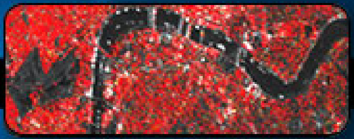
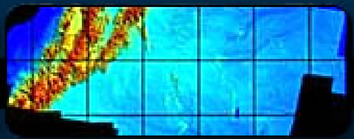
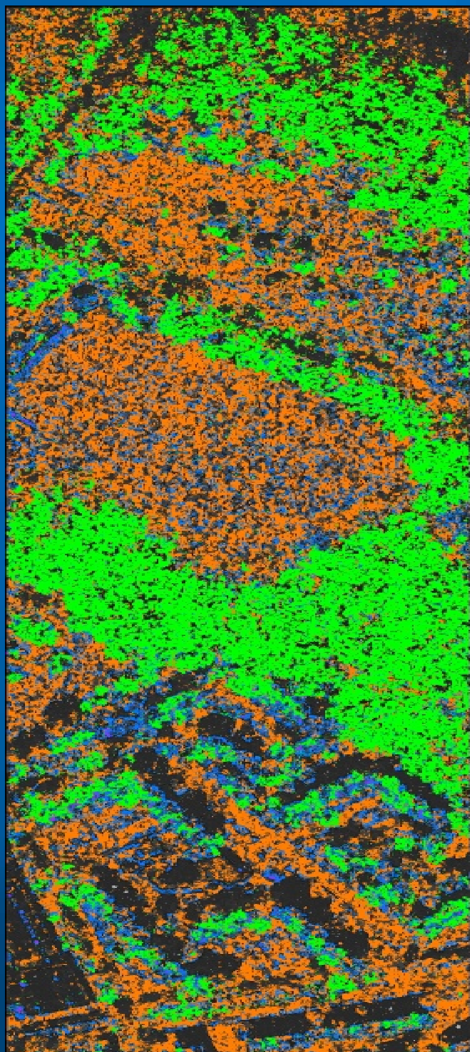




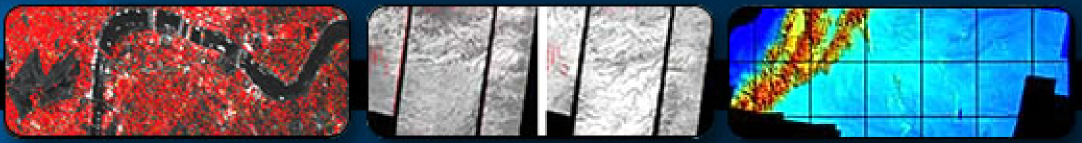
Forest region classification from SRTM data over the U.S.

Carolyn Johnston, Ph.D.
Vexcel Corporation
Boulder, CO
Carolyn.Johnston@vexcel.com
303-583-0273

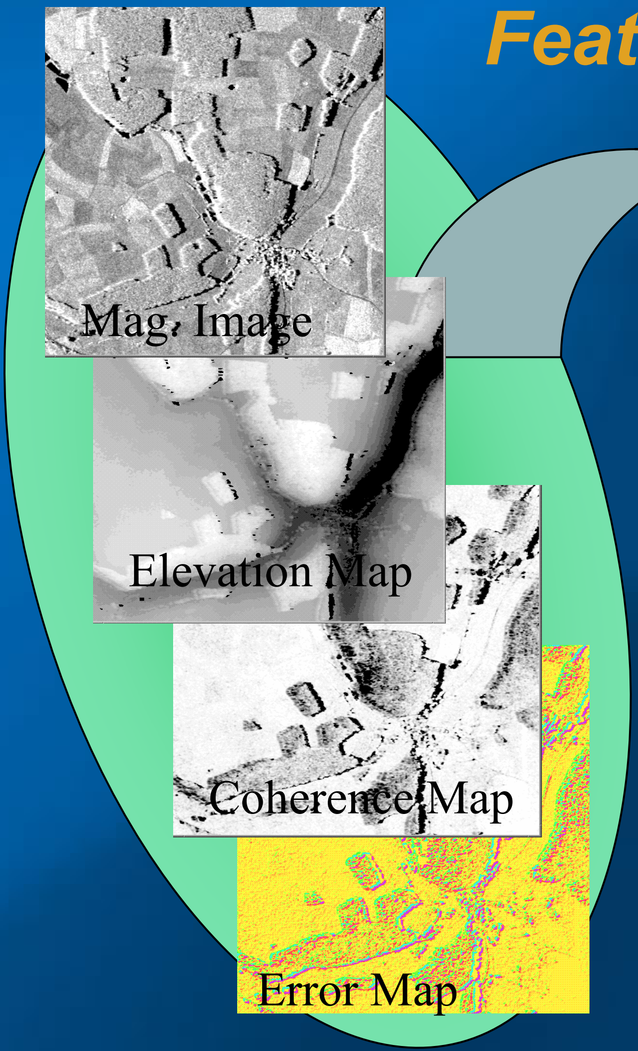


Background and Motivation

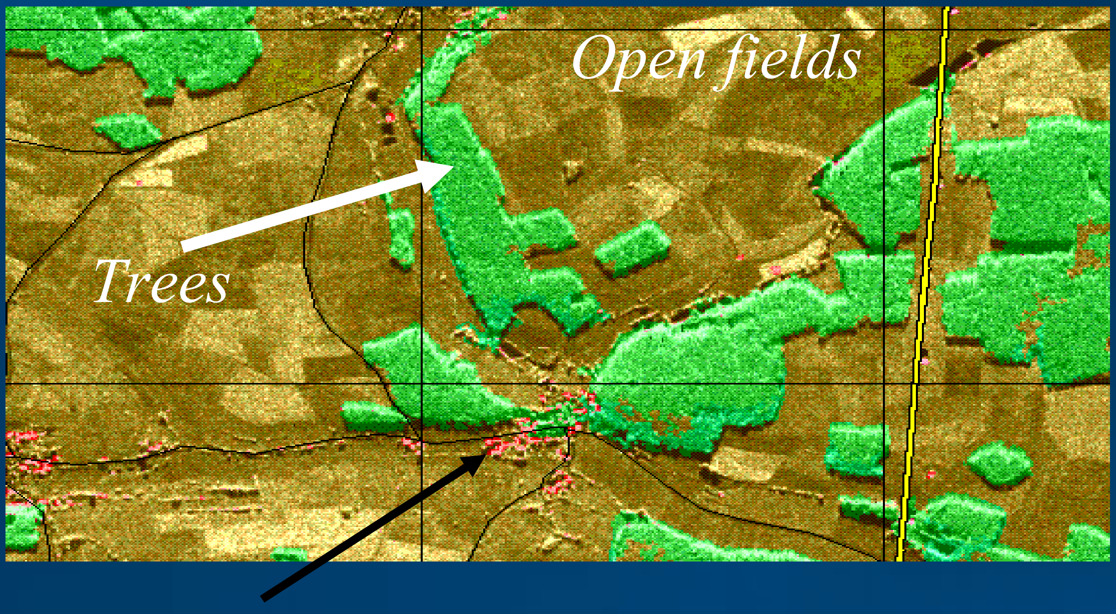
- 1997 Army TEC SBIR Phase 2
- extraction of land-use information from Star-3i (X-band IFSAR)
    - DEM
    - coherence
    - magnitude image
    - derived data
- classes extracted: bald-earth, foliage, low vegetation, urban



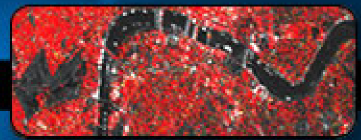
Feature Classification from Star-3i

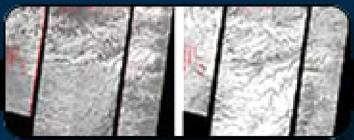


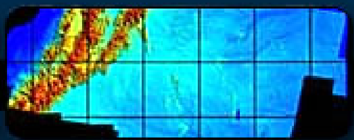
IFMAP Land-Use Classification



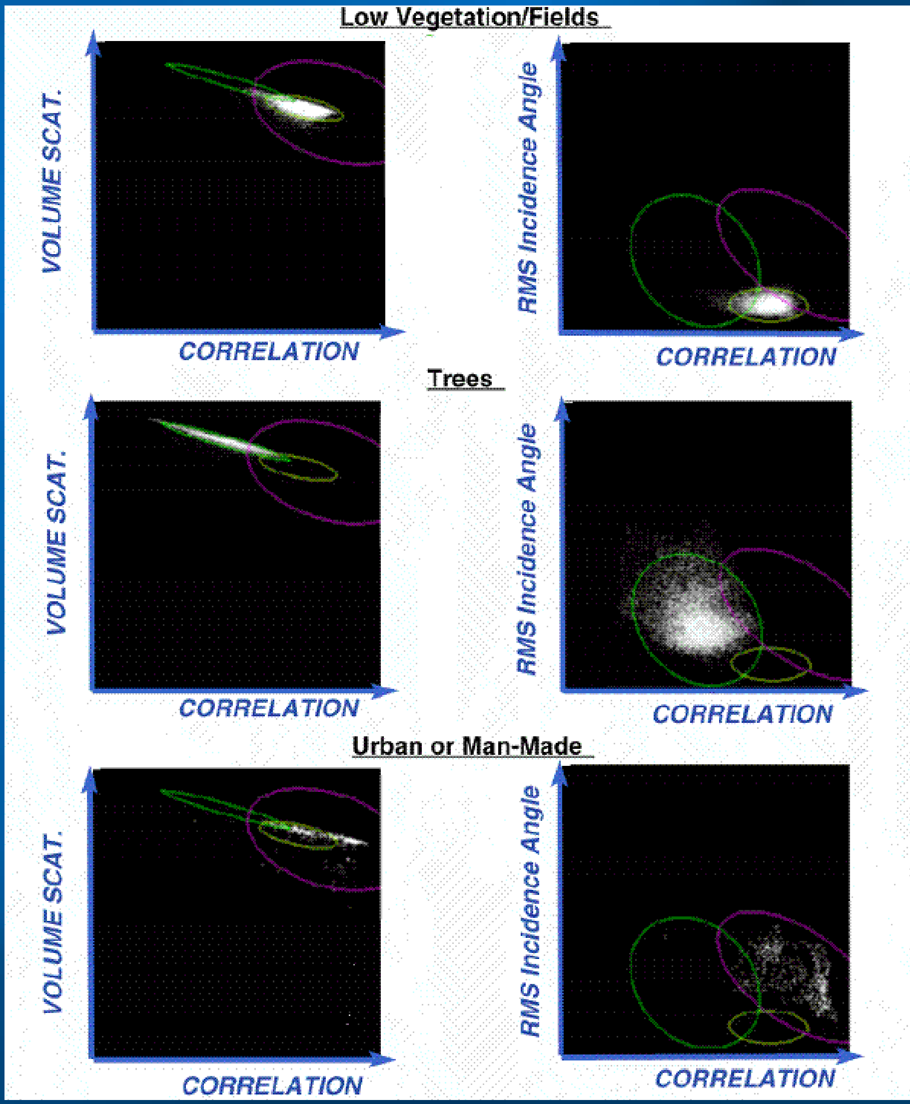
VEXCEL Innovation *in* Remote Sensing



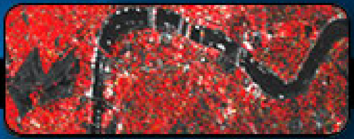


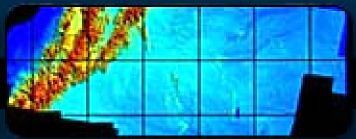


IFSARE discriminating features



This table (from IFMAP Phase 1, 1997) shows discriminating features for low vegetation, trees, and urban/man-made classes.



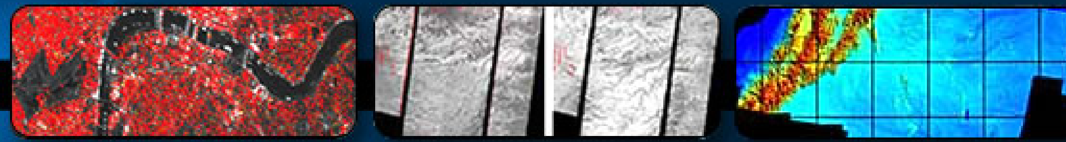


Can SRTM data be used for foliage classification?

- *Potential benefits*
  - worldwide 30m land-use classification for applications such as PTAN
  - IFMAP classification did fairly well without auxiliary data (such as LIDAR)
- *Related work*
  - 2004 JPL study (Rodriguez, Hensley, et. al.) analyzed *SRTM C-band penetration product*, found small-scale calibration ripples hampered classification
  - "Inferring Vegertation Canopy Information Using Combined Lidar and Radar Interferometry Data", Terrestrial Ecology Program Site Review, March 17, 2004.

Interferometric Correlation and THED

- *For single-pass IFSAR: correlation is* $\gamma = \gamma_g \gamma_t \gamma_v$
- *Precise relationship of SRTM correlation and THED product was not known*
  - THED is an aggregate Cramer-Rao estimate of error
  - baseline decorrelation must have been removed, result is derived from combination of thermal and volume decorrelation in multiple SAR images
- *Does vegetation have a detectable effect on THED values?*

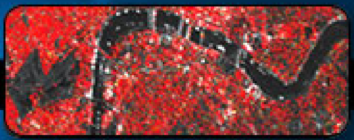


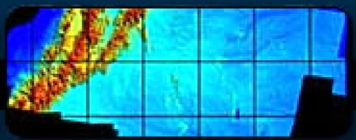
SRTM vs. Star-3i

Difference	Star-3i	SRTM
SNR	high (airborne)	lower (spaceborne)
resolution of products	2.5m	30m
phenomenology	X-band (volume decorrelation mainly in canopy)	C-band (volume decorrelation in twigs as well)
coherence data	available	data derived from multiple passes is available (THED)
sensor geometry (baseline, look angle)	available	not available – products are aggregate (baseline decorrelation is removed from THED)

-With lower SNR, volume decorrelation is difficult to isolate from thermal decorrelation

-Absolute THED values vary with latitude because more passes drive down noise





Preliminary sanity checks

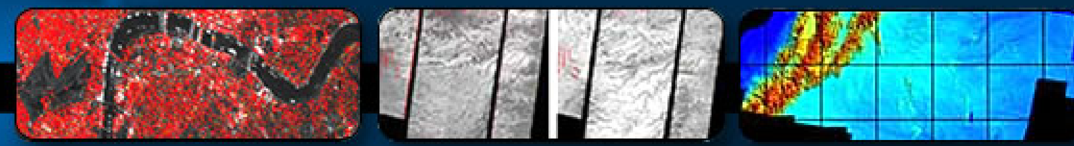
- *Quality maps were constructed for SRTM cells using NGA guidelines*
  - (P. Salamonowicz, NGA: "Use of SRTM Error Data for computing DTED Accuracy Values", July 26, 2002).
  - For datasets studied, the quality map differs from THED by a constant
- *National Land Cover Data (NLCD) was used as landcover truth*
- *Since volume decorrelation is a small effect relative to thermal decorrelation, we restricted consideration to posts that had a relatively low error to start with*

NLCD landcover classes

Class Name		Data Key
Water	Open Water	11
	Perennial Ice/Snow	12
Developed	Low Intensity Residential	21
	High Intensity Residential	22
	Commercial/Industrial/Transportation	23
Barren	Bare Rock/Sand/Clay	31
	Quarries/Strip Mines/Gravel Pits	32
	Transitional	33
Forested Upland	Deciduous Forest	41
	Evergreen Forest	42
	Mixed Forest	43
Shrub land	Shrub land	51
Non-natural Woody	Orchards/Vineyards/Other	61
Herbaceous Upland Natural /Semi-natural Vegetation	GrassLands/Herbaceous	71
Herbaceous Planted/Cultivated	Pasture/Hay	81
	Row Crops	82
	Small Grains	83
	Fallow	84
	Urban/Recreation Grass	85
Wetlands	Woody Wet Lands	91
	Emergent Herbaceous Wetlands	92

Foliated areas = union of all 'forested upland' classes

Bare earth = union of barren classes (+ shrubland and cultivated areas in Oregon)

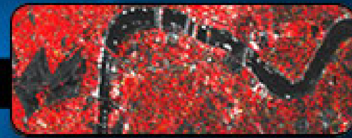
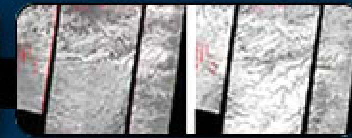
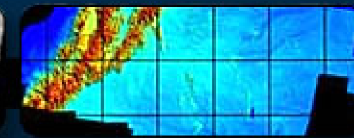


Quality map values – class means

- *Datasets studied were from Mississippi, Texas, Florida*
- *Only points corresponding to included classes were sampled*
- *We looked at the difference in mean of populations that were restricted to increasing thresholds*
- *Mean quality map values for the two classes were consistently only a centimeter apart*
  - The effect disappeared as higher quality map values were admitted to the population

Feature Vectors Study

- *We next considered full feature vectors from the Oregon and Mississippi SRTM datasets*
- *We performed a Principal Components Analysis, and a Linear Discriminant Analysis*
  - PCA: looks for a basis in which coefficients are decorrelated
  - LDA: looks for a basis that maximally separates class centroids, and minimizes variance around class centroids

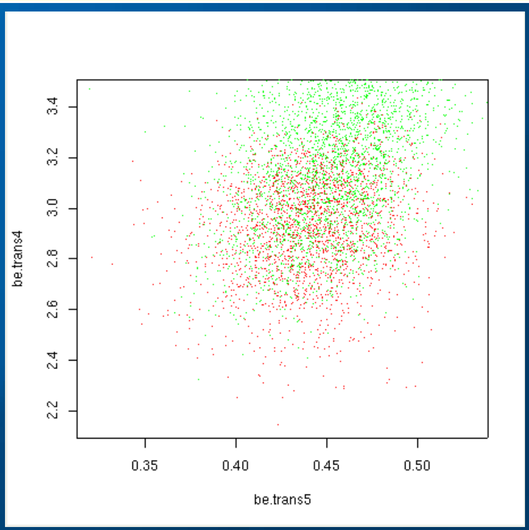




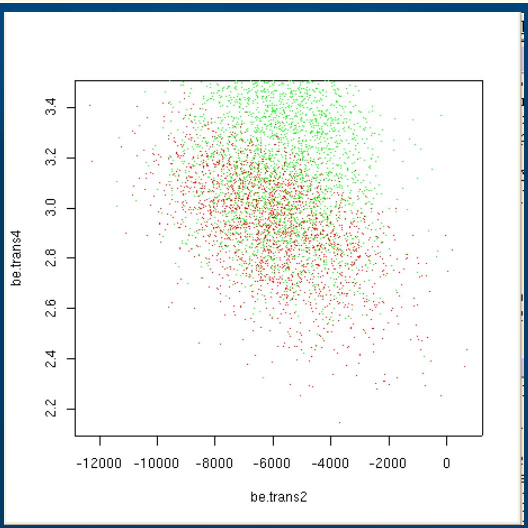
Feature vector fields

Field name	Land Type	THED	AOIM	DOIM	Slope magnitude	Slope aspect
Field description	0 = bald earth 1 = forest	Terrain height error, in decimeters	Magnitude of backscatter in ascending orthoimage	Magnitude of backscatter in descending orthoimage	Magnitude of elevation gradient vector at a point in the DEM, in meters per post	Direction of elevation gradient vector at a point in the DEM, in degrees: -180 through 180
Typical value	0	3.73	26133	27145	0.32	0

PCA analysis results, Oregon

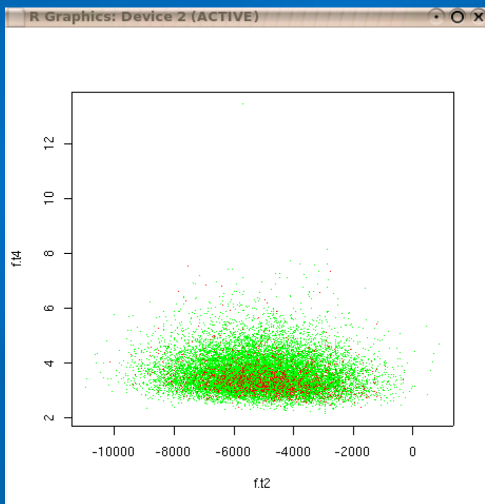
Principal component vector	THED	AOIM	DOIM	Slope magnitude	Slope aspect
PC 4	0.9978	5.90E-005	-2.04E-005	6.61E-002	-8.29E-004
PC 1	1.37E-005	-6.20E-001	-7.84E+000	-7.18E-006	-8.92E-003
PC 2	5.81E-005	-7.84E-001	6.20E+000	7.05E-006	-5.84E-004

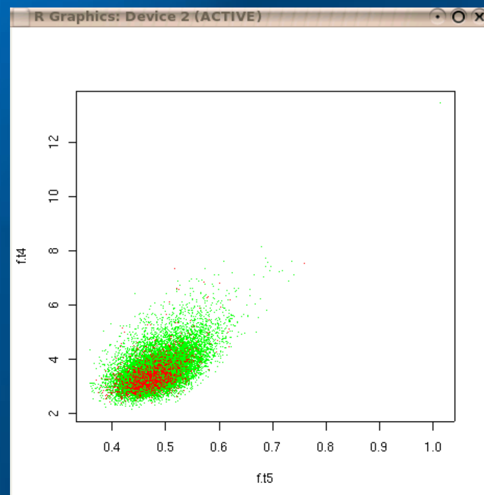




left: PC2, PC4
right: PC1, PC4
green: foliated
red: bare earth

PCA analysis, Mississippi

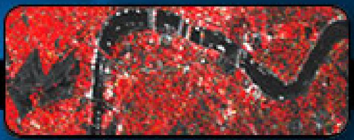


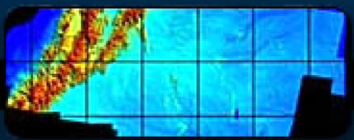


- *Scatterplots show projections of Ms. data onto Oregon PCs*
- *Mississippi PCs are no better*
- *PCA sometimes does not yield good discriminators*

Tried Fisher's LDA

LDA vector for data-set:	Thed	Aoim	Doim	Slope magnitude	Slope as-pect
MS	0.45962	0.00047	0.00026	5.62050	-0.00033
OR	0.38489	0.00025	0.00025	-0.26020	-0.00337



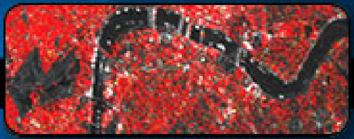


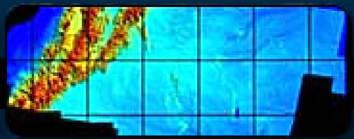
Fisher's LDA analysis

- *Ran LDA analysis on both Mississippi and Oregon datasets*
- *In both cases, found a single discriminator, similar to Oregon PC4: projection distributions were roughly Gaussian*
- *Second table shows means, variance of projections onto the discriminator*

LDA vector for dataset	THED	AOIM	DOIM	Slope magnitude	Slope aspect
MS	0.45960	0.00047	0.00026	5.62050	-0.00033
OR	0.38489	0.00025	0.00025	-0.26020	-0.00337

	Mean of bald earth	Mean of foliation	Variance of bald earth	Variance of foliation
MS	21.37	22.23	0.89	1.01
OR	14.51	15.20	~1.0	~1.0





Conclusions

- *The best linear discriminant in high and low terrain heavily weights THED and slope magnitude values; other discriminators from PCA can aid classification in some locations*
- *We tried K-means clustering in MS, which performed poorly with and without initialization to cluster centers*
- *Discrimination is better at higher latitudes*
- *Discrimination at lower latitudes is better if analyzed posts are restricted to low THED values*