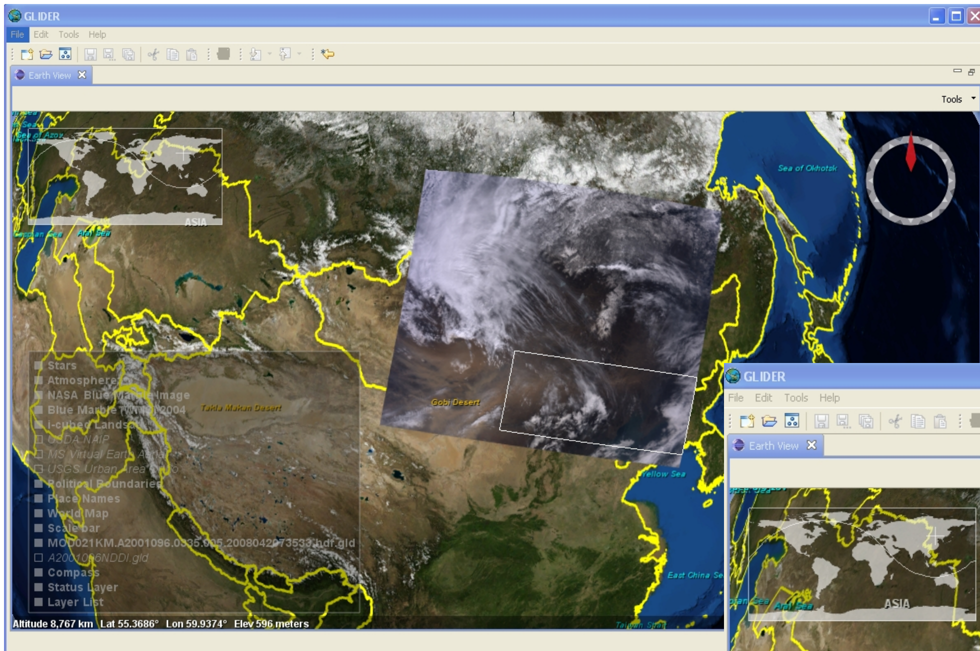


UAHuntsville

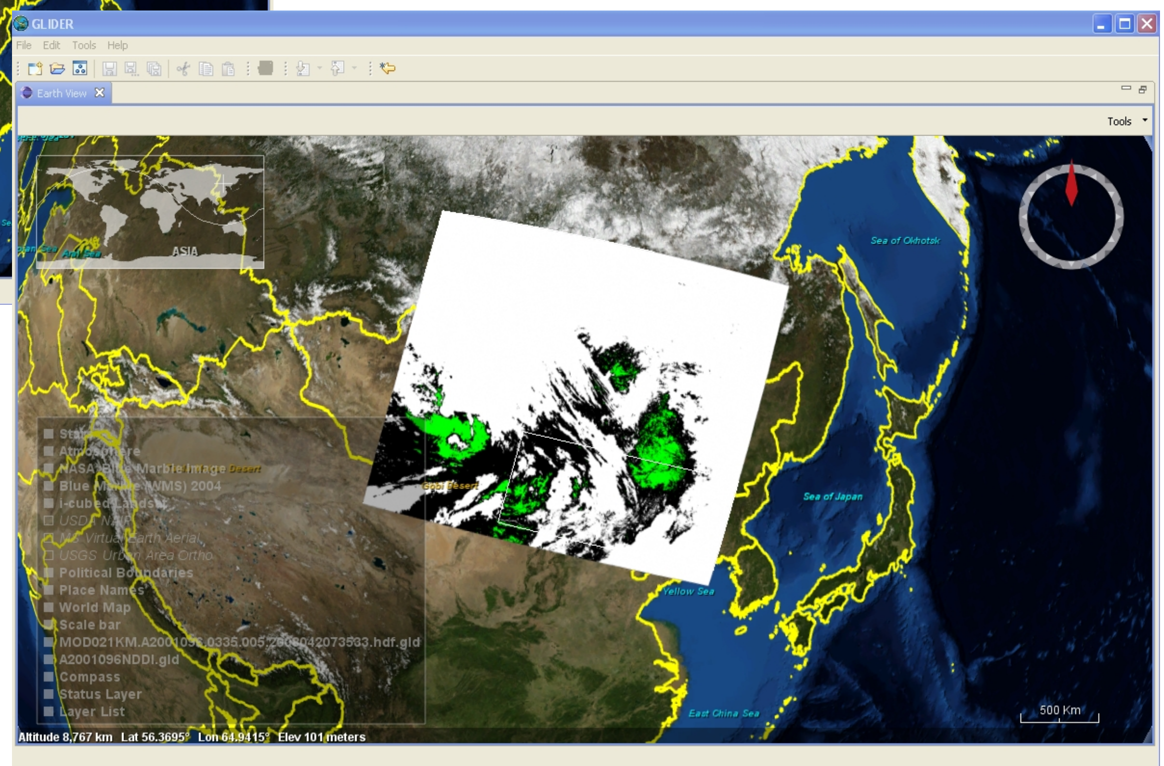
Visualize Result in Image View and Apply a Custom Color Map



Display Original and NDDI image on Earth View



Toggle Layer On/Off
To visually inspect
how well NDDI works in
Detecting dust



DIY

- Apply NDDI using Band Math feature in GLIDER to the other MODIS granule

Learning Module 4

"Ashes to Ashes" – Part 1

–Look at Ash/Steam Plume event from Iceland's Eyjafjallajoekull Volcano

–Learn how to subset imagery both spatially and spectrally

–Apply clustering algorithm to generate classification maps

What is Cluster Analysis?

- Cluster: a collection of data objects
  - Similar to one another within the same cluster
  - Dissimilar to the objects in other clusters
- Cluster analysis
  - Grouping a set of data objects into clusters
- Clustering is unsupervised classification: no predefined classes

Source – Dr. John Rushing, ITSC/UAHuntsville

Similarity and Dissimilarity Between Objects

- Distances are normally used to measure the similarity or dissimilarity between two data objects
- Some popular ones include: *Minkowski distance*:

$$d(i,j) = \sqrt[q]{(|x_{i_1} - x_{j_1}|^q + |x_{i_2} - x_{j_2}|^q + ... + |x_{i_p} - x_{j_p}|^q)}$$

where $i = (x_{i1}, x_{i2}, ..., x_{ip})$ and $j = (x_{j1}, x_{j2}, ..., x_{jp})$ are two p-dimensional data objects, and q is a positive integer

- If $q = 1$, d is Manhattan distance

$$d(i,j) = |x_{i_1} - x_{j_1}| + |x_{i_2} - x_{j_2}| + ... + |x_{i_p} - x_{j_p}|$$

Source – Dr. John Rushing, ITSC/UAHuntsville

Similarity and Dissimilarity Between Objects (Cont.)

- *If q = 2, d* is Euclidean distance:

$$d(i,j) = \sqrt{(|x_{i_1} - x_{j_1}|^2 + |x_{i_2} - x_{j_2}|^2 + ... + |x_{i_p} - x_{j_p}|^2)}$$

  – Properties
    - $d(i,j) \geq 0$
    - $d(i,i) = 0$
    - $d(i,j) = d(j,i)$
    - $d(i,j) \leq d(i,k) + d(k,j)$

What should one look out for when using distance measures?

- Also, one can use weighted distance, parametric Pearson product moment correlation, or other disimilarity measures

Source – Dr. John Rushing, ITSC/UAHuntsville

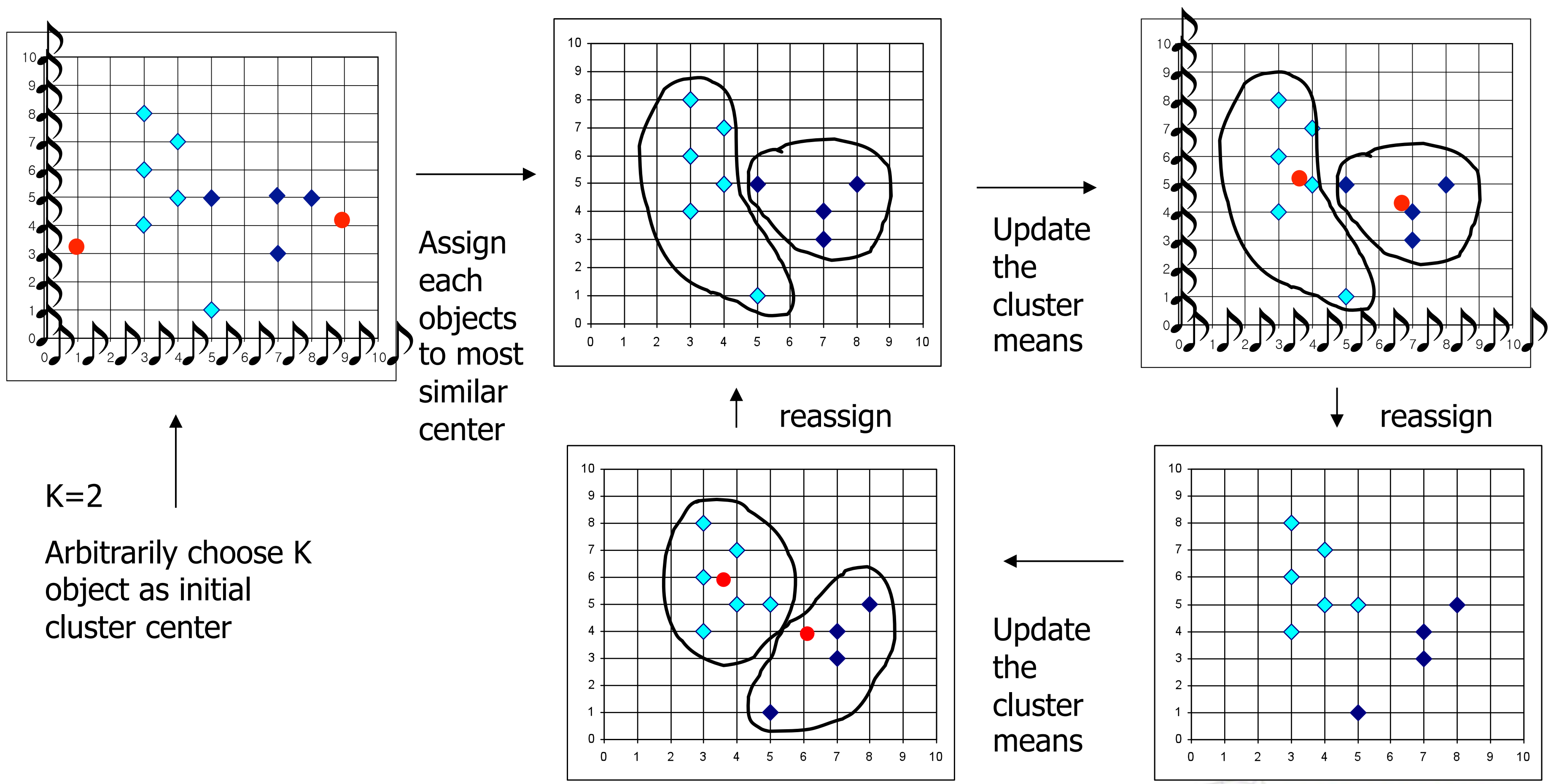
The *K-Means* Clustering Method

- Given *k*, the *k-means* algorithm is implemented in four steps:
  - Partition objects into *k* nonempty subsets
  - Compute seed points as the centroids of the clusters of the current partition (the centroid is the center, i.e., *mean point*, of the cluster)
  - Assign each object to the cluster with the nearest seed point
  - Go back to Step 2, stop when no more new assignment

Source – Dr. John Rushing, ITSC/UAHuntsville

The *K-Means* Clustering Method

- Example



Source – Dr. John Rushing, ITSC/UAHuntsville

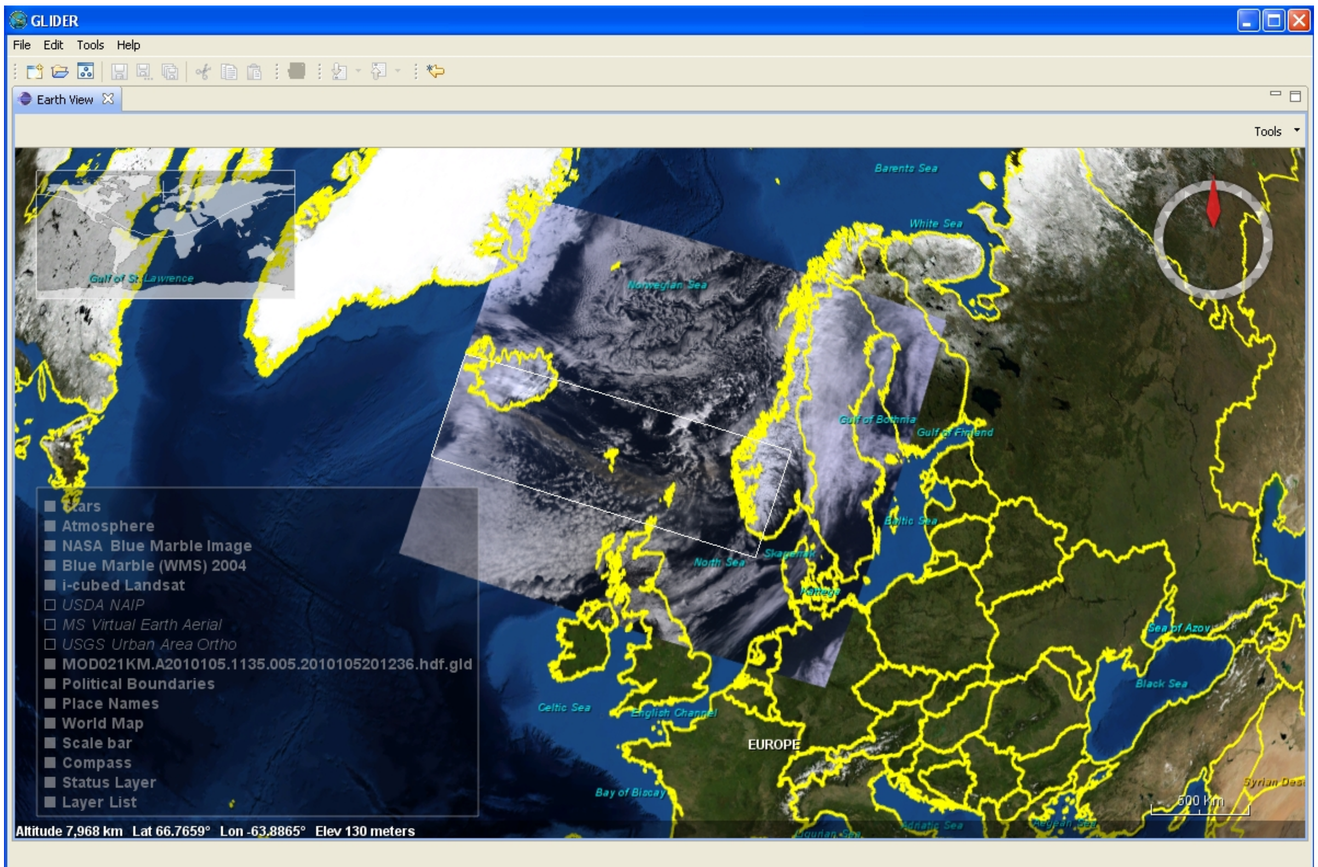
Let’s Apply a Clustering Algorithm

- Goal – Create a thematic/classification map using MODIS L1B data with three classes: Clouds, Ash/Steam and Ocean
- Methodology:
  - Subset the data both spatially and spectrally
  - Apply K-Means with k=5 and let the algorithm find groups in spectral feature space
  - Assign semantic (3) classes to the 5 groups

Lets Apply a Clustering Algorithm

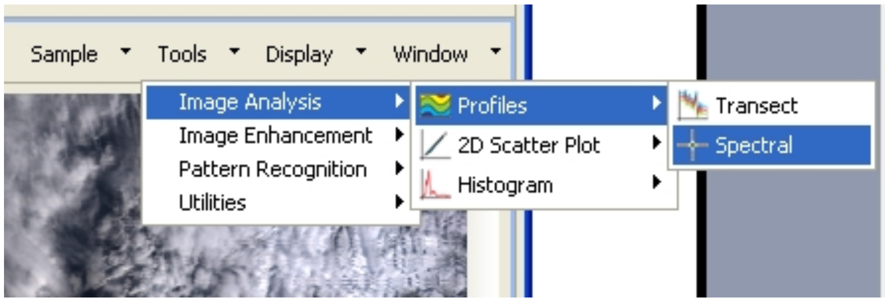
•Open MOD021KM.A2010105.1135.005.2010105201236.hdf.gld in Image View and Earth View

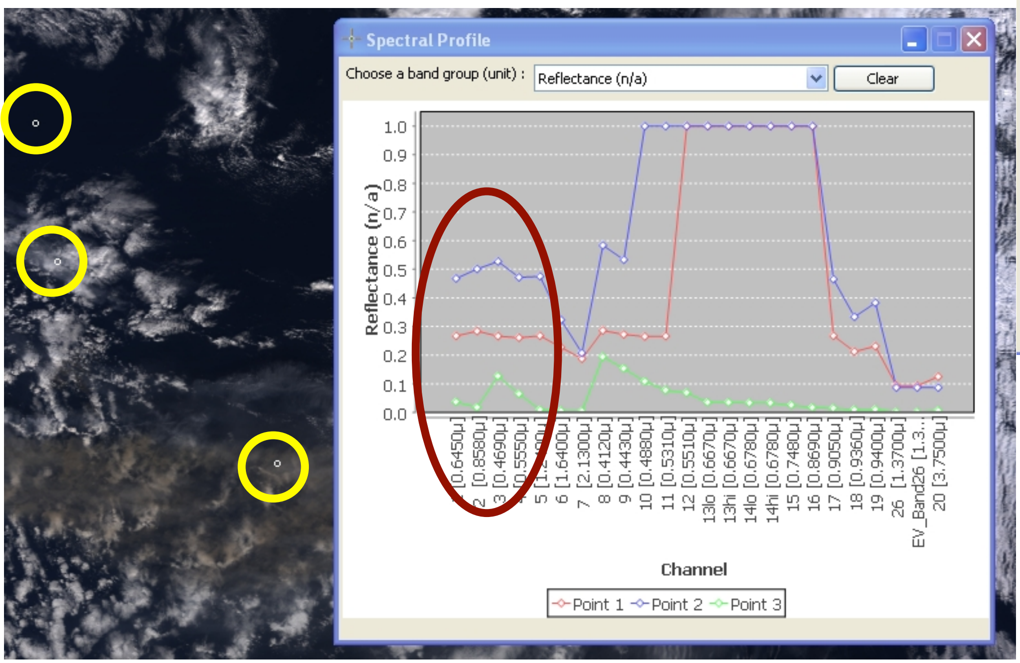
• Locate the Ash/Steam in the image



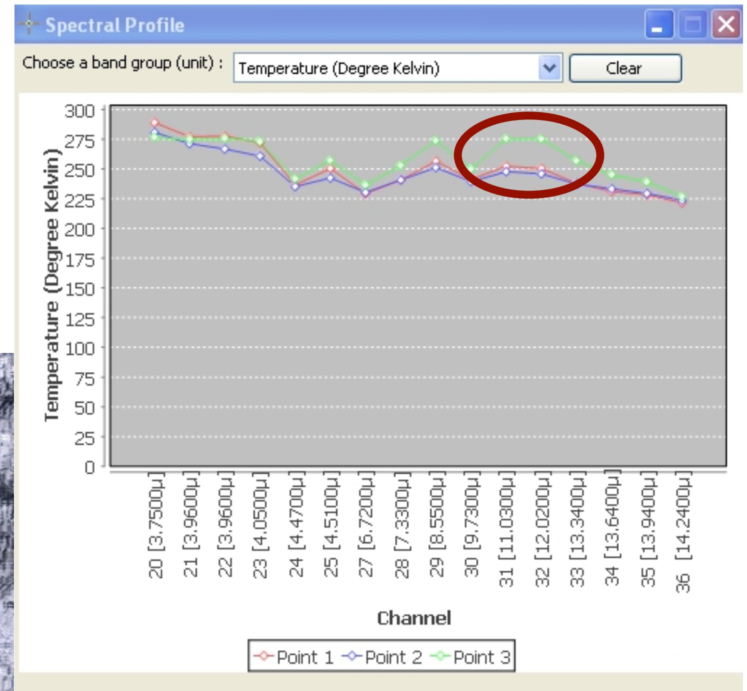
UAHuntsville
THE UNIVERSITY OF ALABAMA IN HUNTSVILLE

Look at the Spectral Signatures for Clouds, Ash/Steam and Ocean



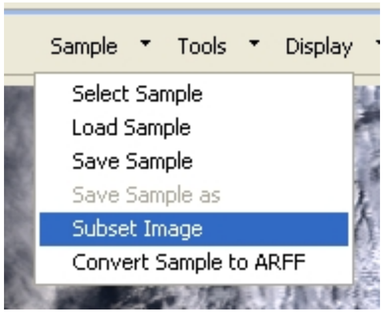


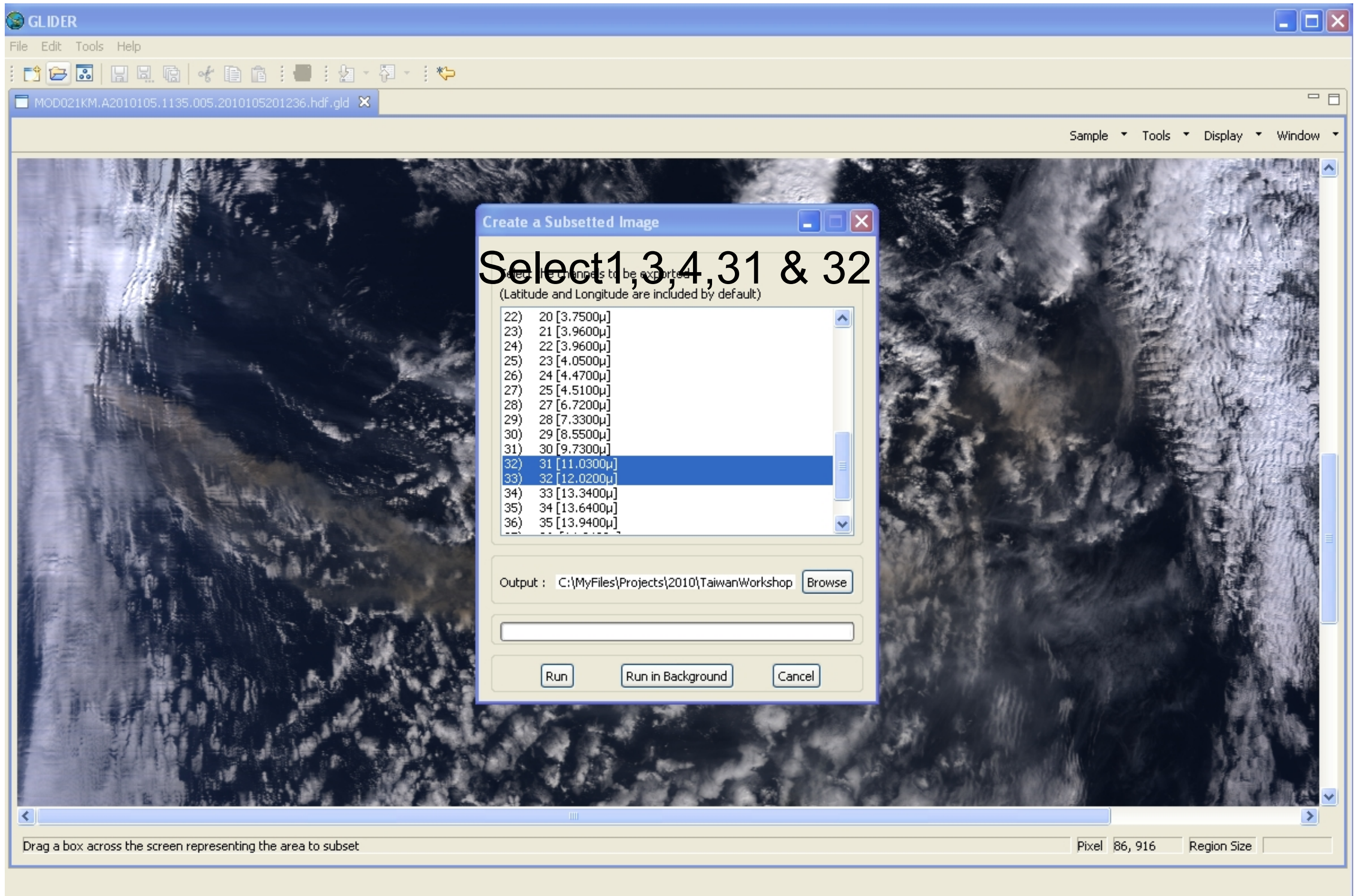
Select representative pixels for Ash, Cloud, Ocean



Lets keep bands 1,3,4,31 & 32

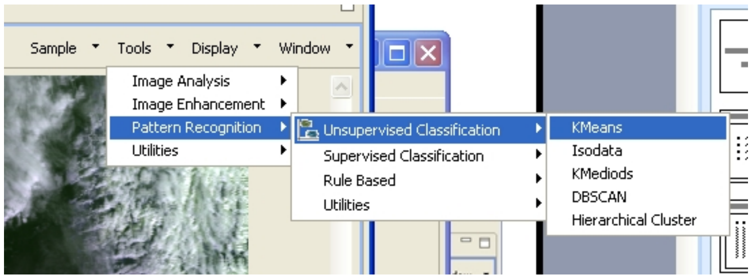
Spatially and Spectrally Subset Data



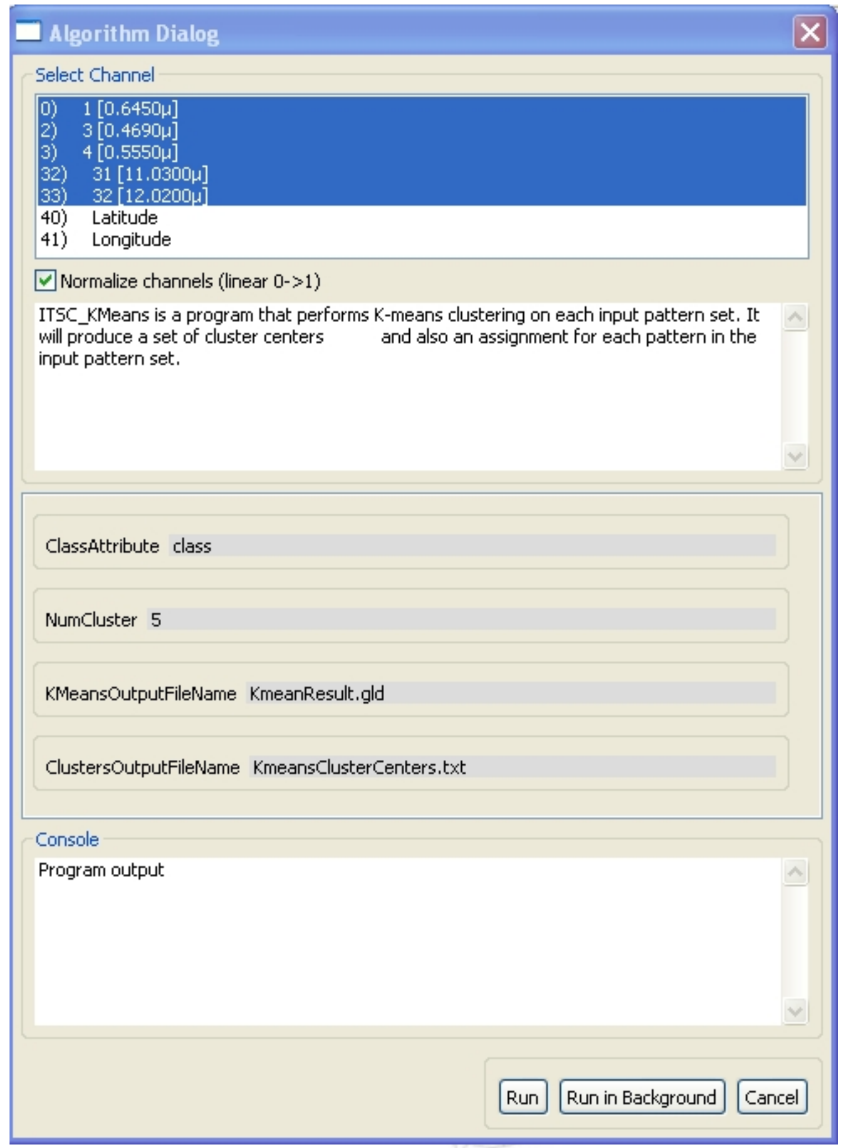


Select an area within the Image View, then select Bands, provide output filename (subset.gld) and hit Run button. Go to Project View and load subset.gld in Image View

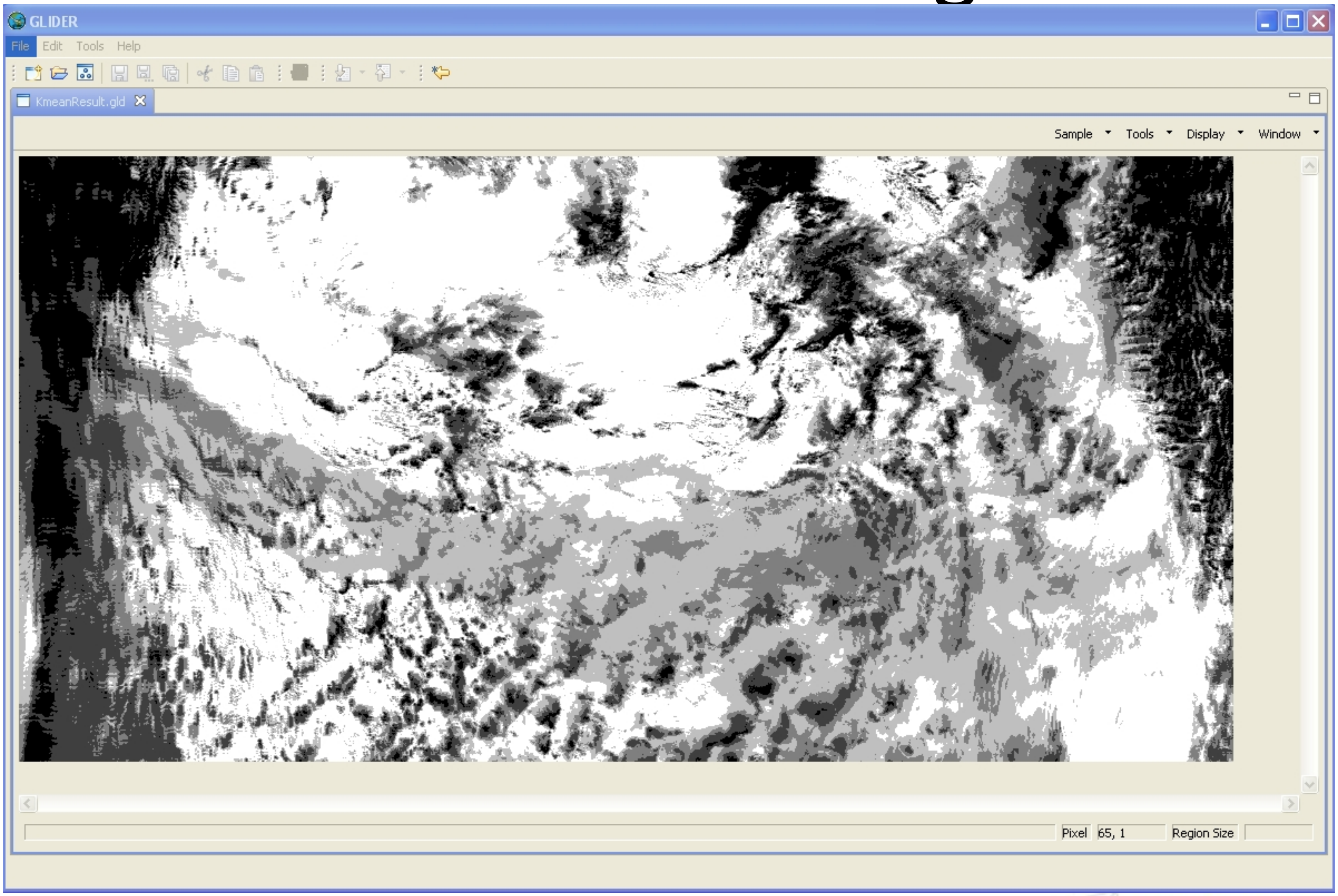
Apply KMeans Algorithm



- Only select the spectral bands
- Make sure you select normalize channels
- Set the # of clusters to 5 even though we only want three final classes
- We will merge clusters at the end!

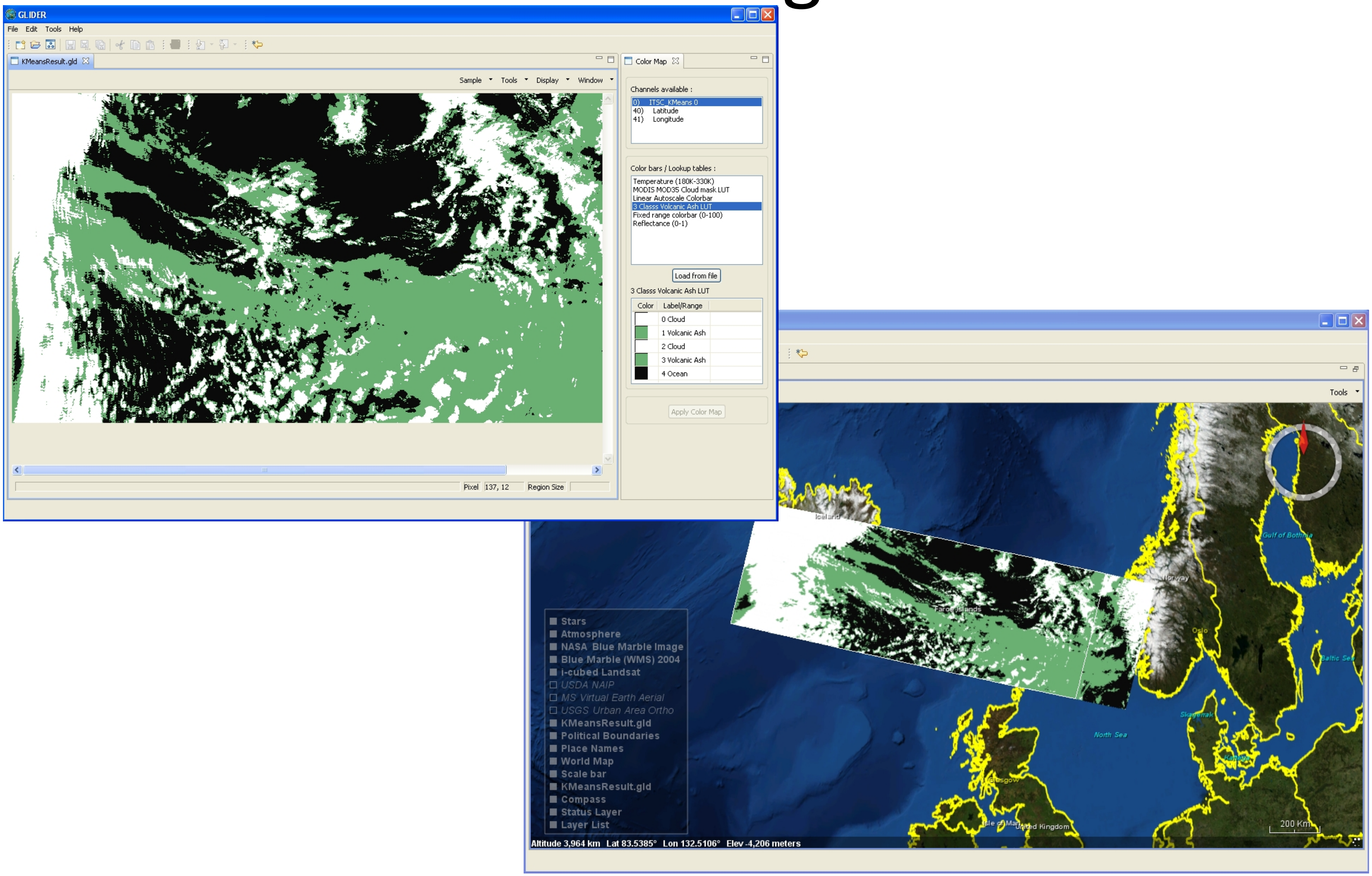


Visualize Result in Image View



Lets merge classes to create a map with only three classes –
Load the ClassLUT.txt Color Map

Final Clustering Result



Learning Module 5

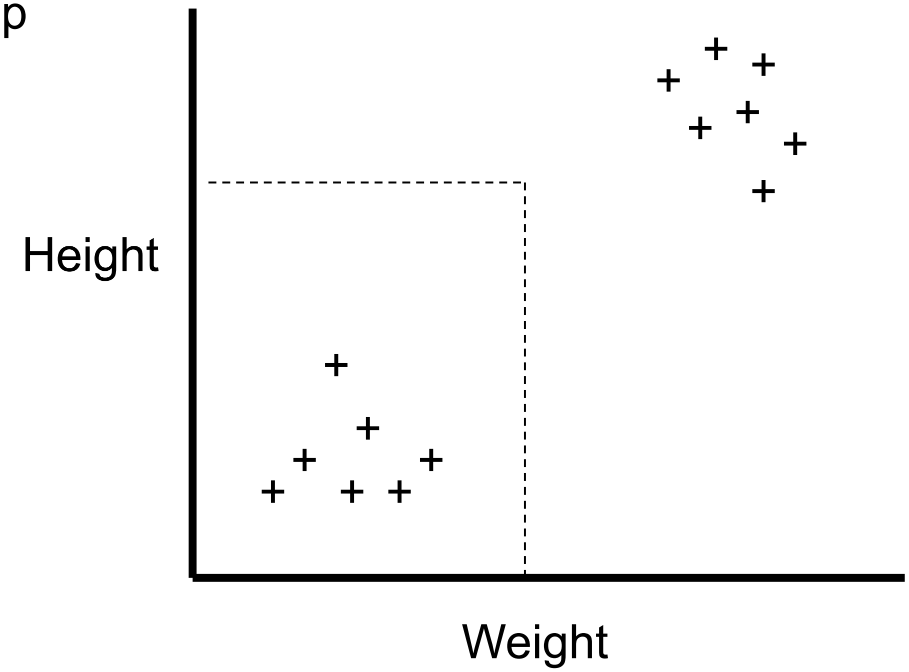
"Ashes to Ashes" – Part 2

– Learn how to construct a supervised classification process
– Learn how take training samples
– Create a *mining recipe/workflow* using visual programming

Simple Classification Example

Given a dataset containing student's names, weight and height, develop rules to classify between Football players and non-Football Players

Name	Height	Weight
Joe Montana	6'4"	230
………	…….	……
Avg Joe	5'9"	180



Classification Rule: If Ht > 6'3" AND Wt > 220 lb THEN Football Player Else Regular Student

Classification Problem

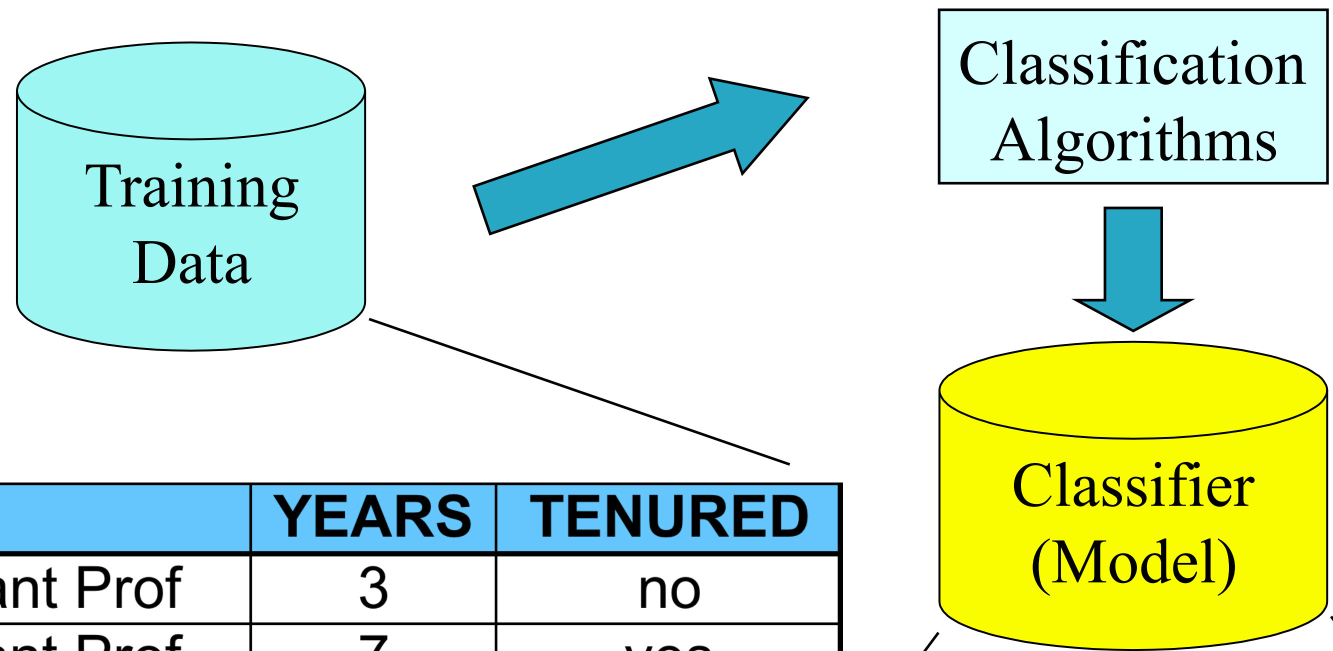
- Satellite Remote Sensing: Features can be spectral bands and other derived parameters (textures, ratios etc)
- Real Life Problems: Features are MANY!
- One can limit the problem using Heuristics (e.g., NDDI)
- Human's cannot visualize beyond 3 dimensions
- Hence, need for Pattern Recognition/Data Mining algorithms

Classification—A Two-Step Process

- Model construction: describing a set of predetermined classes
  - Each tuple/sample is assumed to belong to a predefined class, as determined by the class label attribute
  - The set of tuples used for model construction is training set
  - The model is represented as classification rules, decision trees, or mathematical formulae
- Model usage: for classifying future or unknown objects
  - Estimate accuracy of the model
    - The known label of test sample is compared with the classified result from the model
    - Accuracy rate is the percentage of test set samples that are correctly classified by the model
    - Test set is independent of training set, otherwise over-fitting will occur
  - If the accuracy is acceptable, use the model to classify data tuples whose class labels are not known

Source – Dr. John Rushing, ITSC/UAHuntsville

Classification Process (1): Model Construction

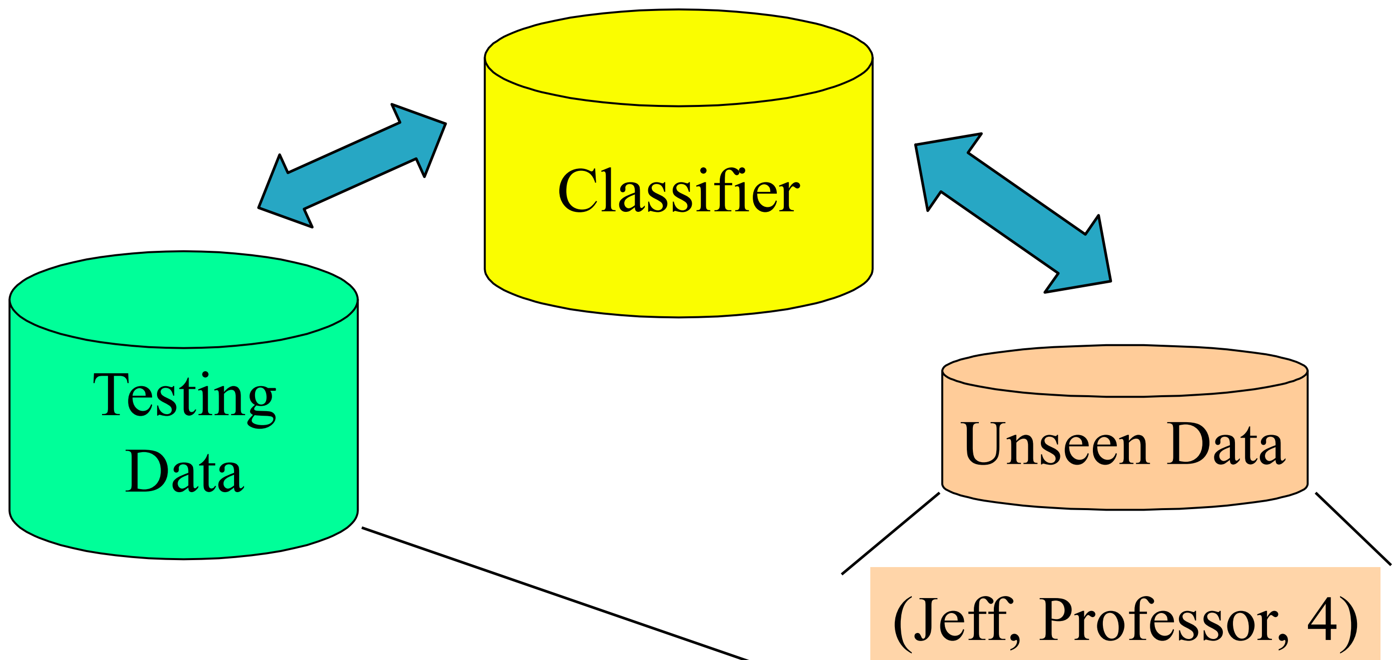


NAME	RANK	YEARS	TENURED
Mike	Assistant Prof	3	no
Mary	Assistant Prof	7	yes
Bill	Professor	2	yes
Jim	Associate Prof	7	yes
Dave	Assistant Prof	6	no
Anne	Associate Prof	3	no

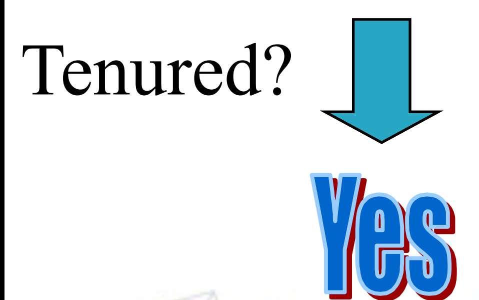
IF rank = 'professor'
OR years > 6
THEN tenured = 'yes'

Source – Dr. John Rushing, ITSC/UAHuntsville

Classification Process (2): Use the Model in Prediction



NAME	RANK	YEARS	TENURED
Tom	Assistant Prof	2	no
Merlisa	Associate Prof	7	no
George	Professor	5	yes
Joseph	Assistant Prof	7	yes



Source – Dr. John Rushing, ITSC/UAHuntsville

Maximum Likelihood/Bayes Classifier

- The *maximum likelihood decision rule* is based on *probability*.
- It assigns each pixel having pattern measurements or features *X* to the class *i* whose units are most probable or likely to have given rise to feature vector *X.*
- In other words, the probability of a pixel belonging to each of a predefined set of *m* classes is calculated, and the pixel is then assigned to the class for which the probability is the highest.
- The maximum likelihood procedure assumes that the training data statistics for each class in each band are *normally distributed* (Gaussian).

- The *maximum likelihood decision rule* is still one of the most widely used supervised classification algorithms.

Source – Introductory Digital Image Processing, Jensen, 2005 Prentice Hall

Maximum Likelihood Classifier

- But how do we obtain the probability information we will need from the remote sensing training data we have collected?
- The answer lies first in the computation of *probability density functions* label samples

Maximum Likelihood Classifier

The estimated *probability density function* for class *wi* (e.g., forest) is computed using the equation:

$$\hat{p}(x \mid w_i) = \frac{1}{(2\pi)^{\frac{1}{2}} \hat{\sigma}_i} \exp\left[-\frac{1}{2} \frac{(x - \hat{\mu}_i)^2}{\hat{\sigma}_i^{\,2}} \right]$$

Source – Introductory Digital Image Processing, Jensen, 2005 Prentice Hall

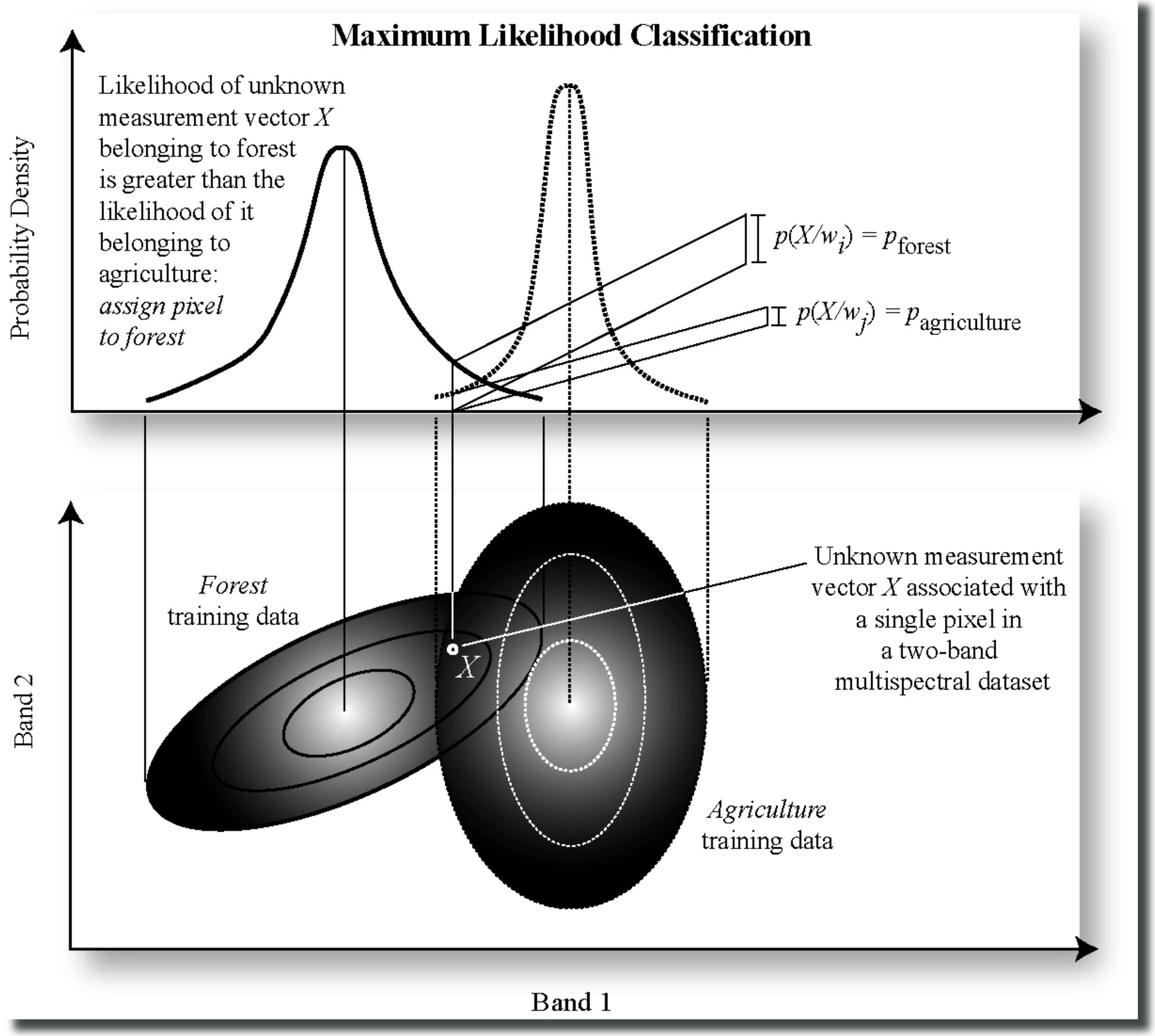
Maximum Likelihood Classifier

But what if the training data consists of multiple bands of remote sensor data for the classes of interest? In this case we compute an *n-dimensional multivariate normal density function* using:

$$p(X \mid w_i) = \frac{1}{(2\pi)^{\frac{n}{2}} \mid V_i \mid^{\frac{1}{2}}} \exp\left[-\frac{1}{2} (X - M_i)^T V_i^{-1} (X - M_i) \right]$$

M: Mean Vector
V: Covariance Matrix

Source – Introductory Digital Image Processing, Jensen, 2005 Prentice Hall



In this case, pixel *X* would be assigned to forest because the probability density of unknown measurement vector *X* is greater for forest than for agriculture.

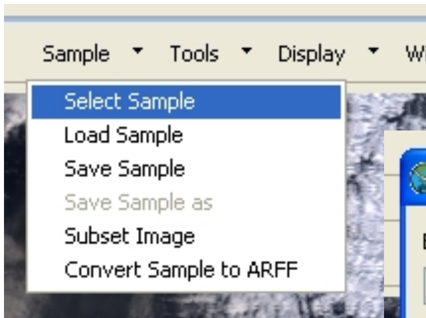
Source – Introductory Digital Image Processing, Jensen, 2005 Prentice Hall

UAHuntsville THE UNIVERSITY OF ALABAMA IN HUNTSVILLE

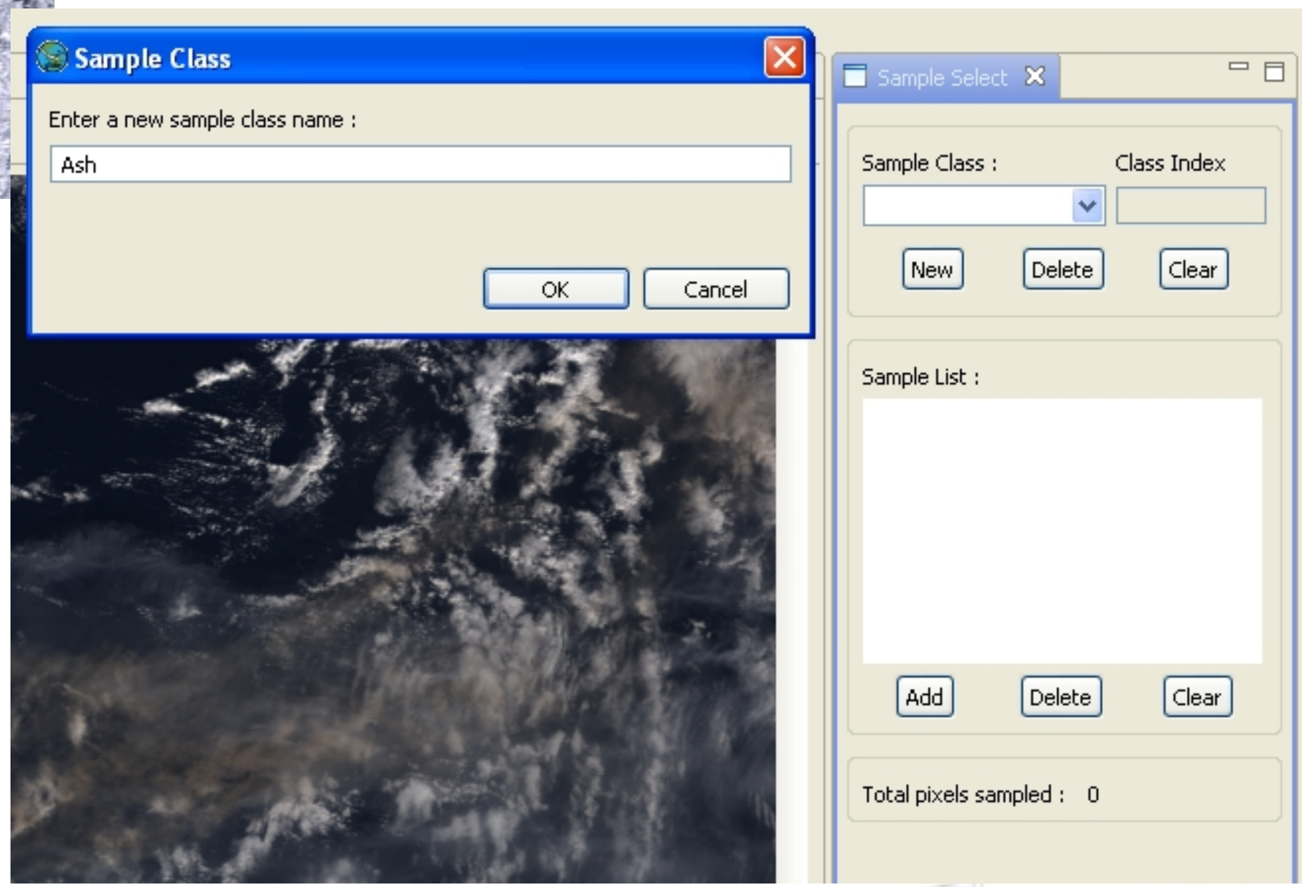
Let's try Supervised Classification

- Goal – Create a thematic/classification map using MODIS L1B data with three classes: Clouds, Ash/ Steam and Ocean
- Methodology:
  – Create a NEW Subsetted data (keep all bands)
  – Take uniform and approximate equal number of samples for the three classes
  – Construct and test the model (train and apply)
  – Use the model for prediction (apply on the original image)

Select Samples for the 3 Classes

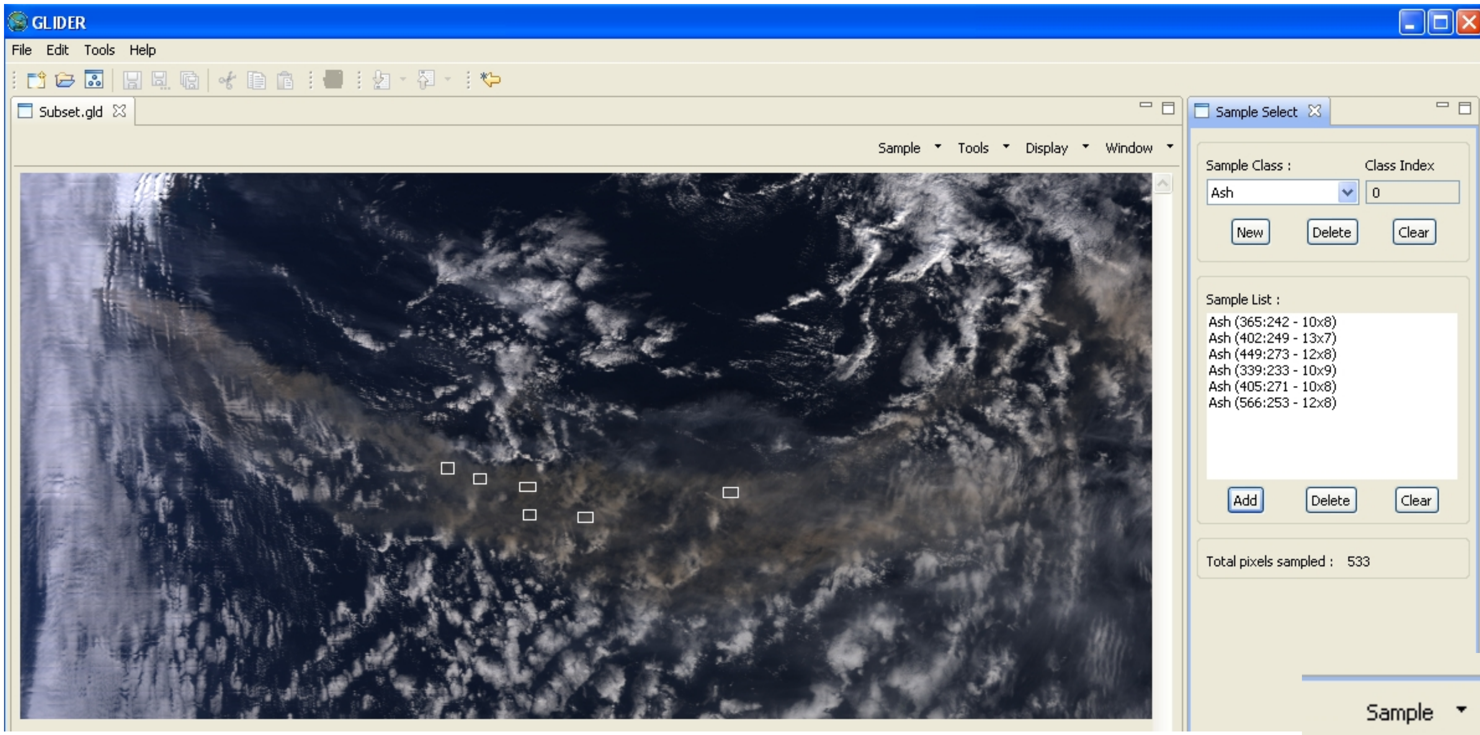


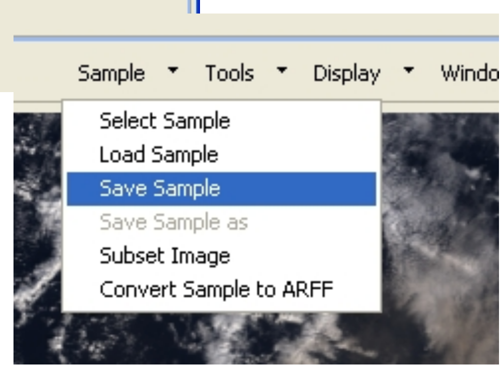
Open Subset.gld in Image View



Take samples by Drawing regions of interest and then clicking Add

Select Samples



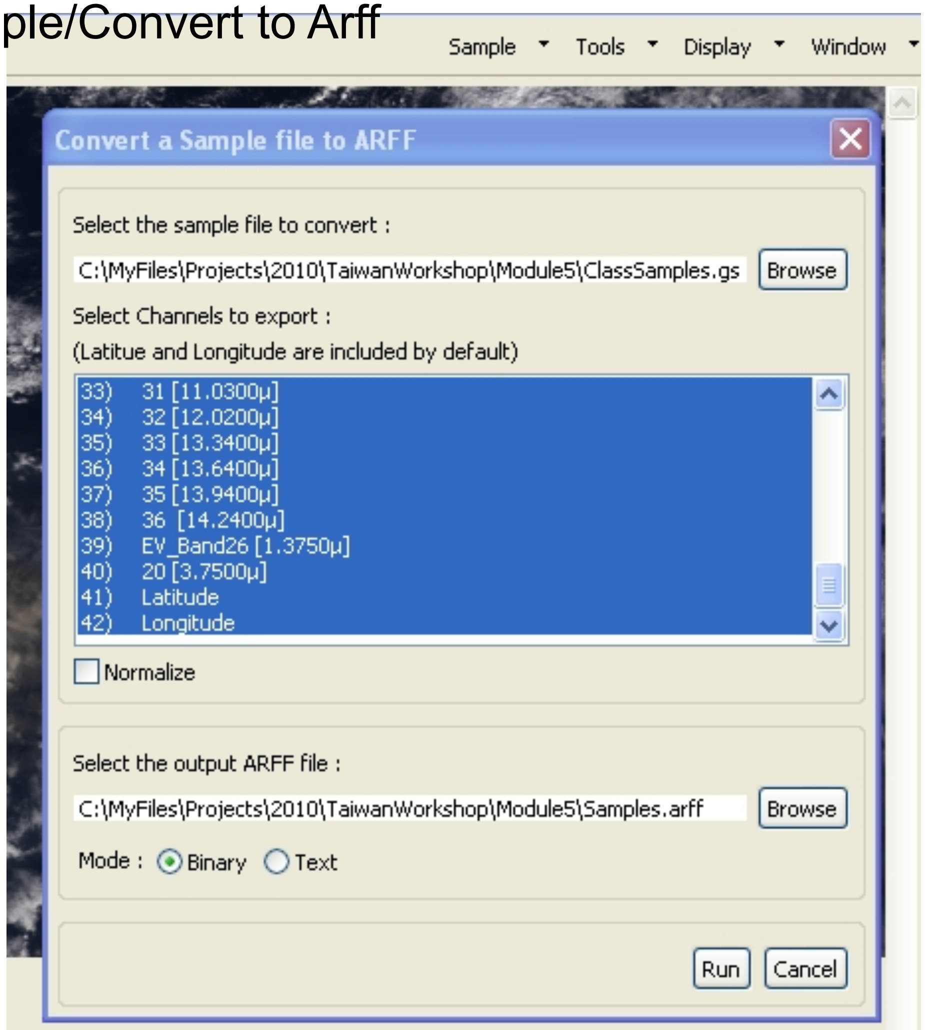


After you have taken enough samples (~1000) for the three classes, save samples as <filename>.gs

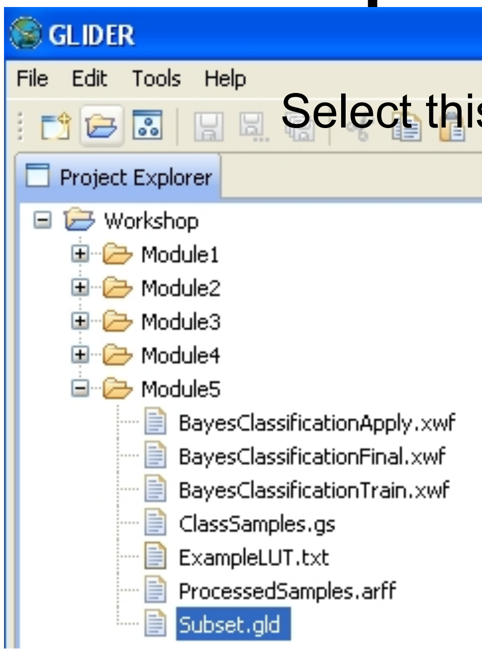
UAHuntsville

Convert Samples to ARFF

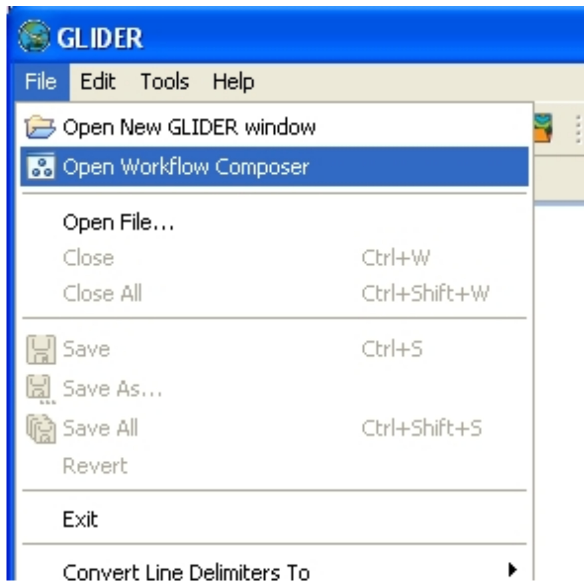
Sample/Convert to Arff

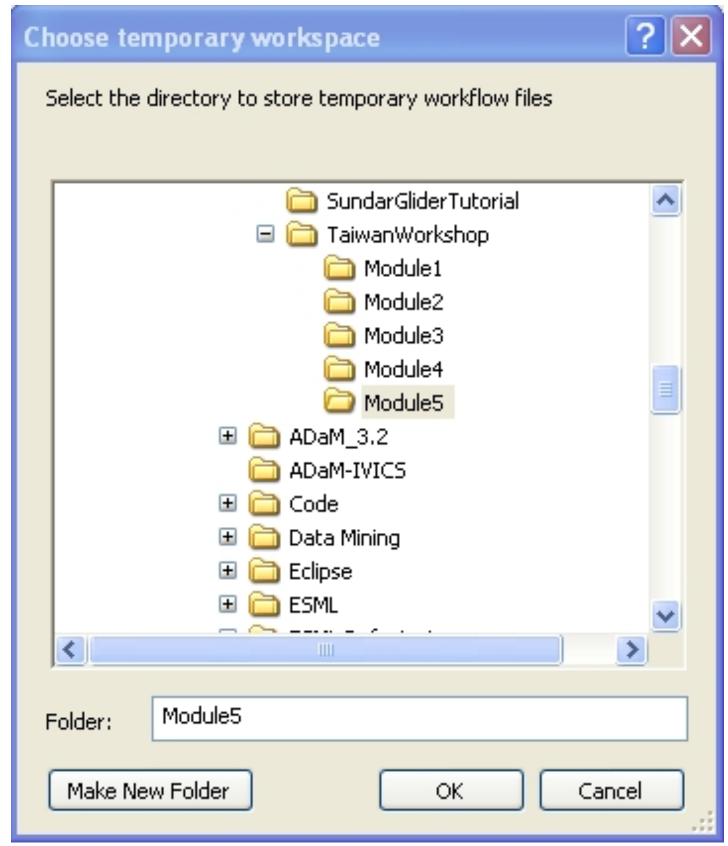


Open Workflow Composer



Select this icon

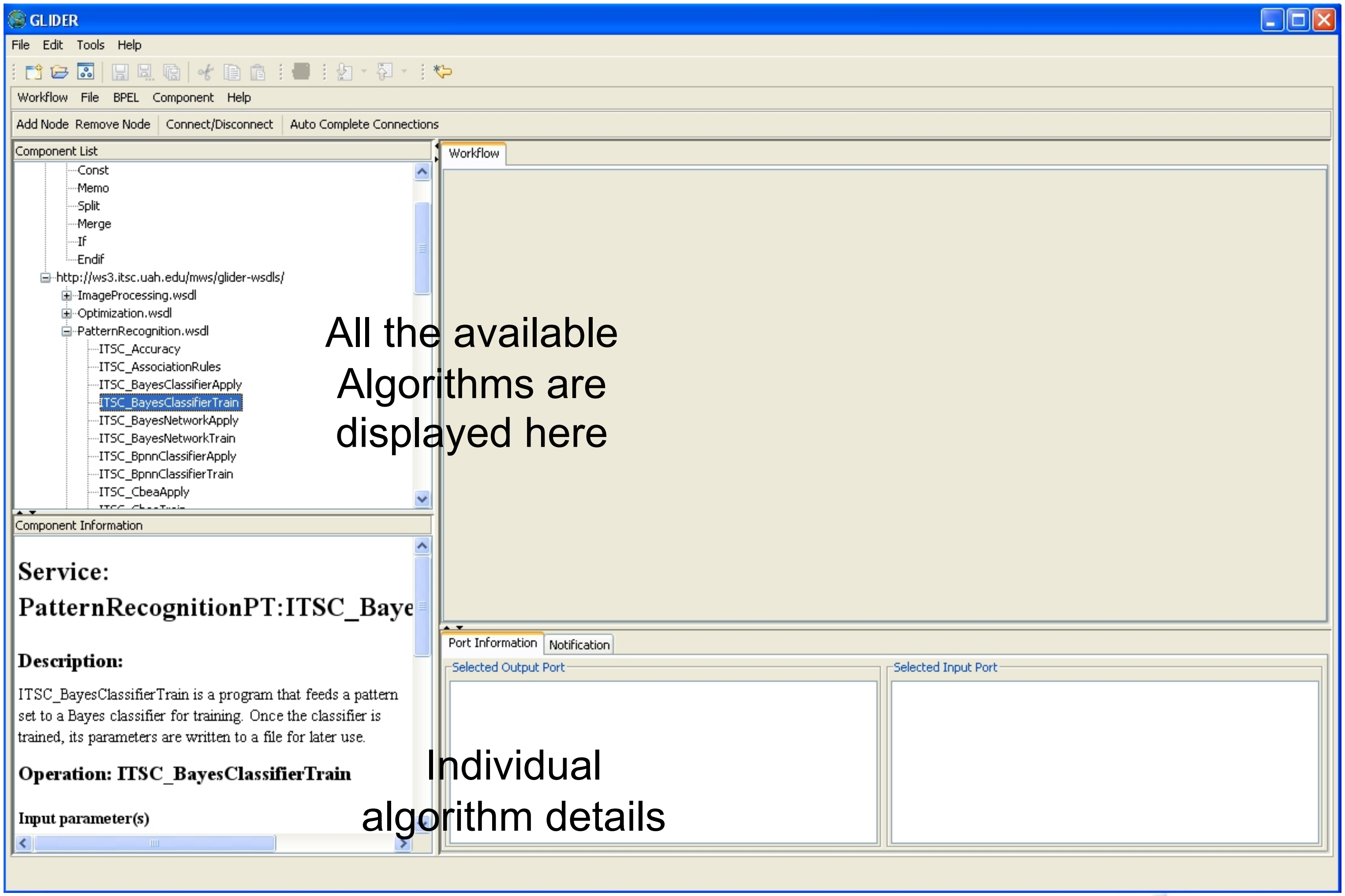




Set your
workspace to your
project folder

UAHuntsville

Workflow Composer

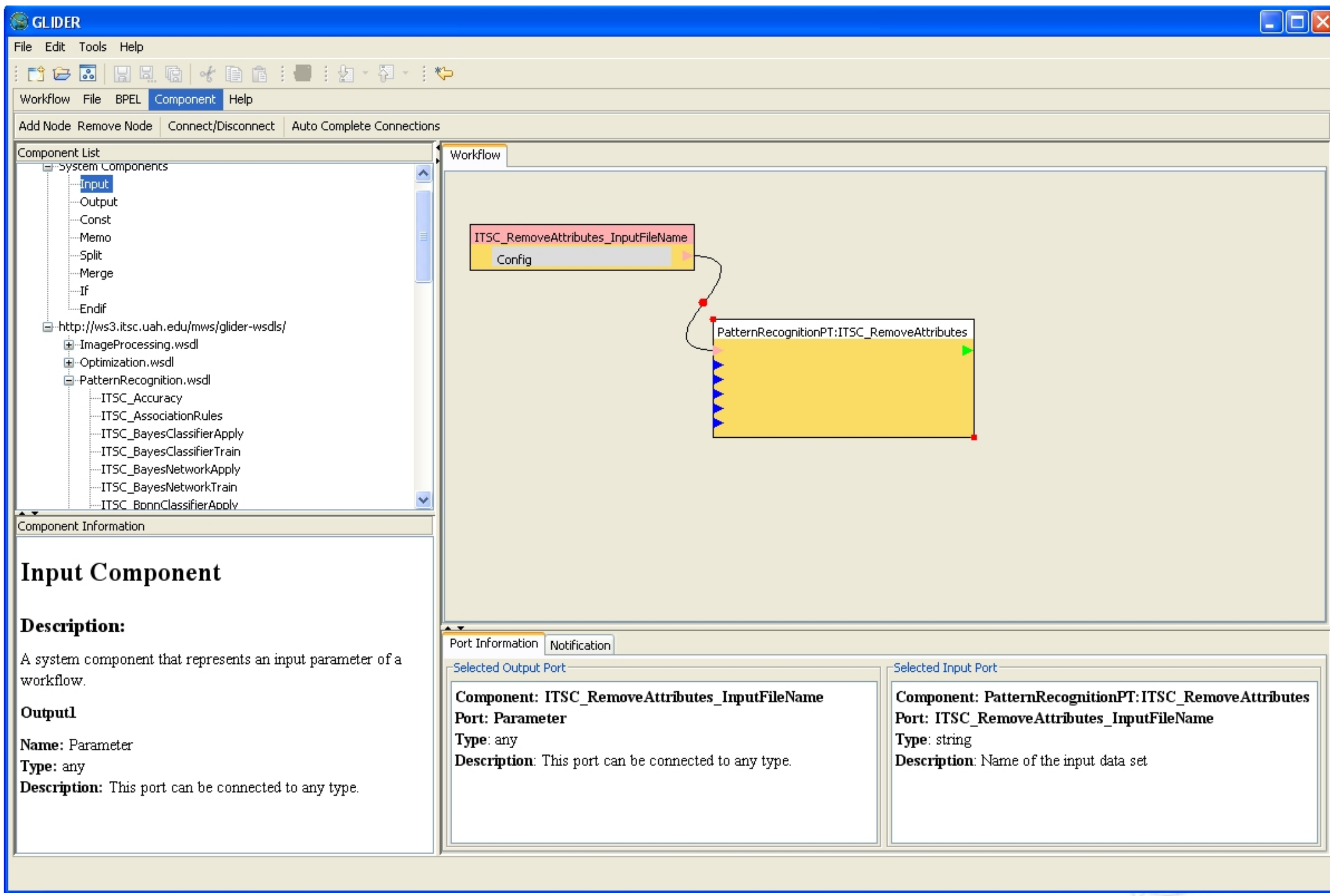


All the available Algorithms are displayed here

Individual algorithm details

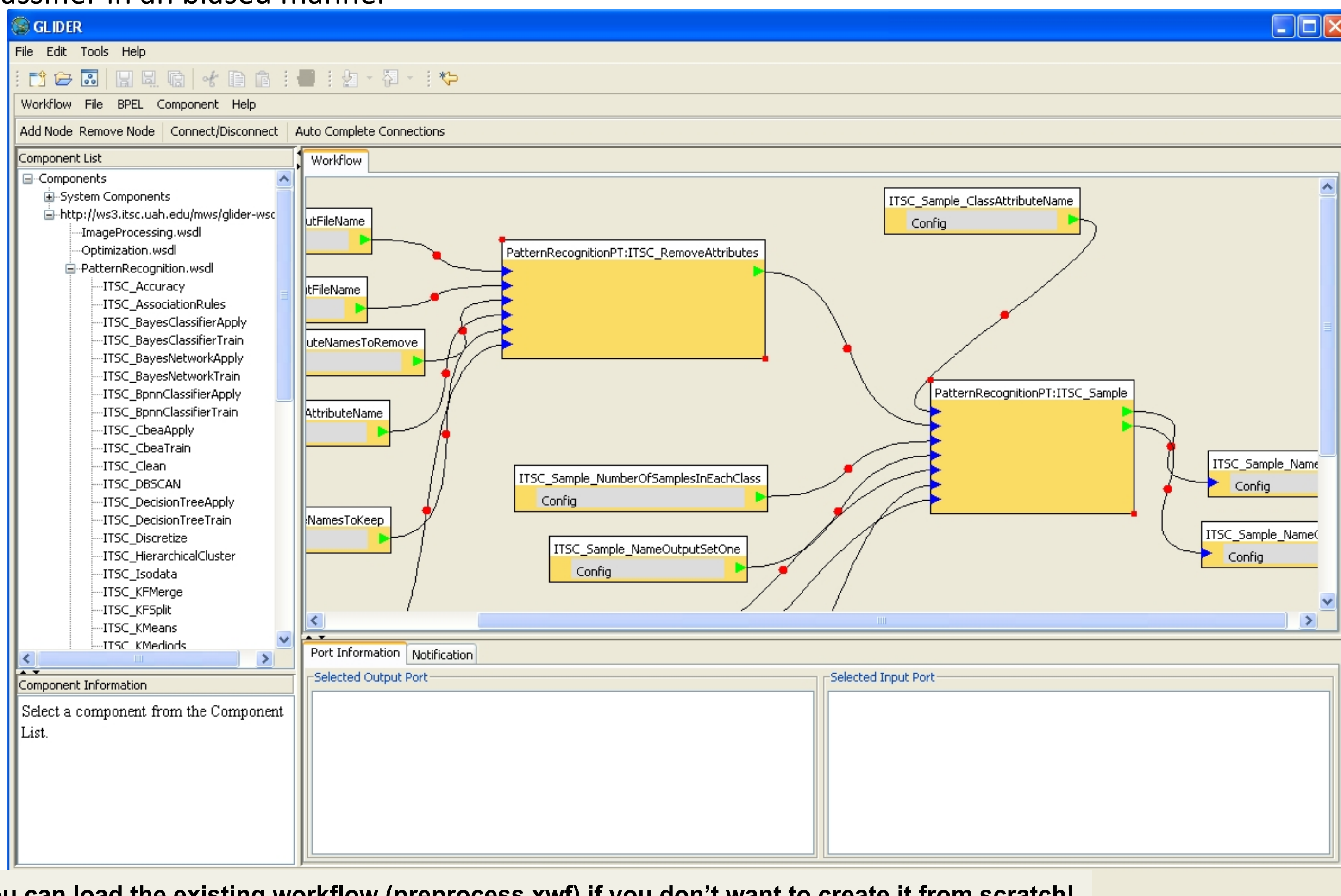
Getting Started with Workflows

- Drag and drop ITSC_RemoveAttributes from the list to the canvas
- Drag input to the canvas and connect it to one of the ports on the algorithm
- Once you have all the algorithms on canvas, you can use the auto completion feature!



First Workflow

- Two step workflow: Preprocess.xwf
  - Remove the unwanted spectral bands
  - Split samples randomly into two files – one for training the classifier, the other for testing the classifier in an biased manner



NOTE: you can load the existing workflow (preprocess.xwf) if you don't want to create it from scratch!

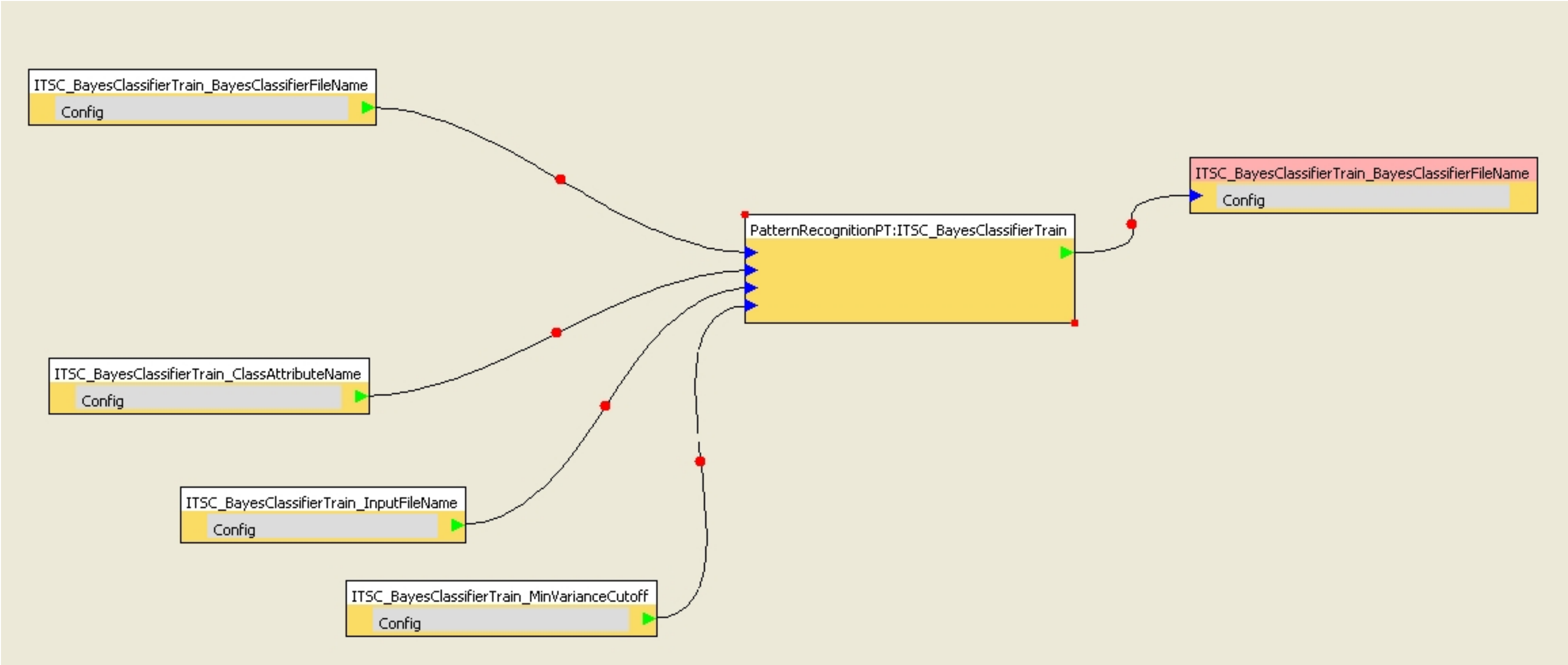
Workflow1: Preprocess - Parameters

- Algorithm1 - ITSC_RemoveAttributes
  - ITSC_RemoveAttributes_InputFileName - \<your path>\ProcessedSamples.arff
  - ITSC_RemoveAttributes_AttributesMask - 010011111111111111111111111110011111111110
  - ITSC_RemoveAttributes_AttributeNamesToRemove - null
  - ITSC_RemoveAttributes_ClassAttributeName - class
  - ITSC_RemoveAttributes_AttributeNamesToKeep - null
  - ITSC_RemoveAttributes_OutputFileName - \<your path>\SamplesFiltered.arff

- Algorithm2 - ITSC_Sample
  - ITSC_Sample_ClassAttributeName - class
  - ITSC_Sample_InputFileName -should be connected to the output from ITSC_RemoveAttributes
  - ITSC_Sample_NumberOfSamplesInEachClass - null
  - ITSC_Sample_NameOutputSetOne - \<your path>\TrainSamples.arff
  - ITSC_Sample_PortionOfSample - 0.5
  - ITSC_Sample_Seed - null
  - ITSC_Sample_NameOutputSetTwo - \<your path>\TrainSamples.arff

Workflow to Train a Classifier

- One step workflow: BayesClassificationTrain.xwf
  - Provide part of the samples to train the classifier
  - Obtain the Bayesian statistics that will be used in the application



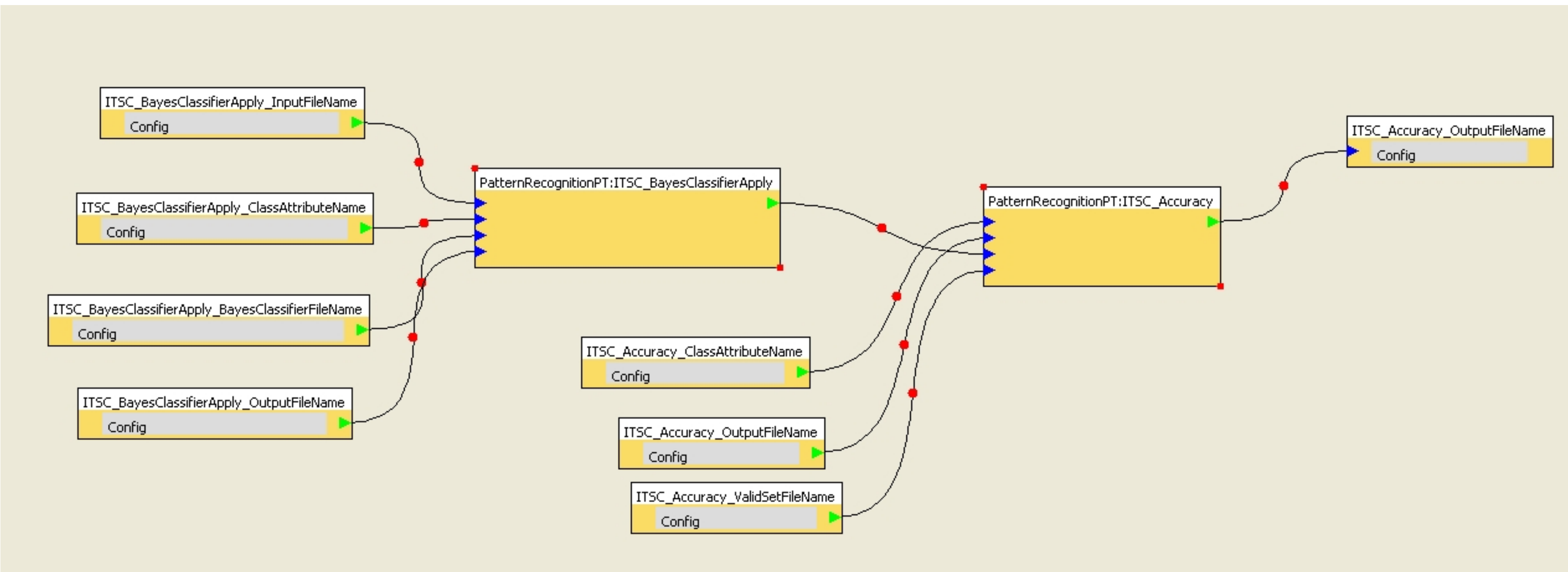
NOTE: you can load the existing workflow (BayesClassificationTrain.xwf) if you don't want to create it from scratch!

Workflow2 - Parameters

- Algorithm - ITSC_BayesClassifierTrain
  - ITSC_BayesClassifierTrain_BayesClassifierFileName - \<your path>\Bayes.txt
  - ITSC_BayesClassifierTrain_ClassAttributeName - class
  - ITSC_BayesClassifierTrain_InputFileName -\<your path>\TrainSamples.arff
  - ITSC_BayesClassifierTrain_MinVarianceCutoff - 0.000000000000000001

Workflow to Test a Classifier

- Two step workflow: BayesClassificationApply.xwf
  - Apply the classifier on the second set of samples
  - Evaluate the classification results (class labels produced by the classifier vs class labels given by the experts)



NOTE: you can load the existing workflow (BayesClassificationApply.xwf) if you don't want to create it from scratch!

Workflow3 - Parameters

- Algorithm1 - ITSC_BayesClassifierApply
  - ITSC_BayesClassifierApply_InputFileName - \<your path>\TestSamples.arff
  - ITSC_BayesClassifierApply_ClassAttributeName - class
  - ITSC_BayesClassifierApply_BayesClassifierFileName - \<your path>\Bayes.txt
  - ITSC_BayesClassifierApply_OutputFileName - \<your path>\BayesResult.arff

- Algorithm2 - ITSC_Accuracy
  - ITSC_Accuracy_ClassAttributeName - class
  - ITSC_Accuracy_OutputFileName - \<your path>\Accuracy.txt
  - ITSC_Accuracy_TestSetFileName - \<your path>\BayesResult.arff
  - ITSC_Accuracy_ValidSetFileName - \<your path>\TestSamples.arff

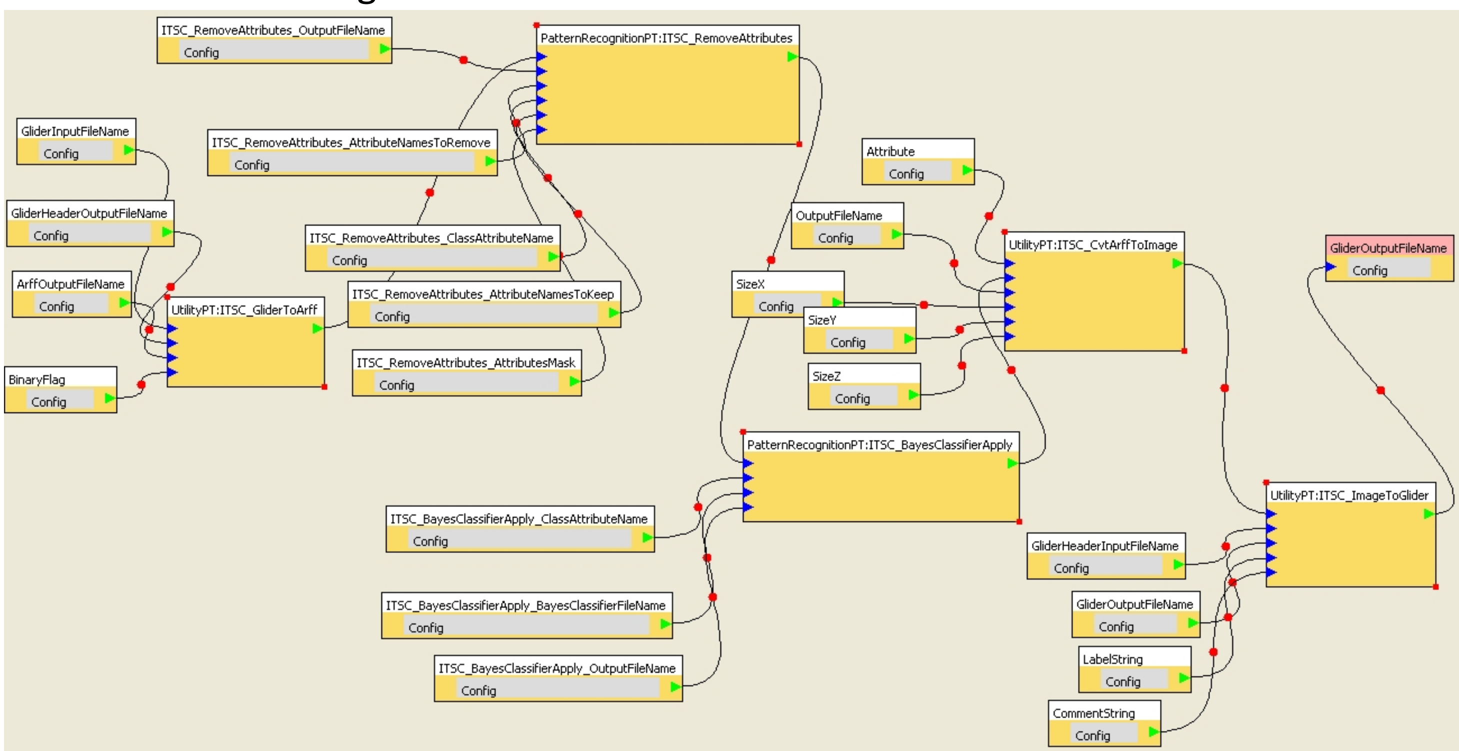
Evaluation Result – Accuracy.txt

```
ITSC_Accuracy - Classes 3, Samples 1540


Confusion Matrix
   |       0      1      2  <--- Actual Class
------------------------------------------
 0 |     487      0      0
 1 |       0    421      0
 2 |       0      3    629

 ^
 |
 +------- Classified As


Accuracy 1537 of 1540 (99.805195 Pct)
```

NOTE: Your numbers may look different!

Workflow to Apply the Classifier on the Image

- Five step workflow: BayesClassificationFinal.xwf
  - Convert data file from GLIDER format to ARFF
  - Remove the spectral bands that you did not use in training the classifier
  - Apply the Bayes Classifier using the Bayesian statistics generated during the training
  - Convert the classification result to image
  - Convert the image to GLIDER format for visualization

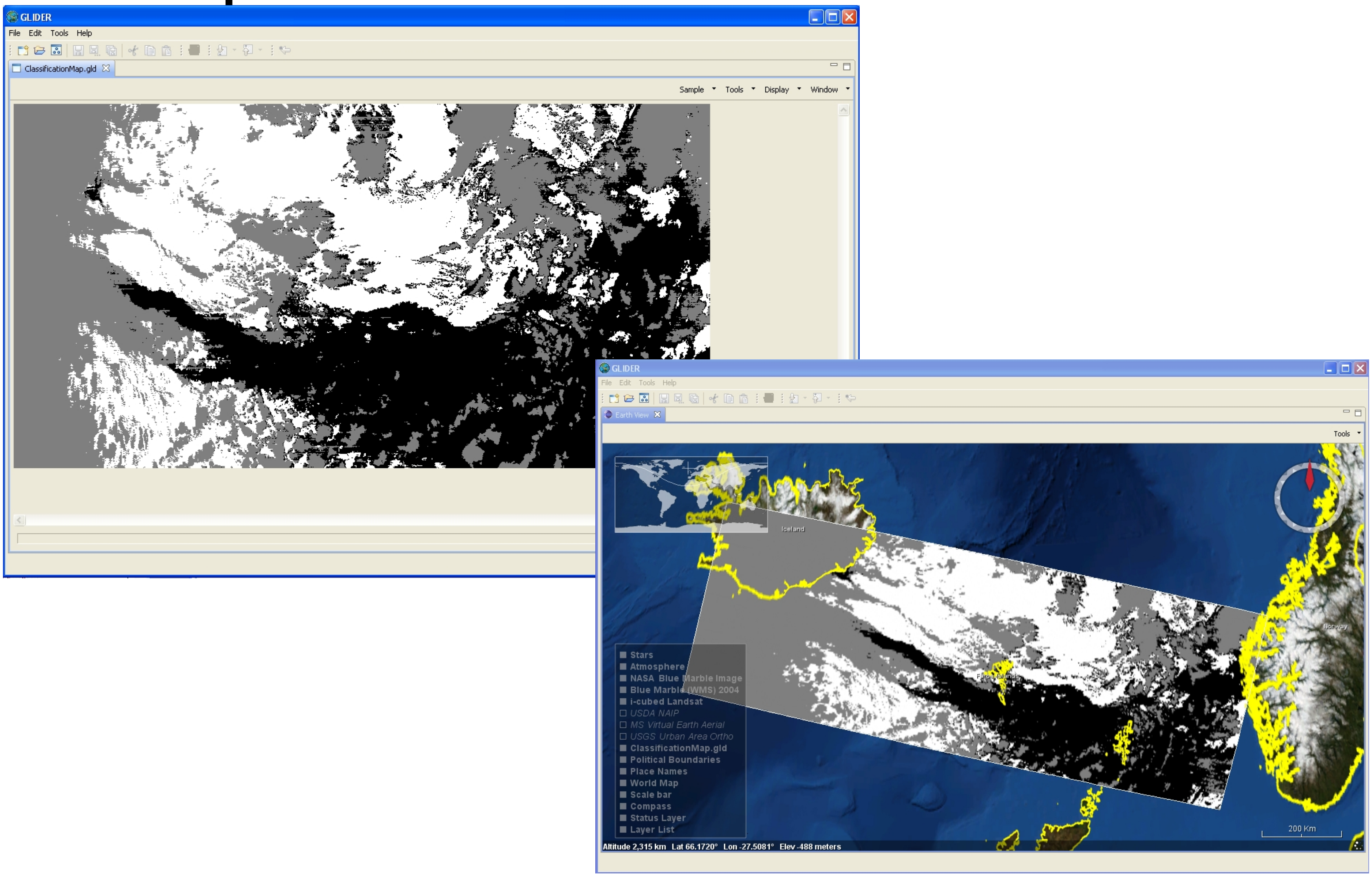


NOTE: you can load the existing workflow (BayesClassificationFinal.xwf) if you don't want to create it from scratch!

Workflow4 - Parameters

- Algorithm1 - ITSC_GliderToArff
  - GliderInputFileName - \<your path>\Subset.gld
  - GliderHeaderOutputFileName - \<your path>\Subset.gh
  - ArffOutputFileName - \<your path>\Subset.arff
  - BinaryFlag - true
- Algorithm2 - ITSC_RemoveAttributes
  - ITSC_RemoveAttributes_InputFileName - should be connected to the output from ITSC_GliderToArff
  - ITSC_RemoveAttributes_AttributesMask - 01001111111111111111111111111110011111111110
  - ITSC_RemoveAttributes_AttributeNamesToRemove - null
  - ITSC_RemoveAttributes_ClassAttributeName - class
  - ITSC_RemoveAttributes_AttributeNamesToKeep - null
  - ITSC_RemoveAttributes_OutputFileName - \<your path>\SubsetFilterered.arff
- Algorithm3 - ITSC_BayesClassifyApply
  - ITSC_BayesClassifierApply_InputFileName - should be connected to the output from ITSC_RemoveAttributes
  - ITSC_BayesClassifierApply_ClassAttributeName - Class
  - ITSC_BayesClassifierApply_BayesClassifierFileName - \<your path>\Bayes.txt
  - ITSC_BayesClassifierApply_OutputFileName - \<your path>\SubsetBayesResult.arff
- Algorithm4 - ITSC_CvtArffToImage
  - Attribute - class
  - ArffInputFileName - should be connected to the output from ITSC_BayesClassifyApply
  - OutputFileName - \<your path>\ClassificationMap.img
  - SizeX - null
  - SizeY - null
  - SizeZ - null
- Algorithm5 - ITSC_ImageToGlider
  - ImageInputFileName - should be connected to the output from ITSC_CvtArffToImage
  - GliderHeaderInputFileName - \<your path>\Subset.gh
  - GliderOutputFileName - ClassificationMap.gld
  - LabelString - ClassificationMap
  - CommentString - 3 Class result

Supervised Classification Result



Music Trivia Answers

- Midnight Oil – Rock Band from Sydney Australia, also known for their political activism especially regarding environmental causes
- Smoke on the Water – Song from Deep Purple, famous for it's guitar riff
- Dust in the Wind – Song from Kansas
- Ashes to Ashes – Song from David Bowie and has Major Tom (astronaut) references

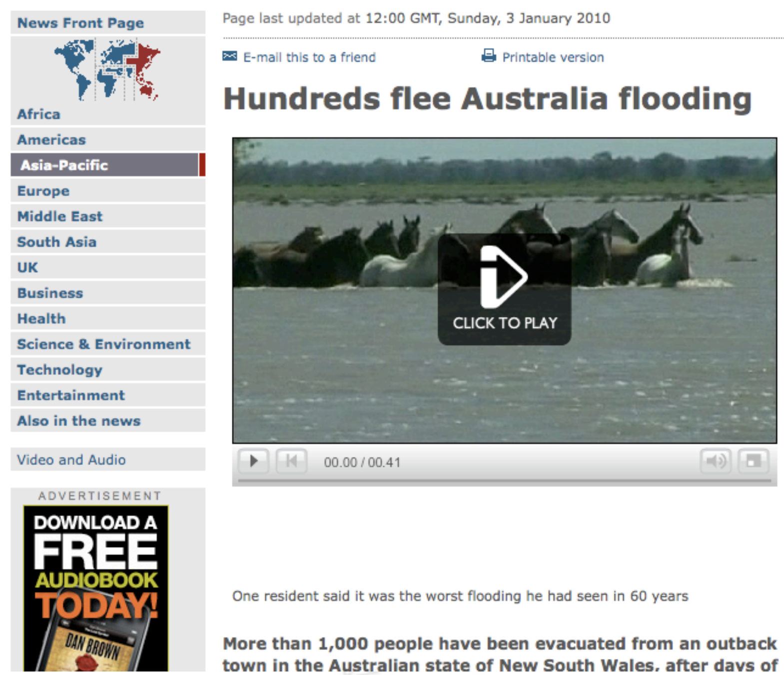
UAHuntsville
THE UNIVERSITY OF ALABAMA IN HUNTSVILLE

Here is your Homework

- Find interesting phenomena observable in satellite imagery
- Order data
- Visualize and analyze using GLIDER
- Submit a microArticle

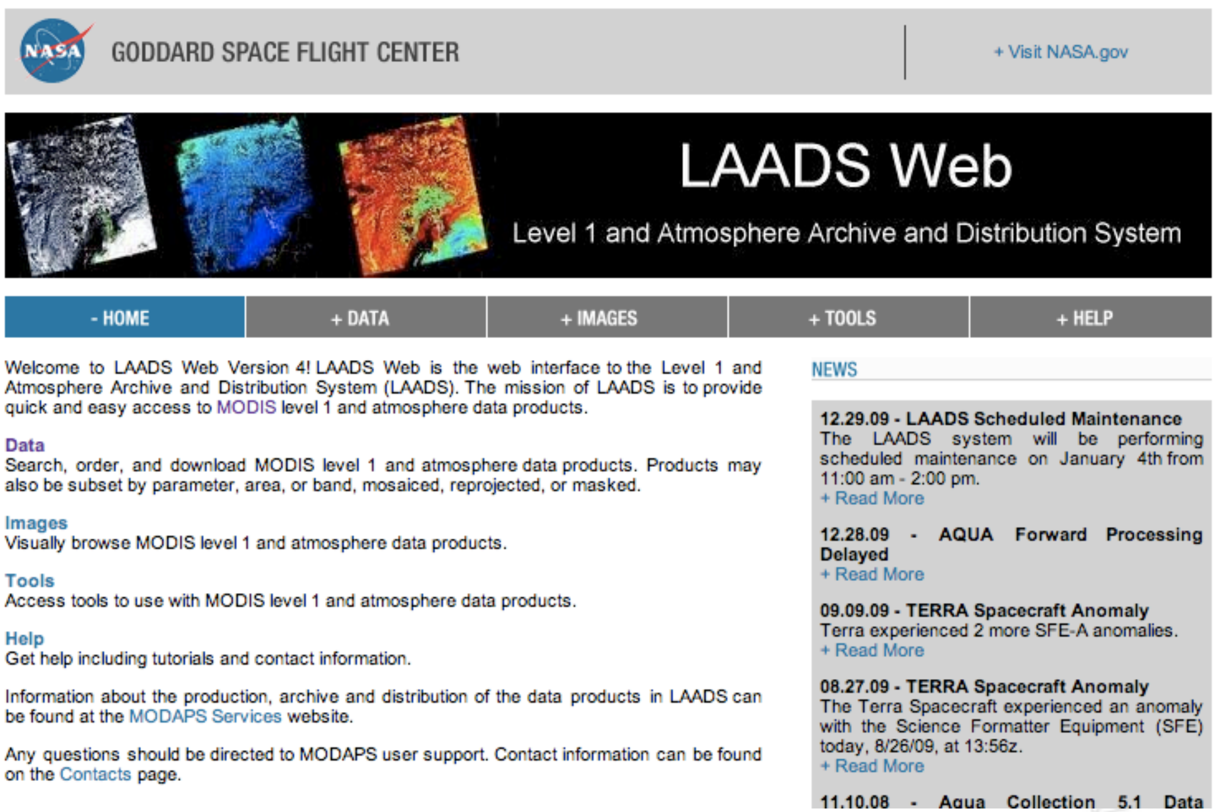
Finding Interesting Cases

- Track environmental news (CNN, BBC)
  - Find the location and time
  - Order data, download,
- Track these websites daily:
  - Earth Observatory: http://earthobservatory.nasa.gov/
  - Operational Significant Event Imagery: http://www.osei.noaa.gov/

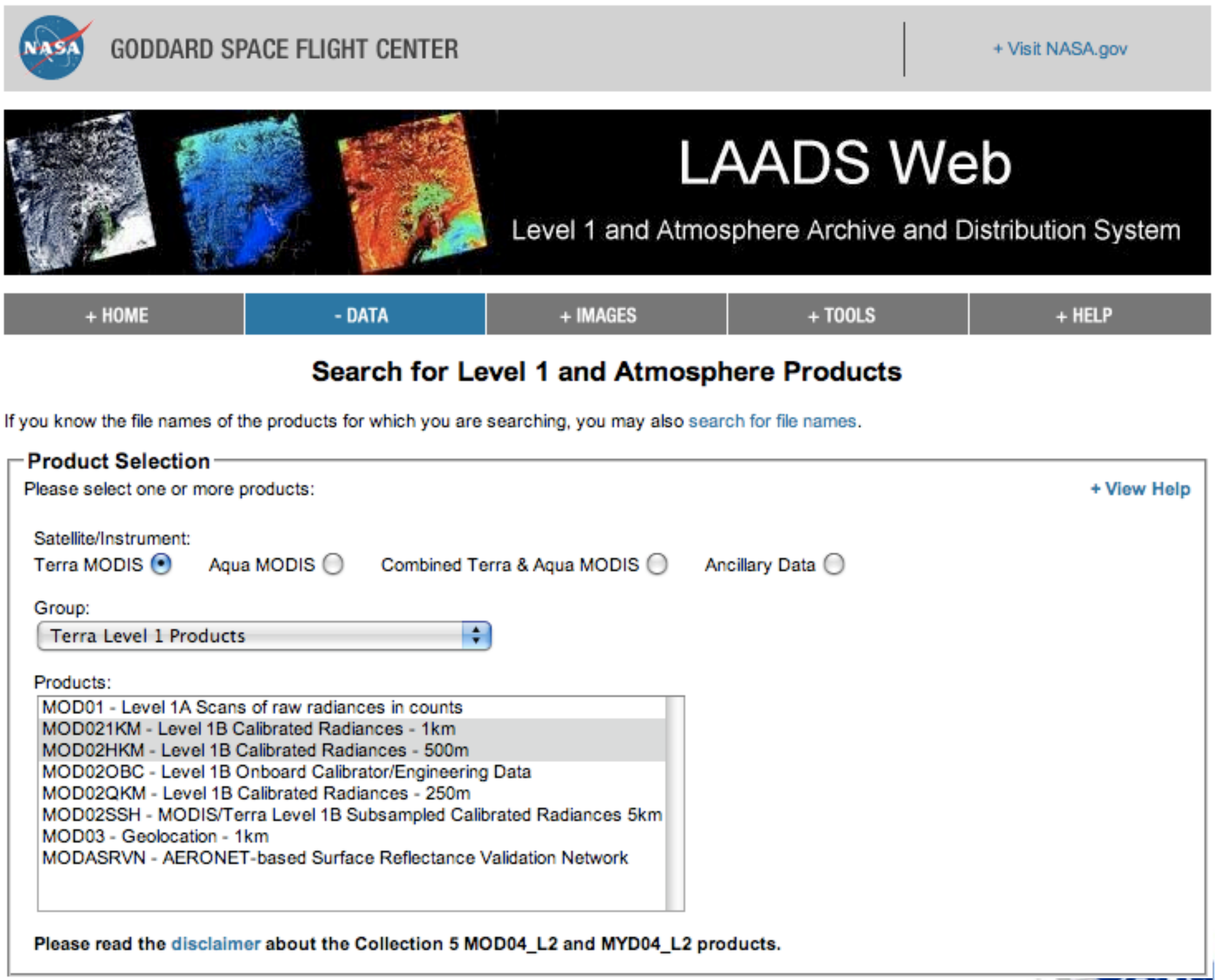


Ordering MODIS Data

- http://ladsweb.nascom.nasa.gov/
- Use L1 data instead of data products
- Information about MODIS (pg 95 – Jensen's Book)

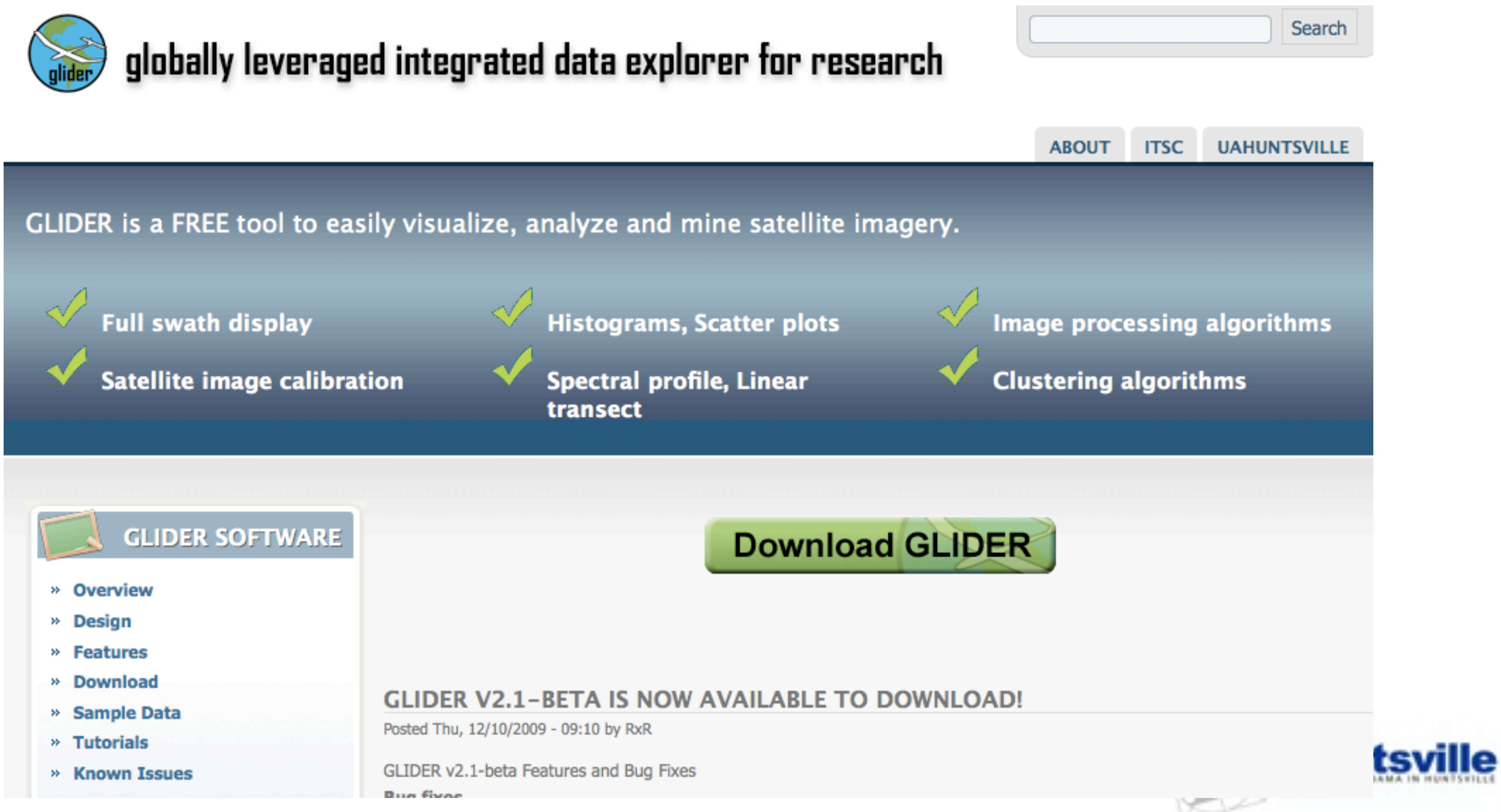


Searching for MODIS Granules



Use GLIDER to examine the data

- GLIDER: http://miningsolutions.itsc.uah.edu/glider/



Submit MicroArticles Here:

www.esphenomena.org

Journal of Earth Science Phenomena

Search

Home | About | Editorial Board | Submission Guidelines | Submission Template | Submit Article | Sign Up | Contact Us

User login

Username: *

Password: *

Log in

- Create new account
- Request new password

Micro-Articles

- What is a micro-article?
- Benefits of micro-article

Text Resize

Phenomena Locations

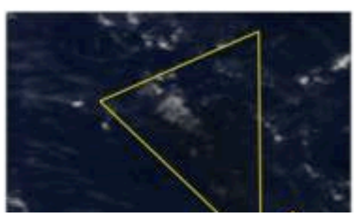
Journal of Earth Science Phenomena (JESP) is an open access journal aimed at publishing micro-articles cataloging interesting and unique phenomena that are observed in Earth science data. The primary aim of this journal is not to report a detailed scientific analysis, but to promote further enquiry, document unique phenomena, assist educational activities and compile the information in a manner that is both searchable and citable. The online nature of the journal also provide for including geographic context, linkages to other geospatial information through Google Earth Technology and also explicit references to online databases where such information could be retrieved.

Key JESP Features:

- *100% digital journal with NO fees for submitting or accessing the journal content*
- *Published micro-articles are peer reviewed and the time from submission to publication is short*
- *All content is open access and can be re-used with proper attribution*
- *Contents covered by Google Scholar*

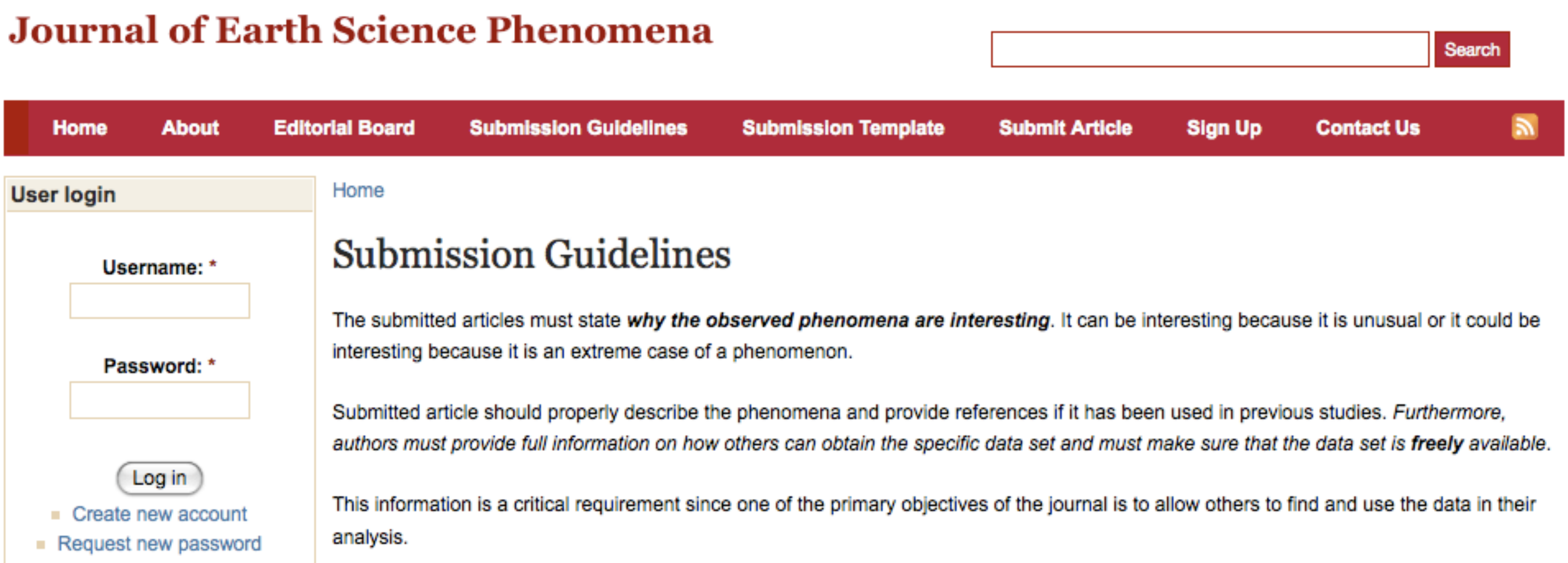
Smoke from an Oil Refinery Fire in Puerto Rico

November 10, 2009 by *rramachandran*



Smoke from an Oil Refinery Fire in Puerto Rico, Rahul Ramachandran, Journal of Earth Science Phenomena, 2009, 15

Submission Guidelines



Journal of Earth Science Phenomena

Search

Home | About | Editorial Board | Submission Guidelines | Submission Template | Submit Article | Sign Up | Contact Us

User login

Username: *

Password: *

Log in

- Create new account
- Request new password

Home

Submission Guidelines

The submitted articles must state ***why the observed phenomena are interesting***. It can be interesting because it is unusual or it could be interesting because it is an extreme case of a phenomenon.

Submitted article should properly describe the phenomena and provide references if it has been used in previous studies. *Furthermore, authors must provide full information on how others can obtain the specific data set and must make sure that the data set is **freely** available.*

This information is a critical requirement since one of the primary objectives of the journal is to allow others to find and use the data in their analysis.

Dr. Rahul Ramachandran

rahul.ramachandran@uah.edu

http://www.rramachandran.com/