Hierarchical Unsupervised Nonparametric Classification of Polarimetric SAR Time Series Data

Ashlin Richardson[1], David G. Goodenough[1,2], Hao Chen[2]

[1]Department of Computer Science, University of Victoria
[2]Pacific Forestry Centre, Natural Resources Canada
ECS Rm. 504, 3800 Finnerty Road, Victoria, Canada, V8W 2Y2
ashy@uvic.ca

International Geoscience and Remote Sensing Symposium (IGARSS 2014) and 35th Canadian Symposium on Remote Sensing
Québec City, Québec, Canada
July 13th-19th, 2014

Natural Resources Canada

University of Victoria

Canada

Why Polarimetric SAR For Forestry?

- All weather
- ***Forest biomass estimation***
- ***Forest land-cover type discrimination***
- Have shown we can ***detect historical fire scars in Canada's western forests (CJRS 2011, IGARSS 2010)***
- Canada is heavily invested in satellite radar

- For Quad-Pol SAR, information is hiding in the interplay between the intensities and phases!
- **More accurate forest information can be retrieved if exploited fully!**
  - **The future:** high resolution multi-frequency multi-temporal tomography: **Reconstruction in time and space**

Quad-Pol Analysis for Forestry

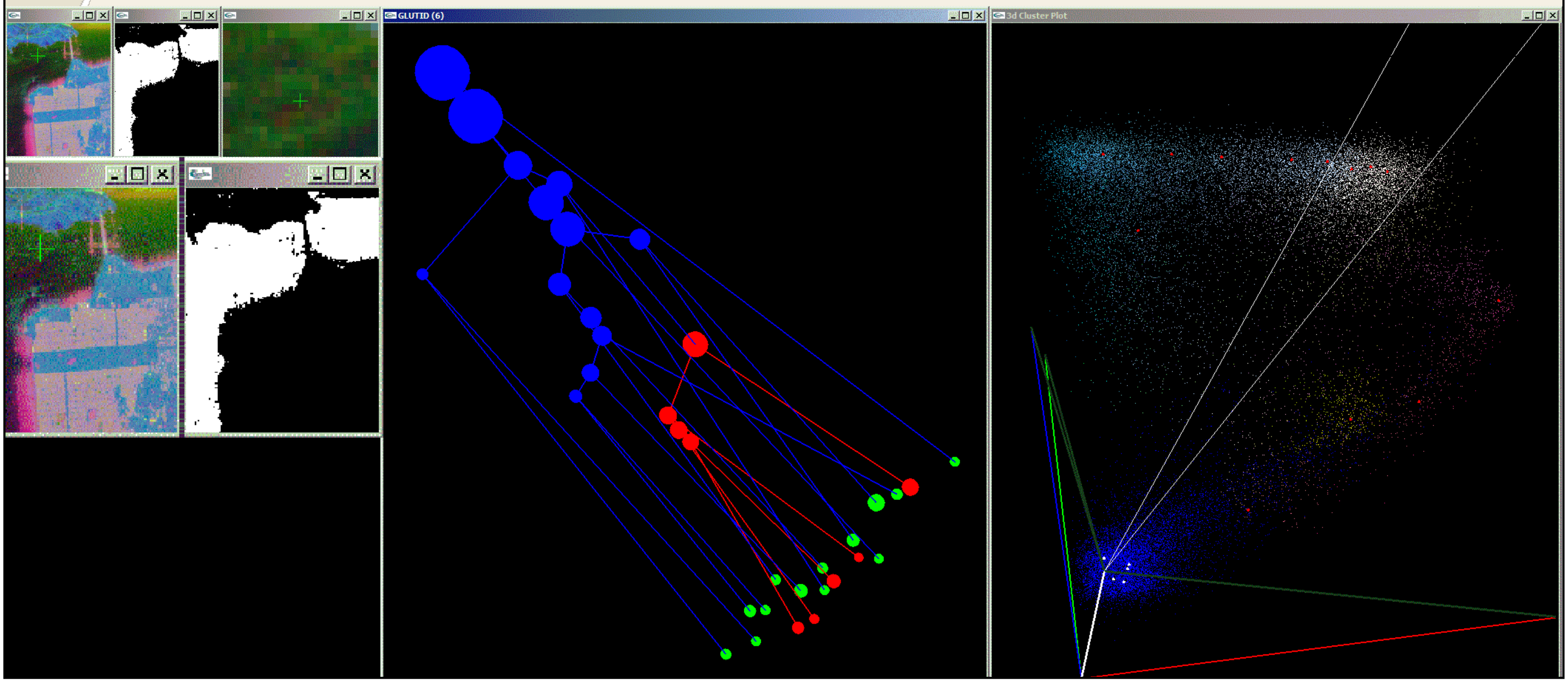
- Quad-Pol analysis.. **Several ongoing challenges:**
  - Need a physically meaningful understanding of decomposition parameters.
  - Need to visualize high dimensional data, and **classify it effectively**

- Present work by University of Victoria (UVic) and Canadian Forest Service (CFS, NRCAN) seeks to address these issues through new algorithms R&D:
  - ***Real-time interactive data visualization***
  - ***Data Driven Classification***

Rapid visual data exploration

Essential for data understanding.. Explore discrimination potential of the data!!!

- Interactive image, dendrogram, and 3-d scatter displays.
- **Change parameters “on the fly”: see resulting clusters immediately.**
- Explore the hierarchy of clusters by clicking on 1) **the image** 2) **the tree** (middle), or 3) cluster centers in **the feature space** (right).

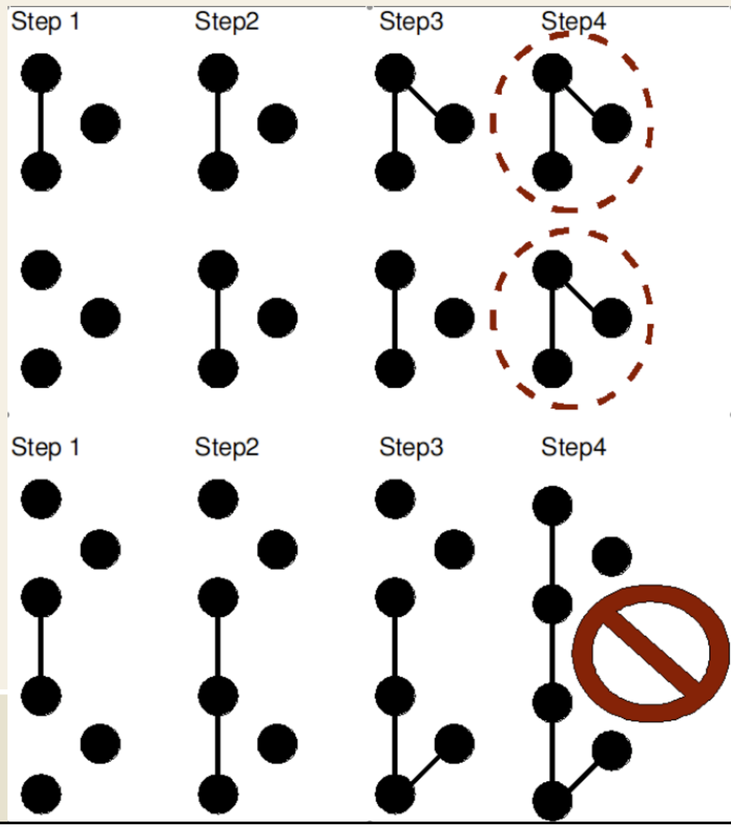


Land Cover Discrimination – why use geometry based methods?

- Classical agglomeration algorithms sensitive to initialization.
- **Poorly separated clusters** reveal mutual inconsistences between: merging procedure, cluster shapes, & distance function. Also: **want to know how clusters connect!!!**

Initialization #1

Initialization #2



$$d(X, Y) = \frac{1}{|X||Y|} \sum_{x \in X} \sum_{y \in Y} d(x, y)$$

Classification – why use non-parametric methods?

- **Data is NOT GAUSSIAN.**
- Textbook assumption: Quad-pol data Wishart/modified-Wishart distributed.. Cluster is described by a single representative
- Wishart assumption implies "round" shaped clusters.. For "non-round" clusters, merging is problematic
- **Cluster description may require multiple elements.**
- **More general approaches e.g. geometry/point-clouds, may give improved results.**

Notes:

- **Want to avoid explicitly setting the number of clusters.**
- Not all clusters may be meaningful (on the ground).

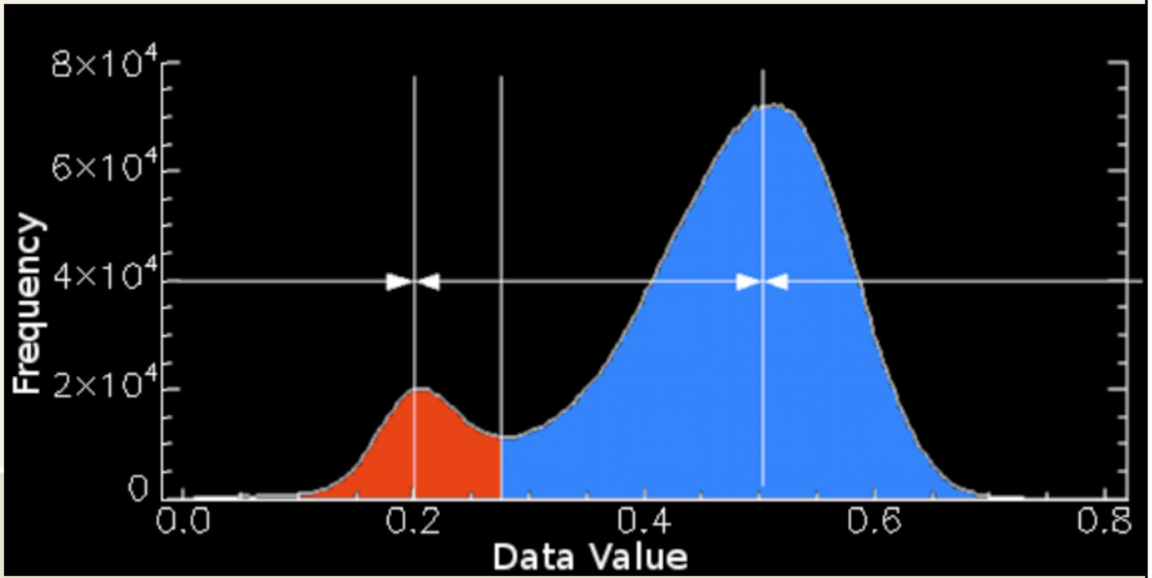

Q: What is mapping?????
Multi-scale
Multi-scale
Multi-scale
Multi-scale

Possible Answer: Modes (Peaks) And Interconnectivities

- Wishart (1969): **clusters are the modes (hilltops) of a density.**
- **A peak (an attractive center) is deemed to represent a cluster..**
- All points attracted to the peak belong to that cluster
- **Method: Climb to the top!!!!**

Histograms represent estimated density functions - ***one dimensional motivational example:***

Shown: two peaks and associated domains of attraction (red & blue).



Geometry-based approach: KNN Graph Clustering (KGC)

IGARSS 2010 (KGC-1)

- Inspired by Wishart's **Mode Analysis.**
- **Data driven;** only one parameter (K).
- **Finds clusters of arbitrary shape** by considering the K-Nearest Neighbors about each point, and their interrelations
- No direct assumption about cluster number, or underlying parameterized distribution (e.g. Wishart).
- Stable in terms of both the number of clusters, and the cluster shapes (with respect to the parameter K).

KGC-1 method: description

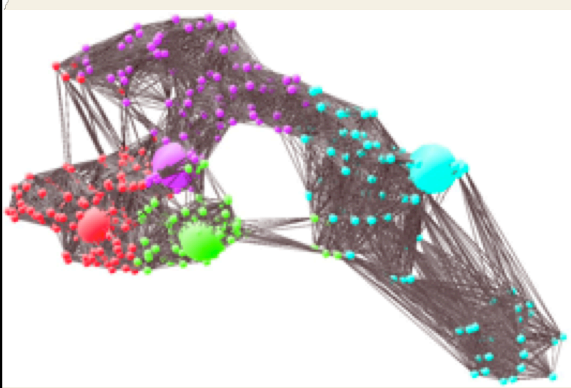
1) Construct the K-NN graph.. Compare everything to everything else!! ***AKA compute the distance matrix.***

2) Estimate density at each data point.

$$\rho(x) = \frac{1}{\frac{1}{K}\sum_{n \in N} d(x,n)}$$

- **x** is a data point, **K** is the number of nearest neighbors of **x**, and **N** is the set of nearest neighbors of **x**
- Density taken as reciprocal of average neighbour-distance

3) Mode seeking on the K-NN graph.. find the peaks!



- **Method: "If my density is higher than my neighbours', I'm a peak! Otherwise, climb up!** (traverse to my highest-density neighbour)"
- **Associate peak with a basin of attraction:** A peak and the data points which "climb up" to that peak.. are given the same color (class label).
- **Left: example with 4 classes!** (K-NN graph in 3-d..K=40) PALSAR data sub-area
- Data points: small spheres. **Peaks: large spheres.**

Geometry-based approach: KNN Graph Clustering (KGC)

<u>CSRS 2013 (h-KGC)</u>

- **<u>Full Hierarchical operation</u>:** Cluster merging consistent with density observations. **Estimation of cluster tree of density!**
- **Interactive pixel & cluster selection, <u>hierarchy traversal</u>.**
- **Application to large data sets (exploiting multi-core architecture) ➔ offload computationally intensive neighbor calculations.**
- All data points are used (IGARSS 2010 used a sampling approach). ➔ **More accurate representation of low-density clusters. Warning: still computationally intensive.**

h-KGC method: Description

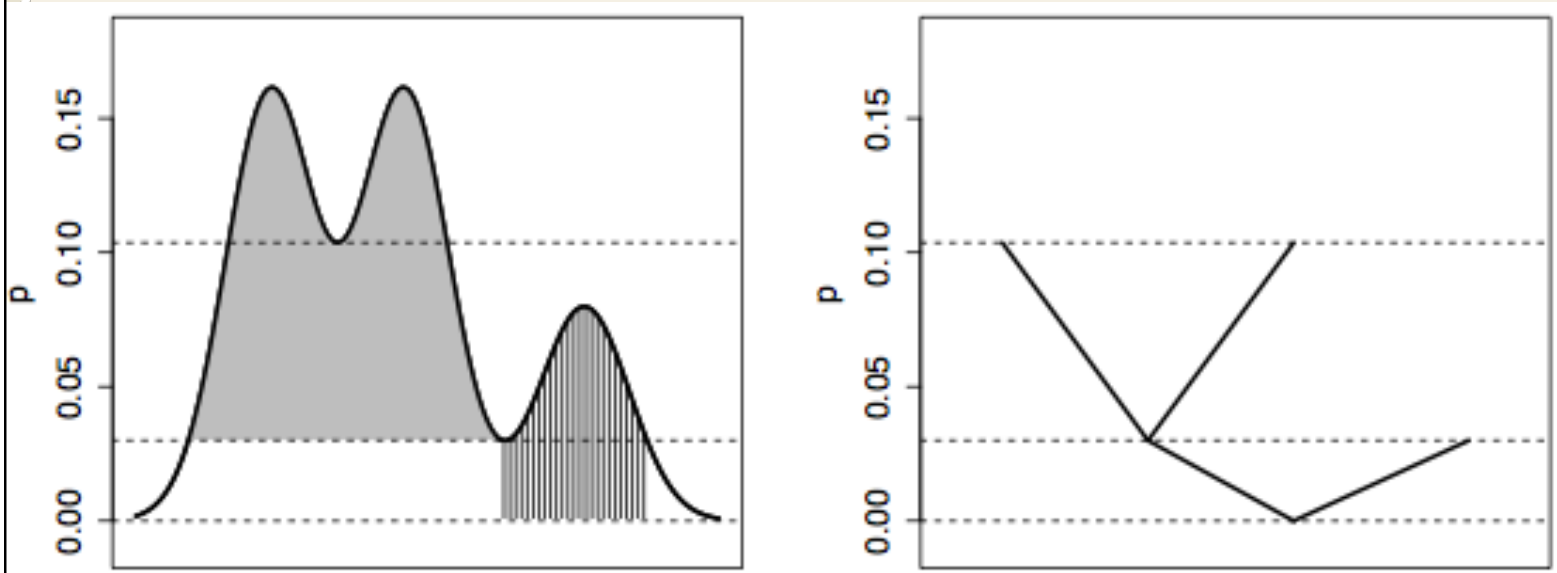
- ***Fully Hierarchical*** operation:

  Estimate Cluster Tree (***DENDROGRAM***) from density estimate!!!

(a) Find modes associated with Density estimate

(b) Estimate the Cluster Tree (math: “persistence of connected components of upper-level sets” $L(\lambda;\rho)=\{x|\rho(x)>\lambda\}$

- The splitting ordered in terms of the density ➔ interpretation in terms of statistical significance. Model-based splitting may not fit the data, and may not have a probability interpretation.



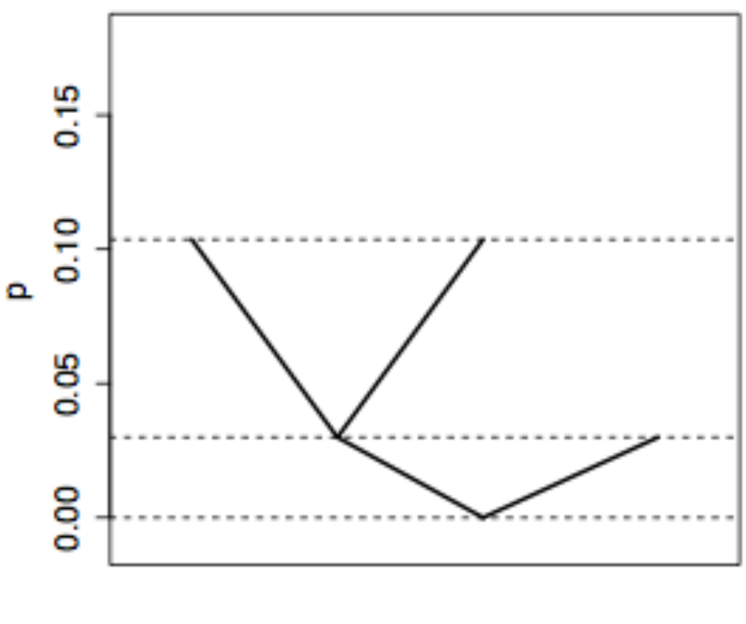
h-KGC method: Description

- **Hierarchical operation: Estimation of the cluster tree (yes, a Binary Partition Tree (BPT)**

Interfaces between modes

Suppose modes M_i, M_j share an interface:

$$\begin{aligned} I_{i,j} = I_{M_i,M_j} &= I_{M_j,M_i} \\ &= (N(M_i) \cap M_j) \cup (N(M_j) \cap M_i) \end{aligned}$$



(b)

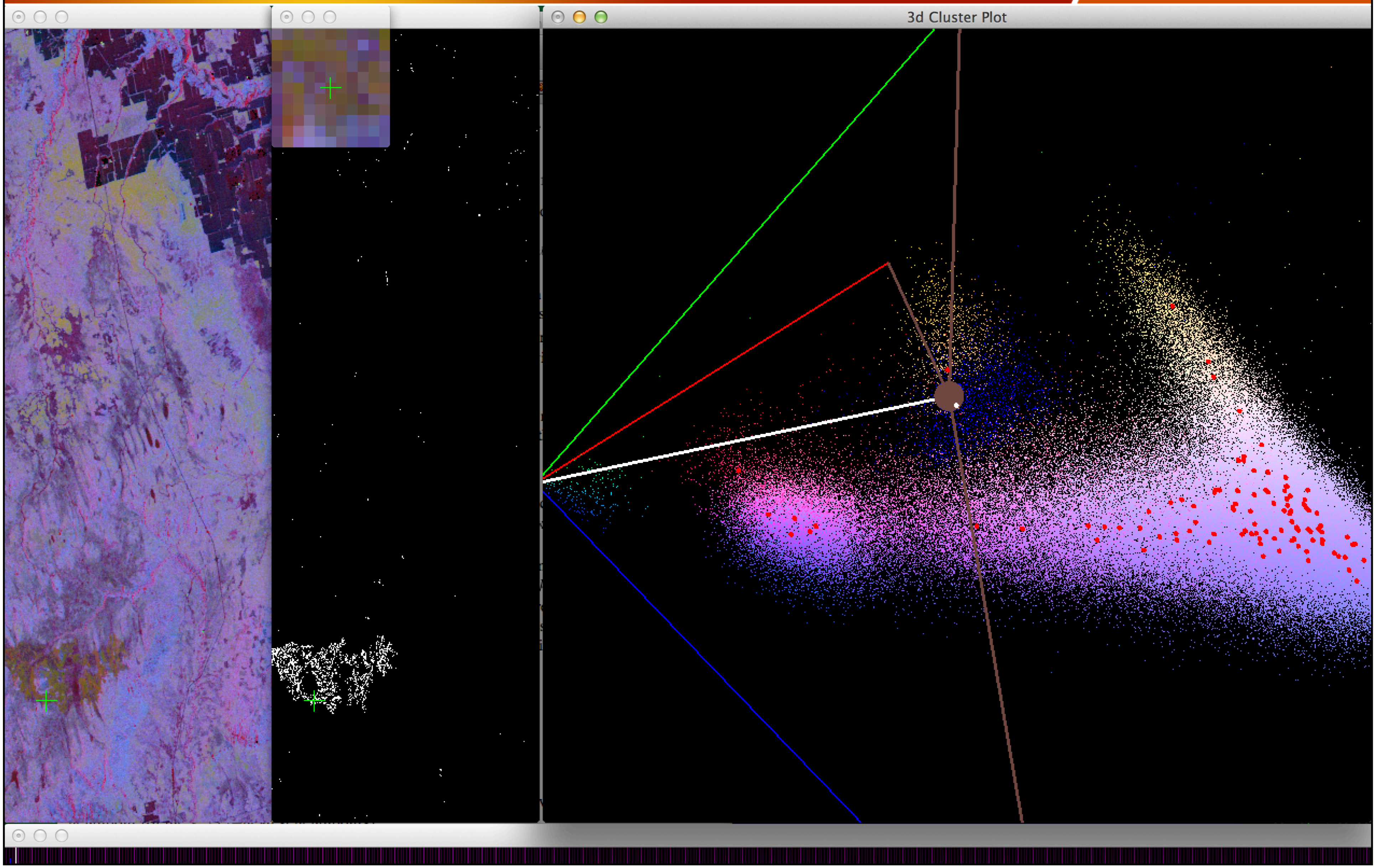
Waterfall Algorithm:

1) **Find interfaces (bridges between peaks)**
2) **Perform merging (in order of highest interface density, to lowest)**

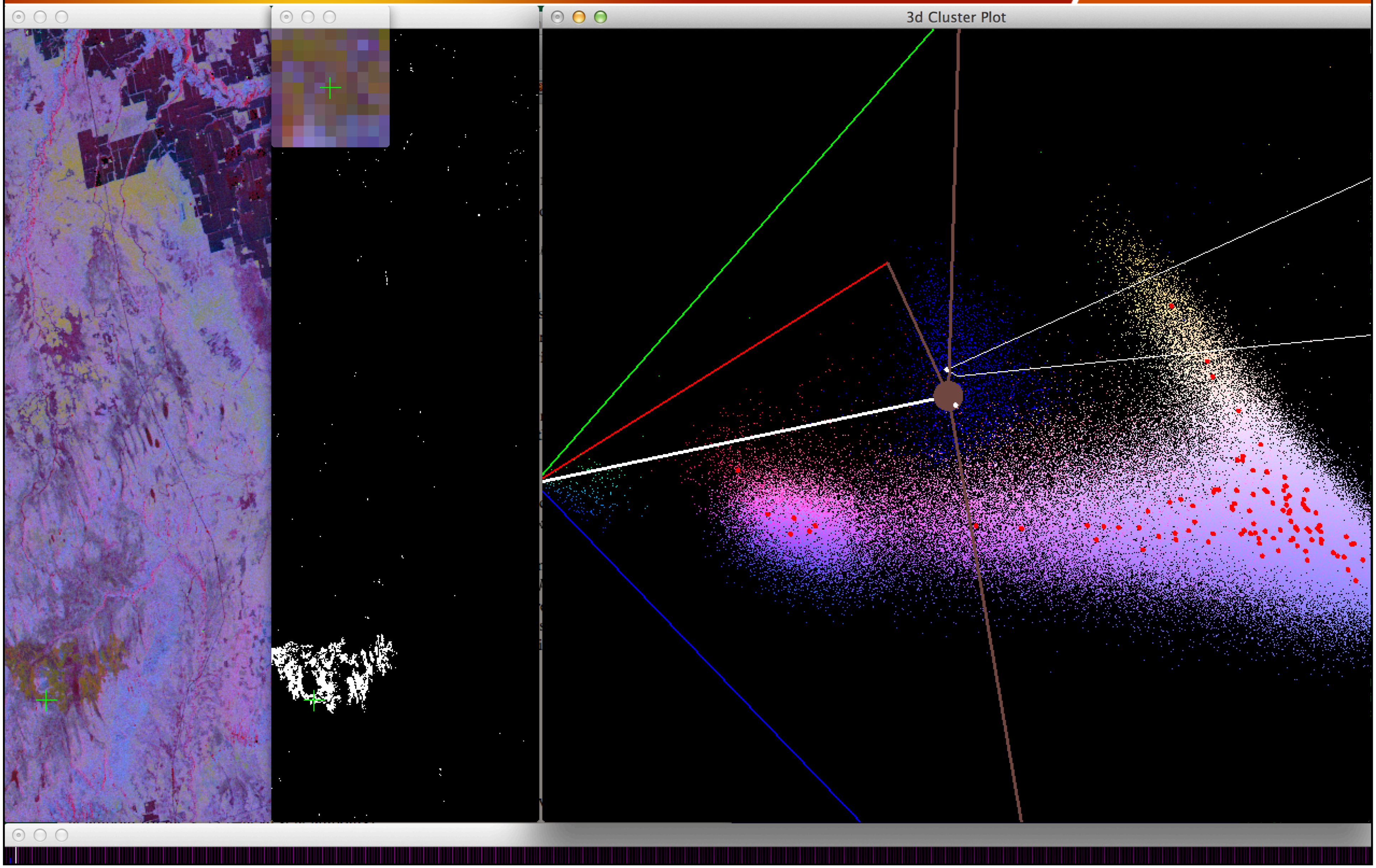
Data structure: disjoint set-forest with path-flattening!

Left: dotted lines ➔ interfaces!!

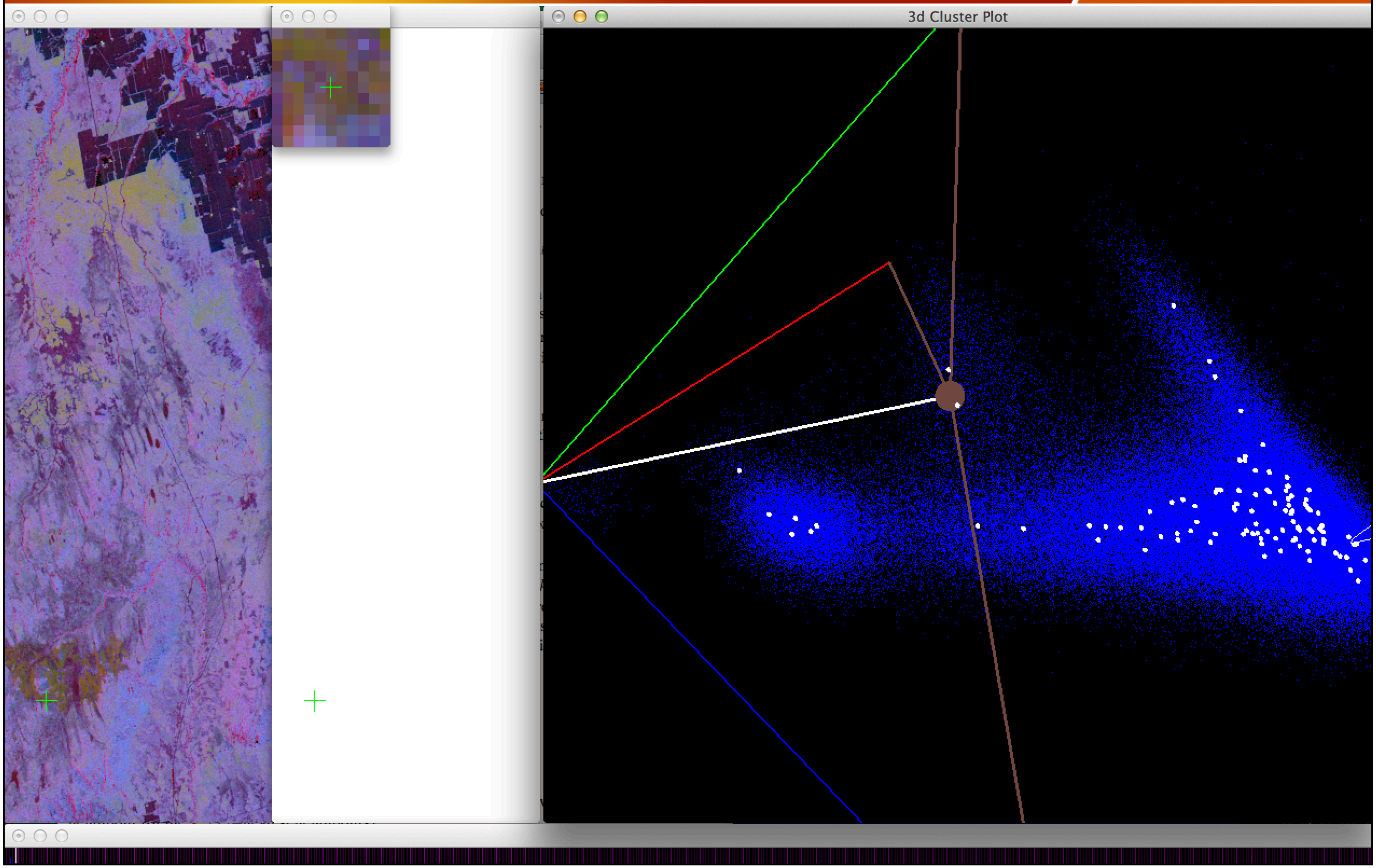
Interactive Node Selection
L-band ALOS near Keg River



Interactive Node Selection
L-band ALOS near Keg River



Interactive Node Selection
L-band ALOS near Keg River



KGC-3 method: description (1/3)

1) Perform nonparametric (local) segmentation

~~*Classical solution: waterfall algorithm on image intensity (← bad.. ignores multivariate interrelationships).*~~

Generic method: given an image metric,

e.g., $d(x,y)=|x-y|^2$

augment that metric with spatial information

e.g., $d(x,y)=|x-y|^2 + |i_x - i_y|^2 + |j_x - j_y|^2$

And perform 1) Distance matrix calculation

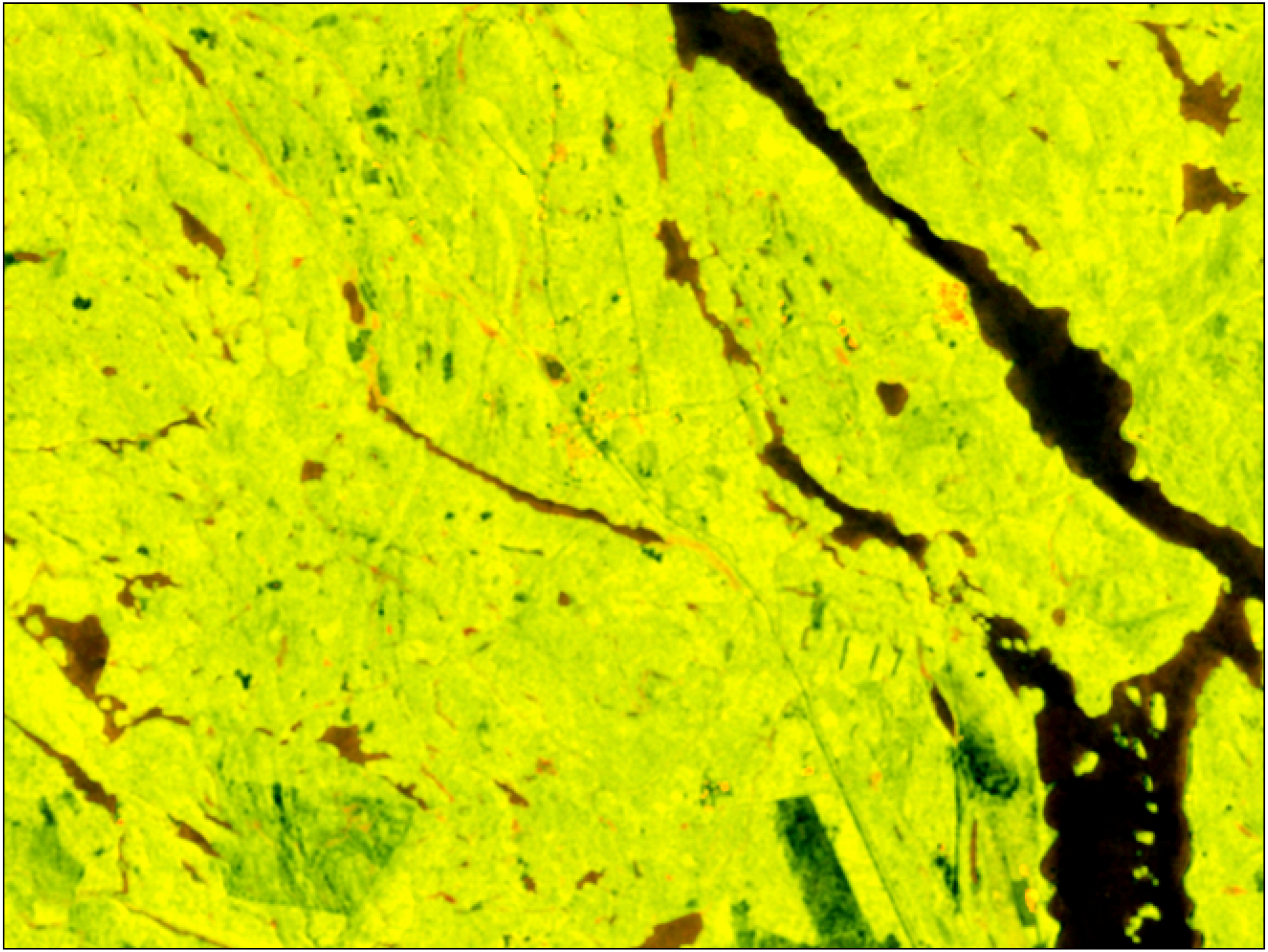
2) Density estimation

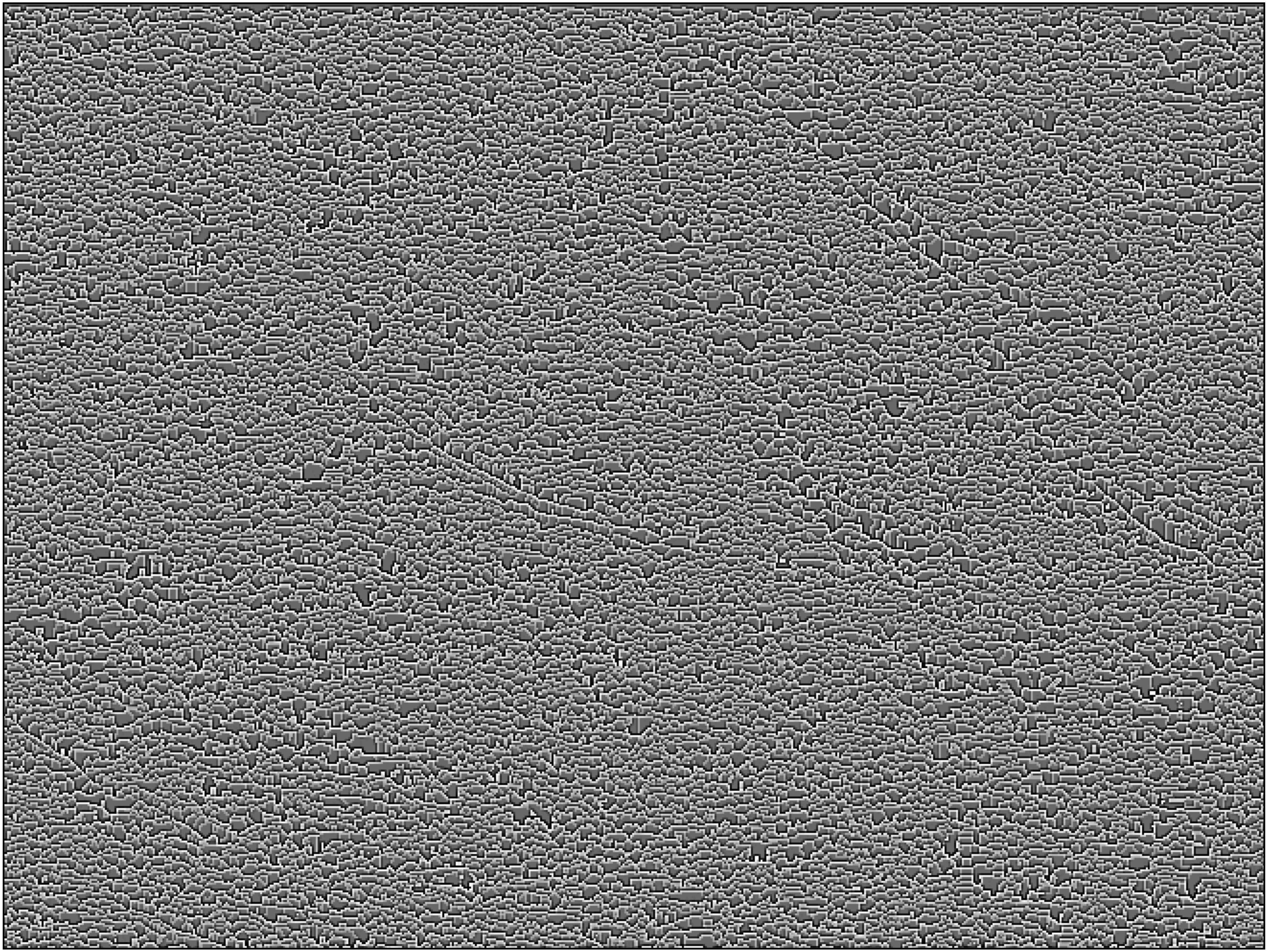
3) Hill climbing

$$\rho(x) = \frac{1}{\frac{1}{K}\sum_{n\in N} d(x,n)}$$

Advantage: fast!!! Consider only neighbors in a small window

Result: small patches. Next slide..Example (3x3 window).





KGC-3 method: description (1/2)

2) Next, on the segments perform:

1) ***Distance matrix calculation***
2) ***Density estimation***
3) ***Hill climbing***

We used:

- ***$d(x,y)=|x-y|^2$ with x and y mean vectors for a pair of patches (segments).***
- ***Same rudimentary (effective) density estimate.***

3) Perform the merging. Result: a tree!

Next slide.. Visualizing the result for R2 FQ9 stack (averaged) parameters (water class selected..)

Results: Data and Comparison

RS2 FQ9 Date
20081231
20090124
20090217
20090711
20090804
20090828
20090921
20091015
20091108
20091202
20100212
20100823
20101010
20101103
20101127
20101221
20111122

Processing

1. ***Multilook 4x2***
2. ***Boxcar 5x5***
3. ***Coregistration***
4. ***Boxcar 3x3***
5. ***Multilook 2x2***
6. ***TS averaging***

7 Parameters:

Alpha, Entropy, Anisotropy,

Shannon Entropy,

RVOG: Biomass, surface, volume

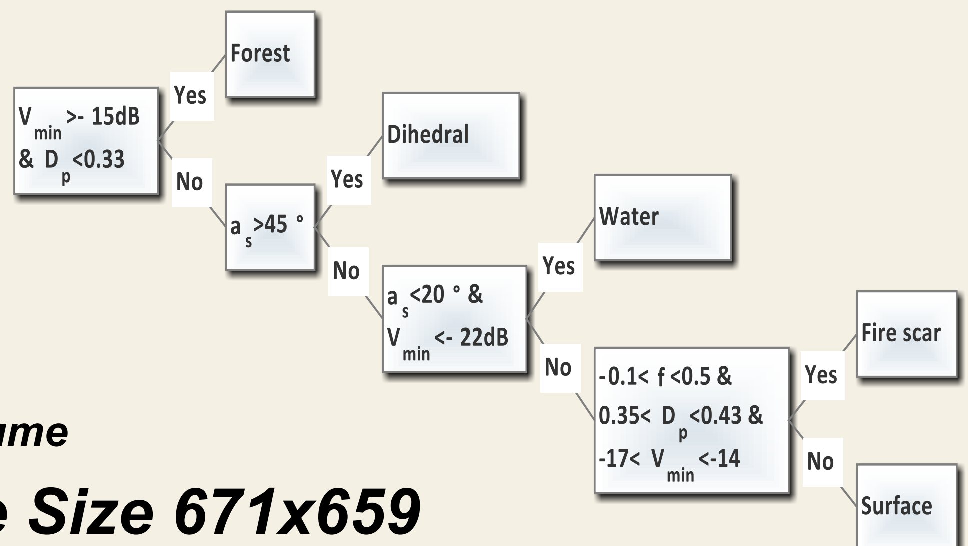
After trimming: Image Size 671x659

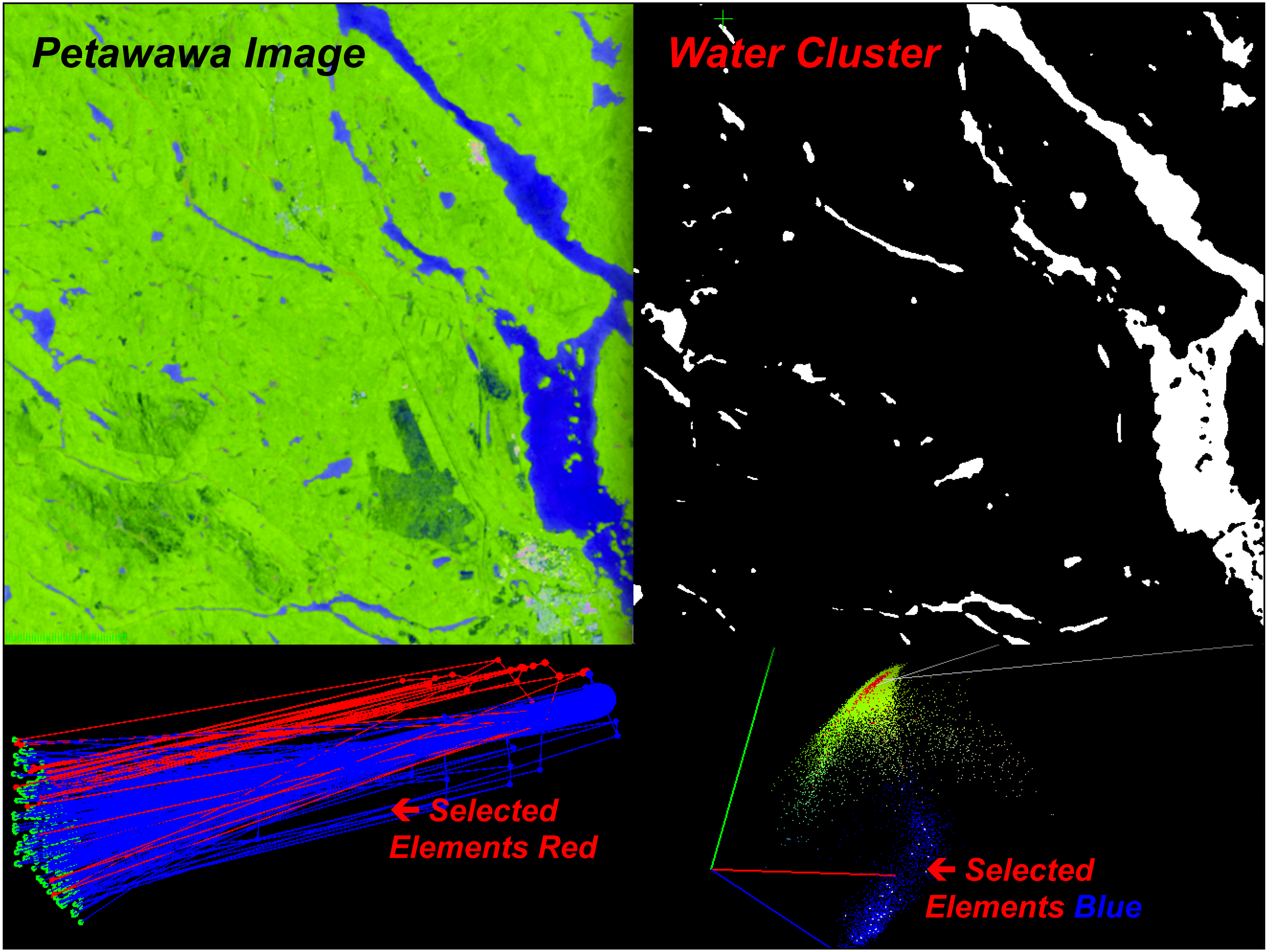
Comparison with rule based classifier:

Simulated Compact-pol parameters

IGARSS 2014: Mapping Forest Fire Scars with Simulated RCM Compact-pol Data

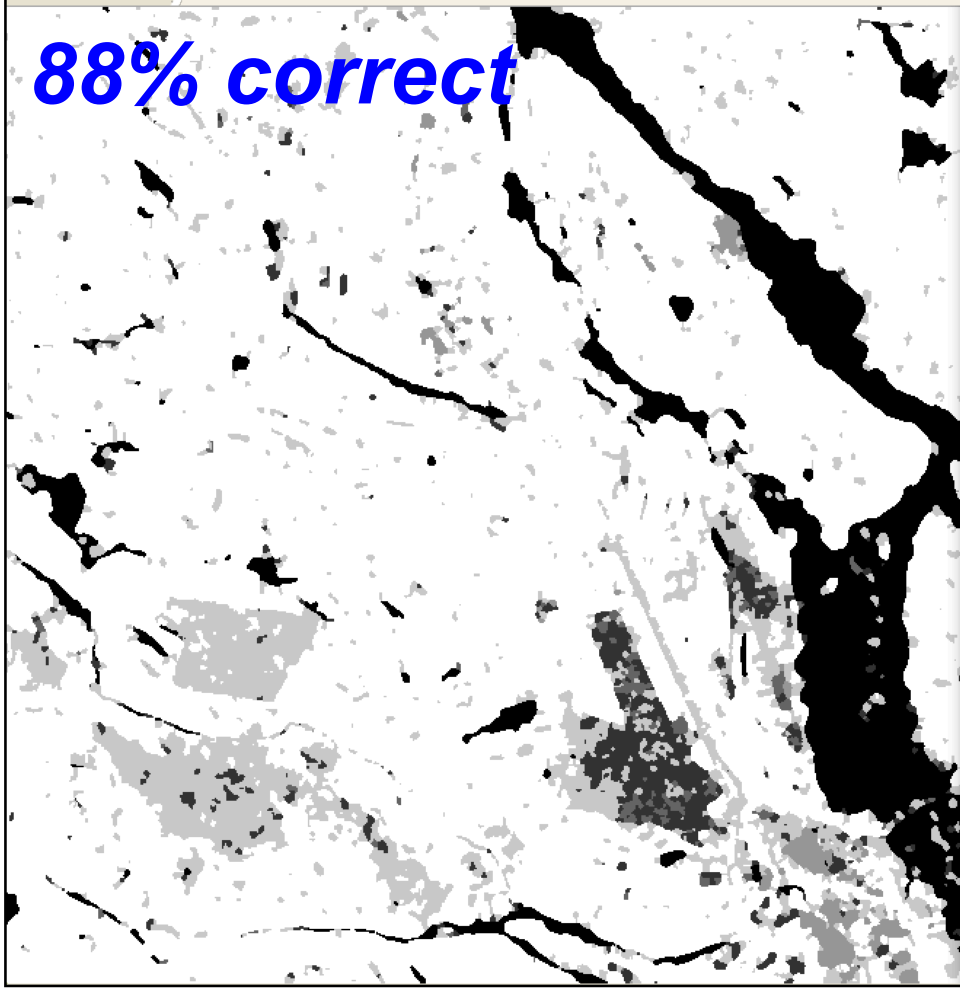
Hao Chen, David G. Goodenough, Shane Cloude, Philip Snead

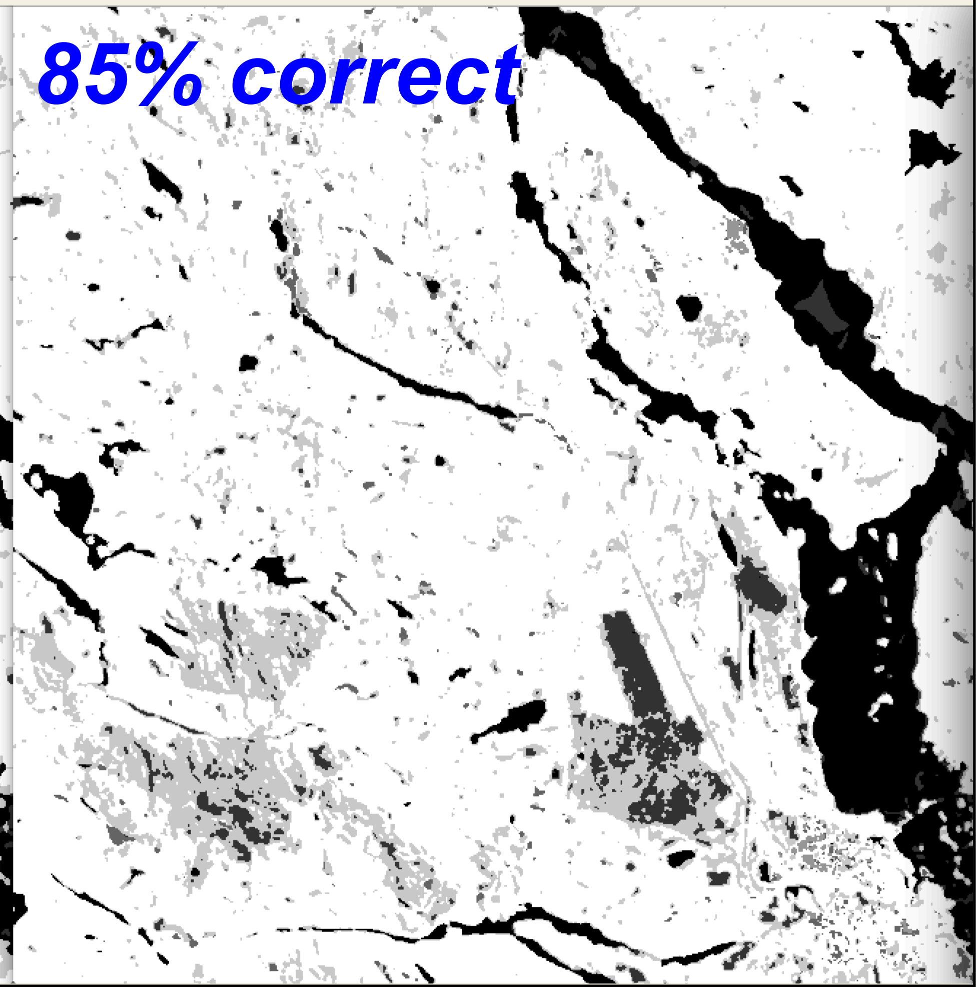



Petawawa Image
Water Cluster
Selected Elements Red
Selected Elements Blue

Classification maps:

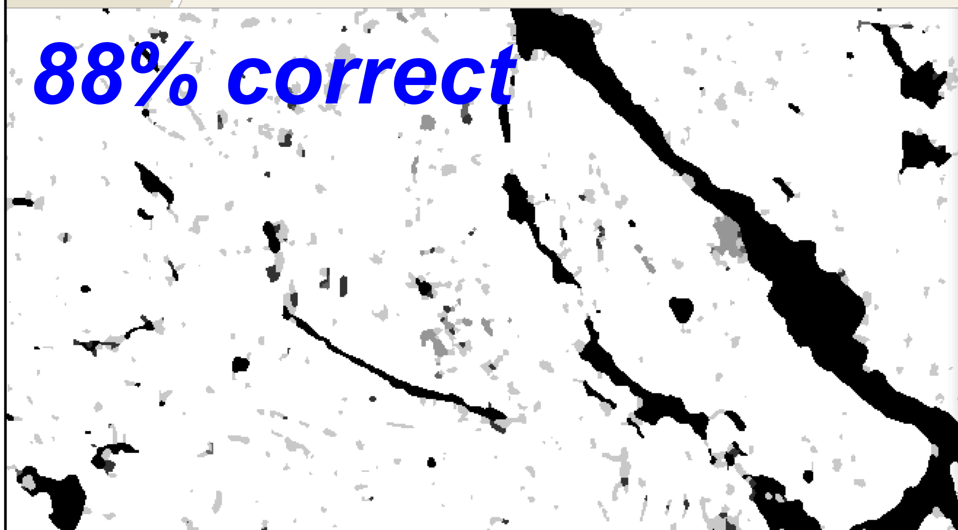
- Left) KGC-3 unsupervised method (with manual cluster assignment)
- Right) rule-based classifier on simulated compact parameters

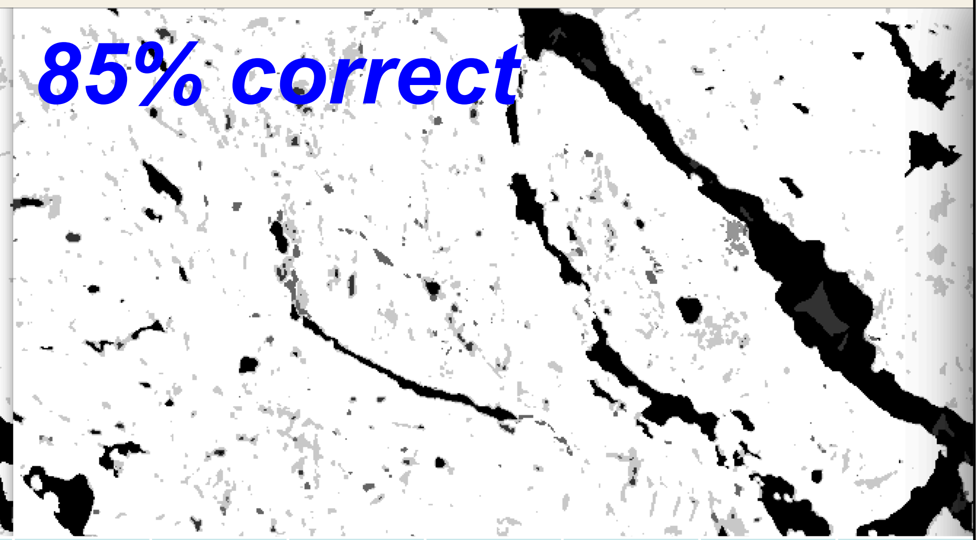




Classification maps:

- Left) KGC-3 unsupervised method (with manual cluster assignment)
- Right) rule-based classifier on simulated compact parameters





Water	Bare	Grass	Urban	Veg	Forest	
4068	0	0	0	0	0	100%
0	50	31	0	74	9	30%
0	272	169	2	391	247	16%
0	0	0	419	26	18	90%
0	376	146	0	1179	617	51%
0	0	10	1	311	13366	98%

Water	Bare	Grass	Urban	Veg	Forest	
4068	0	0	0	0	0	100%
0	41	0	0	109	14	25%
0	0	58	286	10	109	13%
0	0	58	286	10	109	62%
0	565	0	0	972	781	41%
0	21	71	0	434	13162	96%

Advantages: KGC classification

- **Data driven: KGC finds unusually shaped, poorly separated clusters that are difficult to find by existing approaches!!**
- **Nonparametric technique:** Avoid pitfalls of K-means & agglomeration (respectively assuming "roughly spherical" & "well-separated" clusters). For radar these may fail, especially when using decomposition parameters **(clusters behave strangely in decomposition parameter space..)**
- **Hierarchical information:** geometry approach: **merging reflects connections between clusters! Merge order has statistics interpretation (significance).**
- **More important features pop out first!!!!**
- **KGC: New nonparametric (local) segmentation step, reduces data volume, utilizes spatial information, and takes a stab at the filtering problem (less filtering required).**

Summary

- **KGC-3 – new hierarchical data driven method for unsupervised detection of arbitrarily shaped clusters** implemented and applied to Radarsat-2 time series data.
- Studies 2010-present ➔ consistent delineation of historical Fire Scars using Radarsat-2, ALOS-1 PALSAR, and other sensors
- Results show **effectiveness for forest applications, esp.. Fire Scar detection.**
- **Next steps:**
  - **Data fusion:** optical, radar combination
  - **Perform KGC-3 generic (metrical) clustering directly on time-series signatures**

Acknowledgements

Thanks for Financial support from:

- The Canadian Space Agency (CSA).
- Natural Resources Canada (NRCAN).
- The Department of Computer Science (Uvic).

Thanks for your attention.