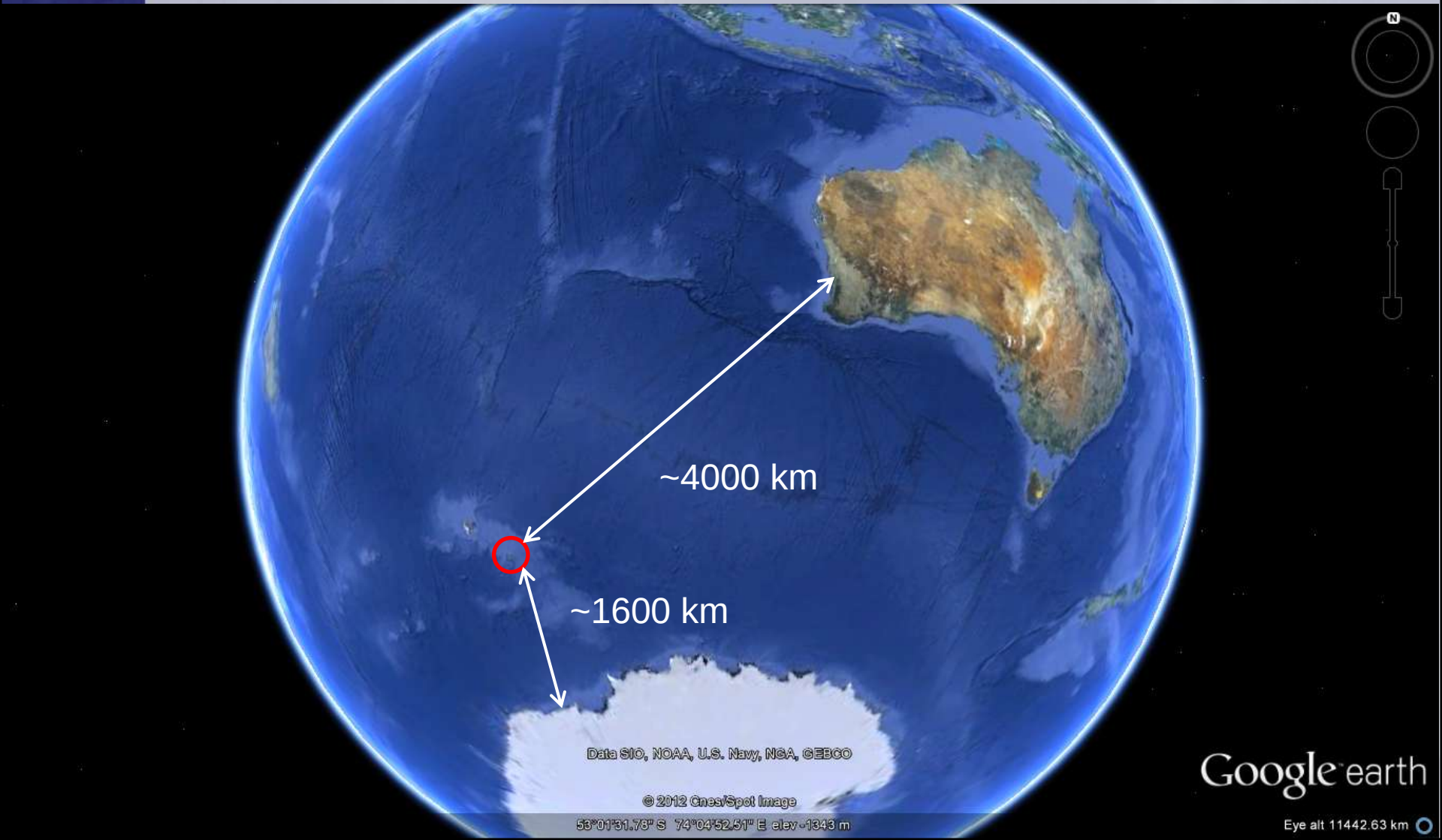
Geographic object-based image analysis (GEOBIA) of Worldview-2 imagery of Heard Island for sub-Antarctic vegetation mapping

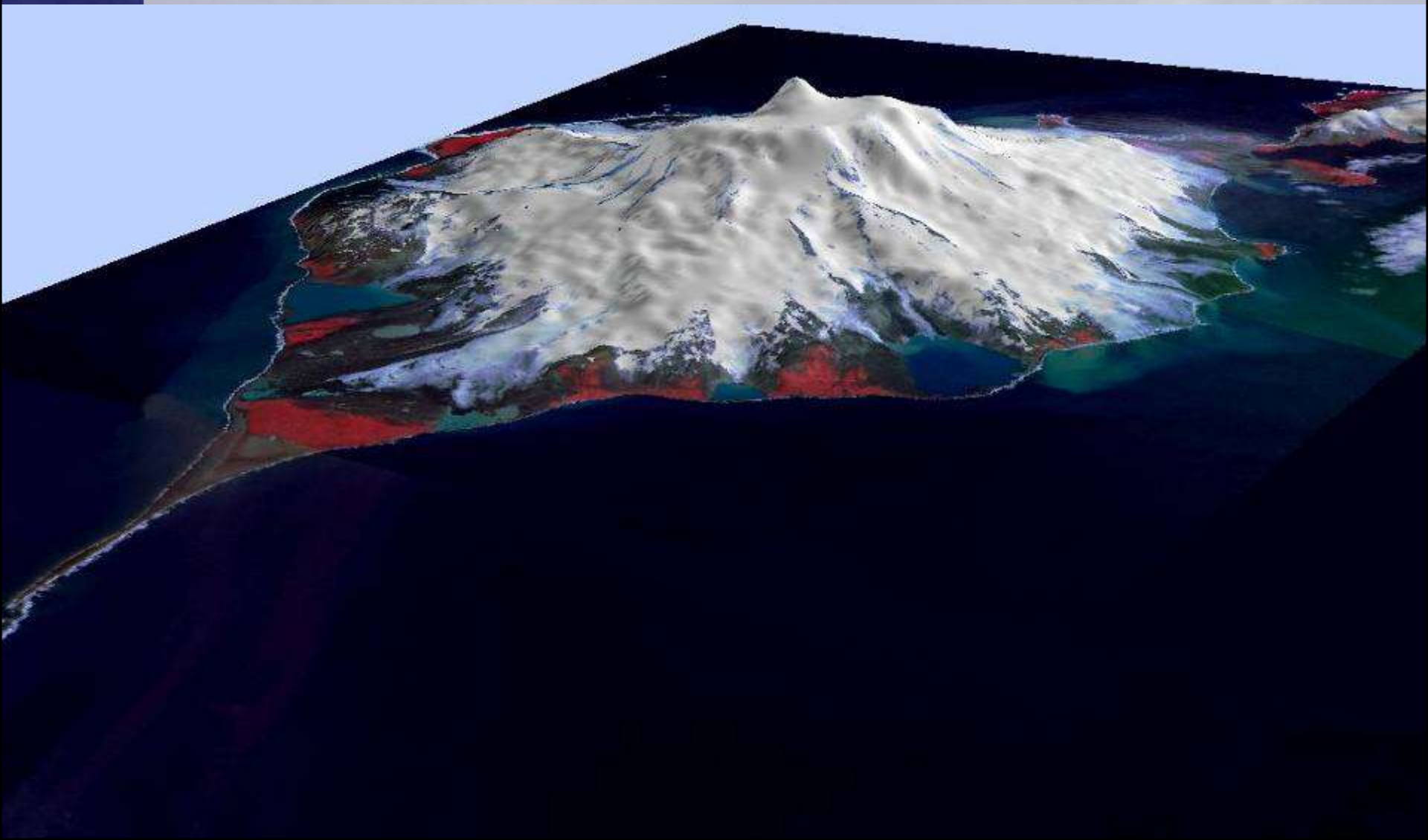
Arko Lucieer
School of Geography and Environmental Studies,
University of Tasmania

The 12th International Circumpolar Remote Sensing Symposium
Levi, Finland, 14 – 18 May 2012

Heard Island: where?

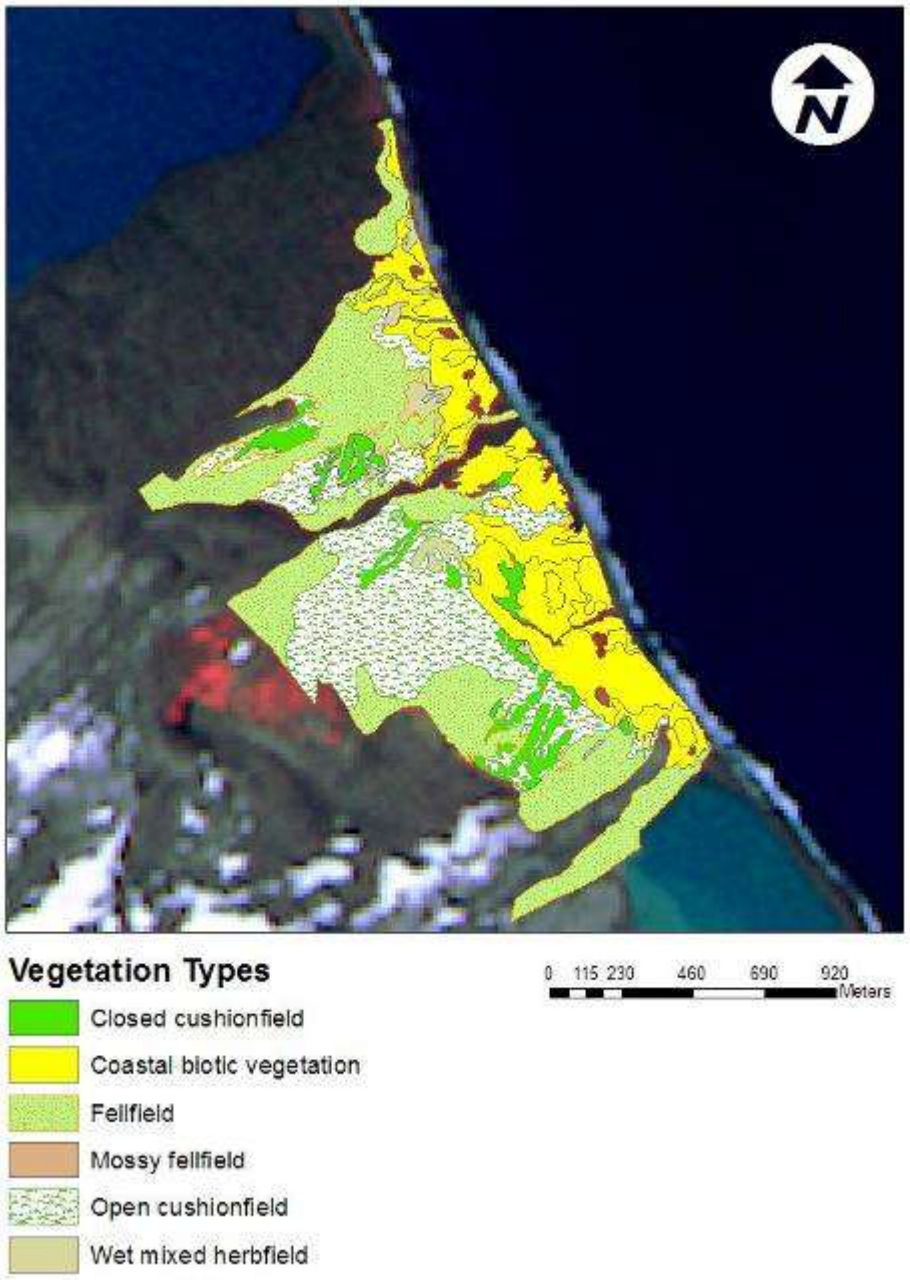


SPOT 1988



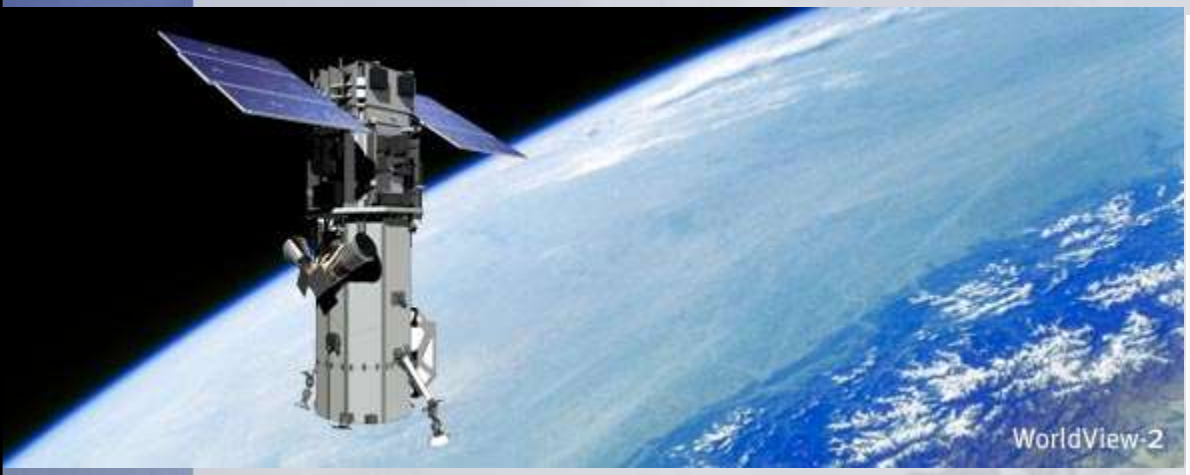

Source: Jenny Scott, UTas

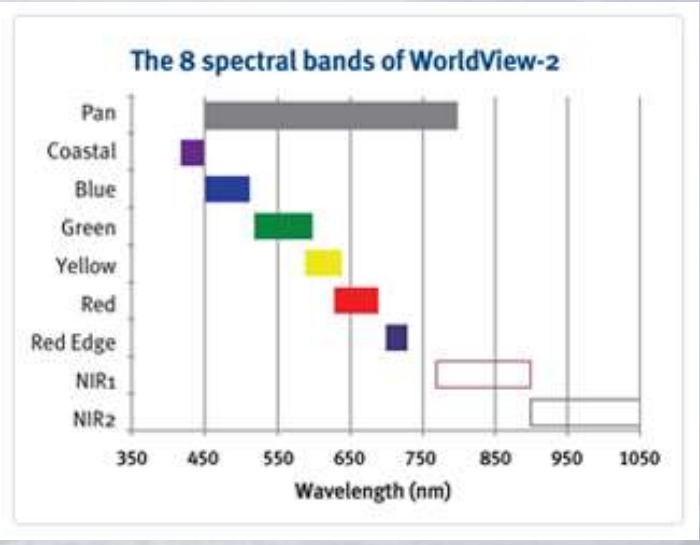
Vegetation M



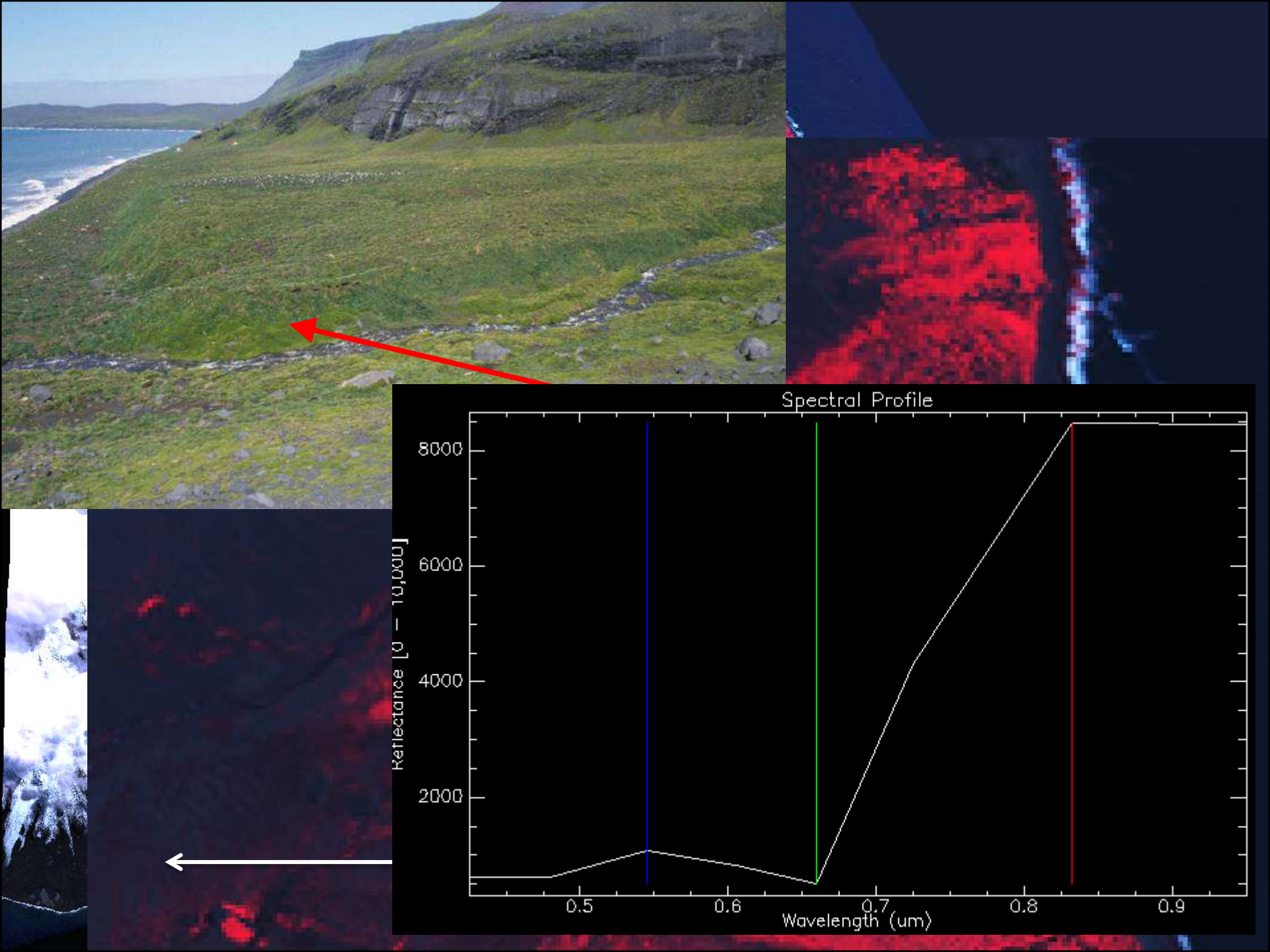
WorldView-2

- 1 panchromatic band: 0.46 m
- **8** multispectral bands: 1.84 m
- 1- 4 day revisit cycle
- 4.6 – 10.7 m positional accuracy
- Launched 8 October 2009

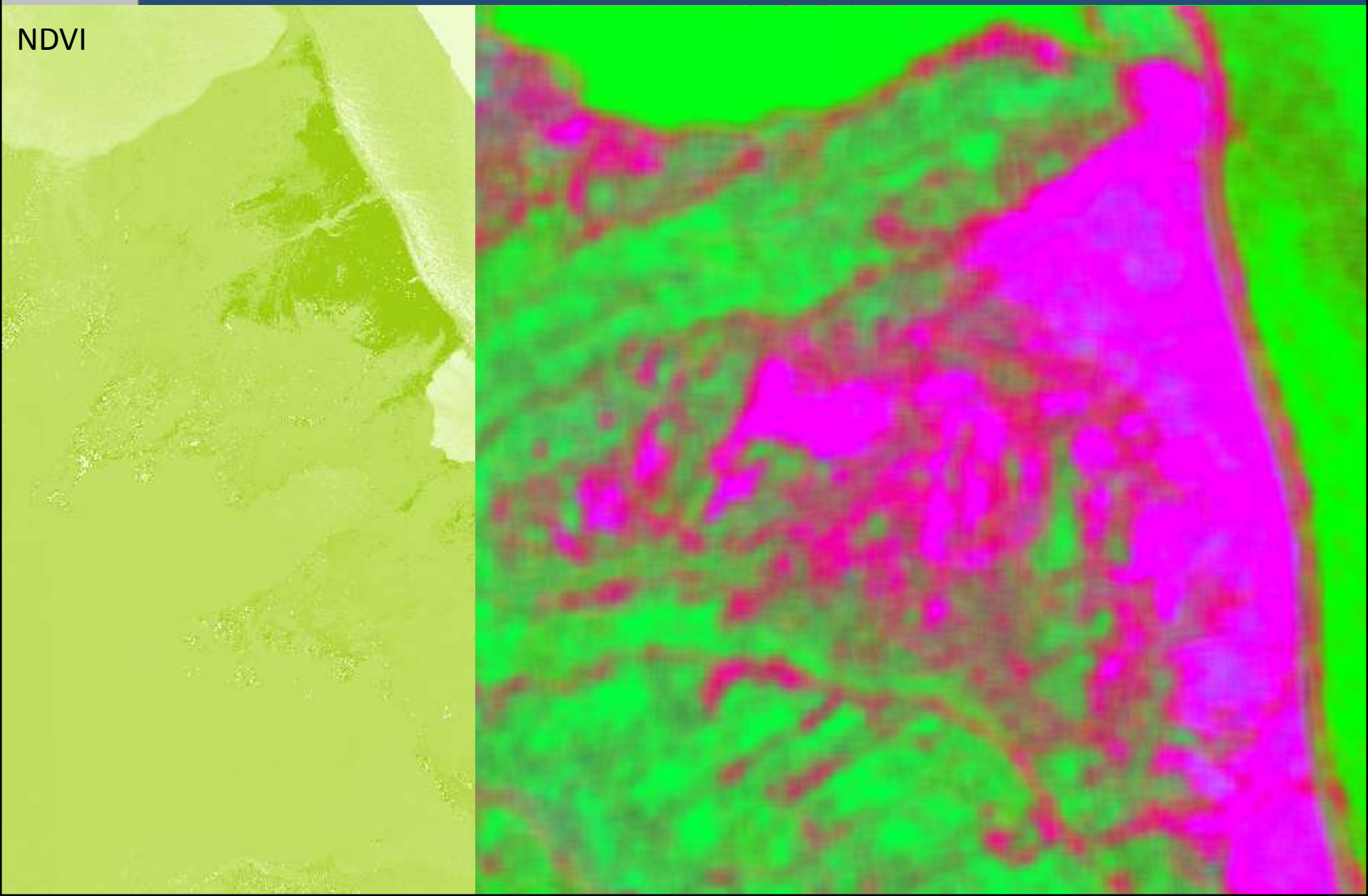




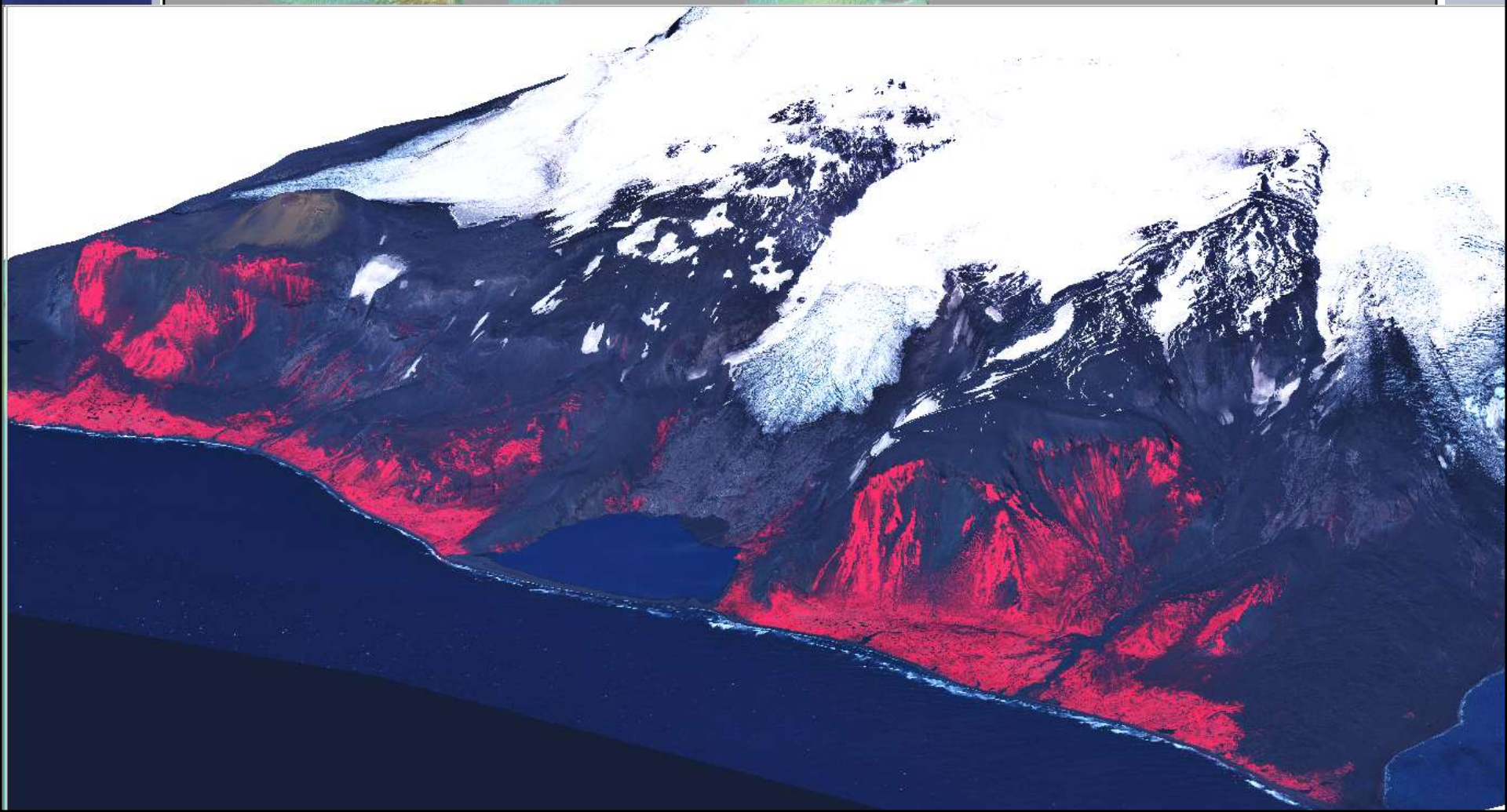
http://worldview2.digitalglobe.com


Spectral Profile
8000
6000
4000
2000
Reflectance [0 – 10,000]
0.5
0.6
0.7
0.8
0.9
Wavelength (um)

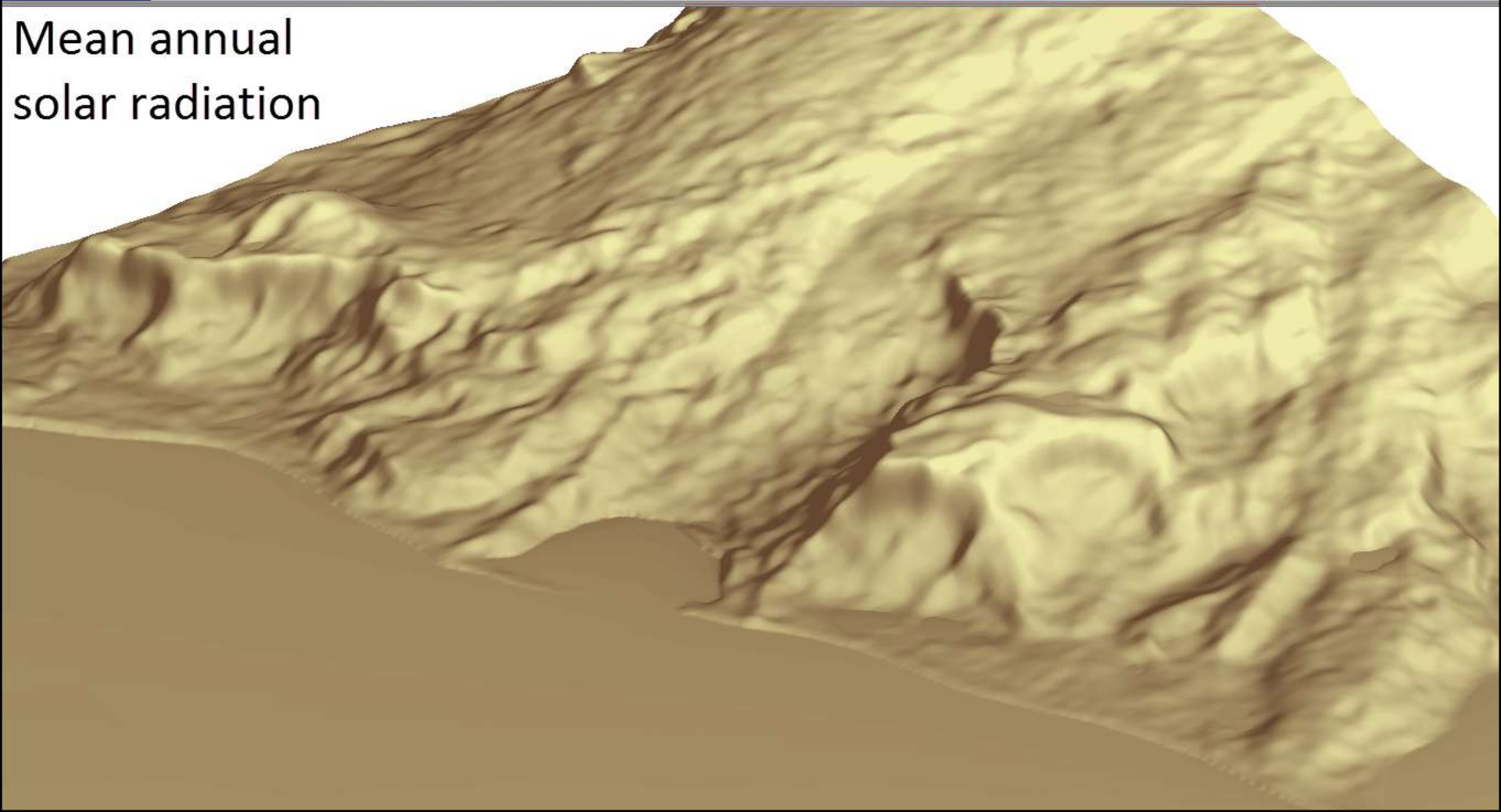
WV2 image derivatives



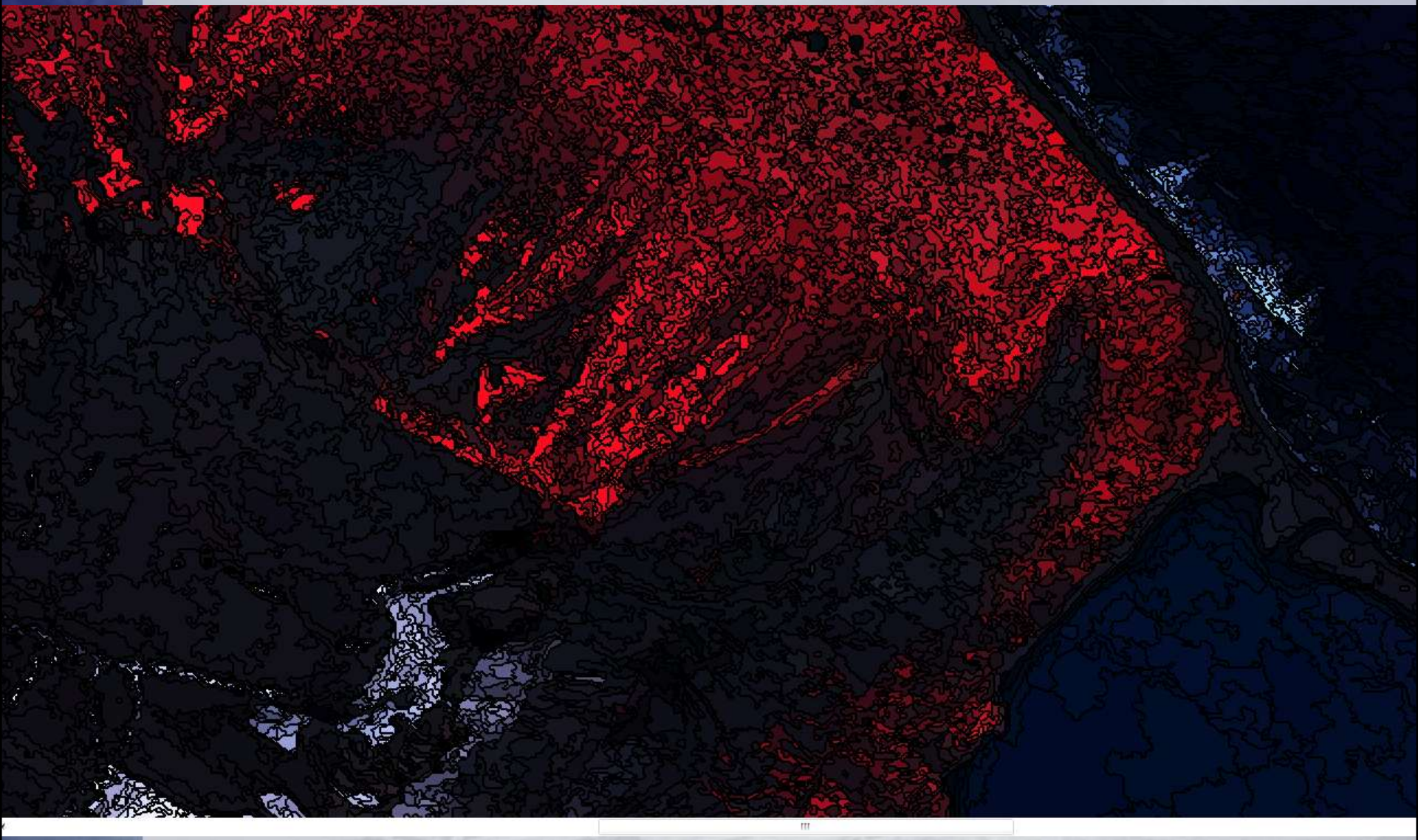
TerraSAR-X Digital Elevation Model

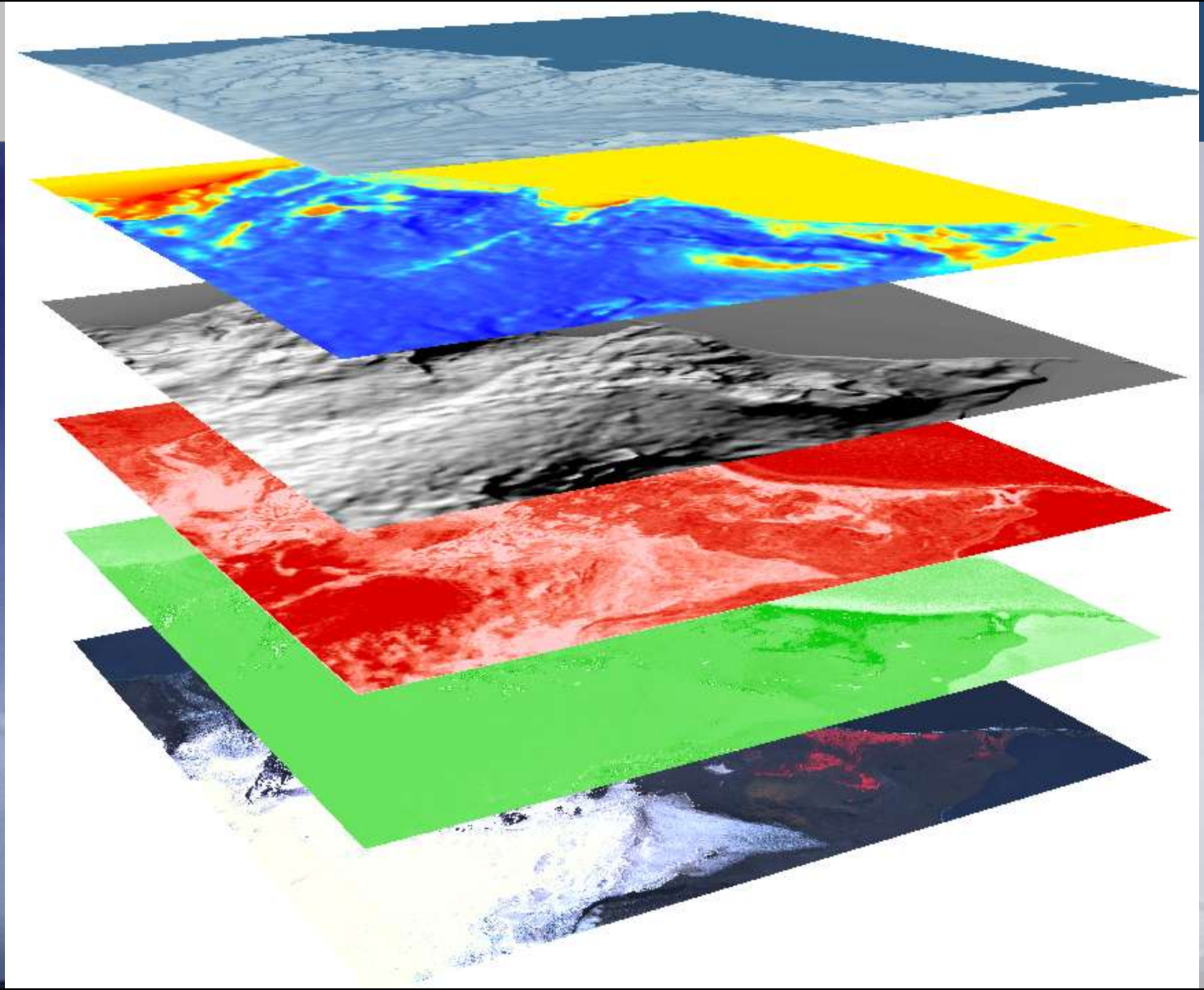


DEM derivatives

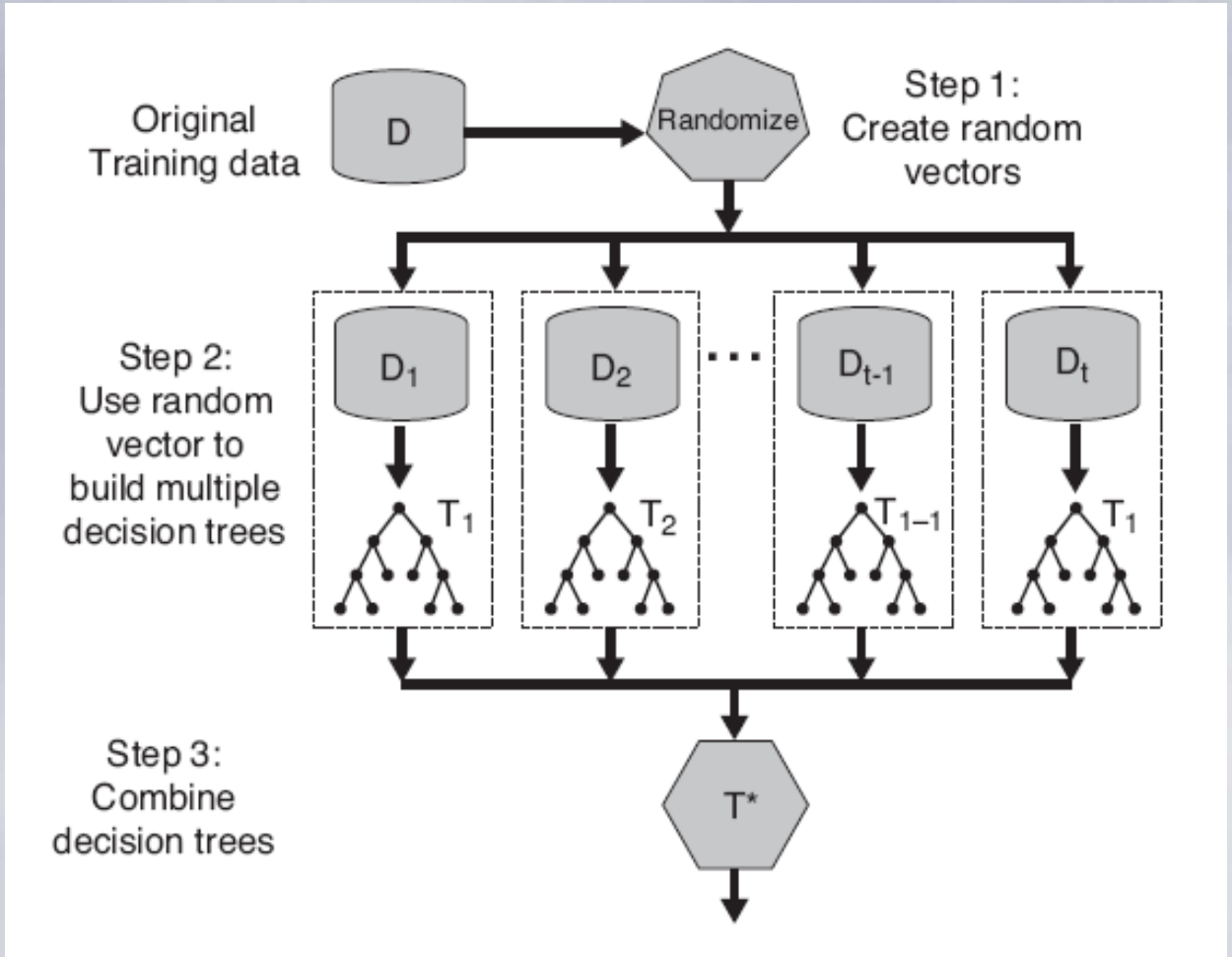


Geographic object based image analysis





Random Forest classifier



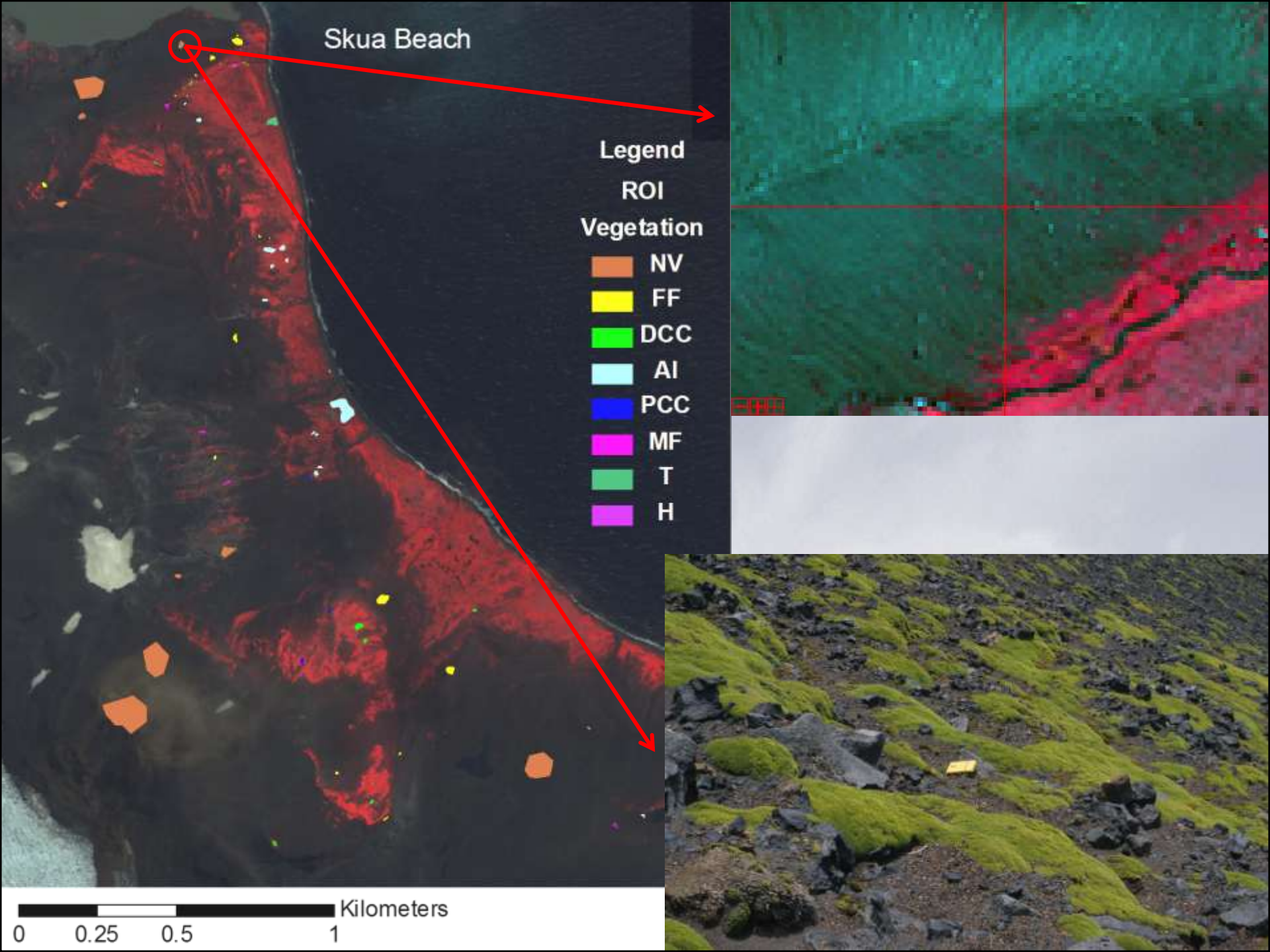
Breiman, Leo (2001). "Random Forests". *Machine Learning* **45** (1): 5–32. doi:10.1023/A:1010933404324.

Random Forest algorithm

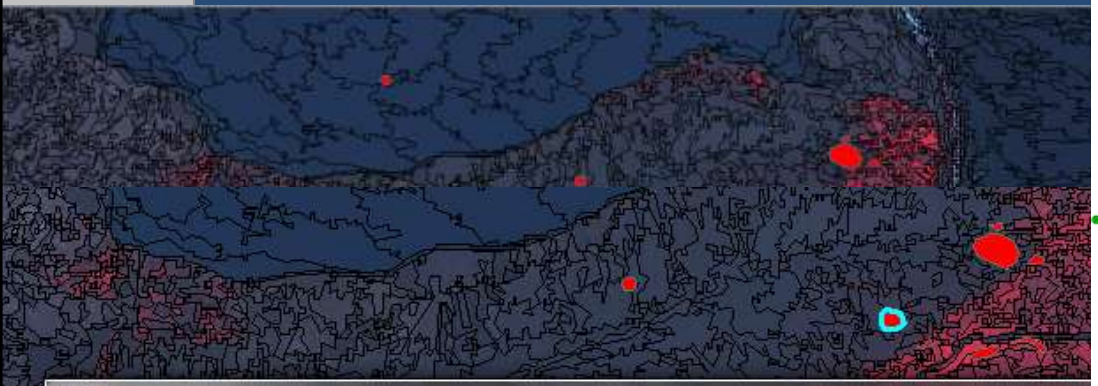
- Ensemble decision tree classifier based on multiple decision trees (CART, ID3, C4.5)
- Introduce two sources of randomness: “Bagging” and “Random input vectors”
  - Each tree is grown using a *bootstrap* sample of training data (random pixels 2/3 training, 1/3 testing, out-of-bag OOB error)
  - At each node, best split is chosen from random sample of m_{try} variables instead of all variables (reduce computational complexity and reduce correlation between trees)
- Make predictions according to majority vote of the set of trees (mode, probability)
- Importance of variable *m* can be estimated by randomly permuting all the values of the m^{th} variable in the out of bag samples for each classifier

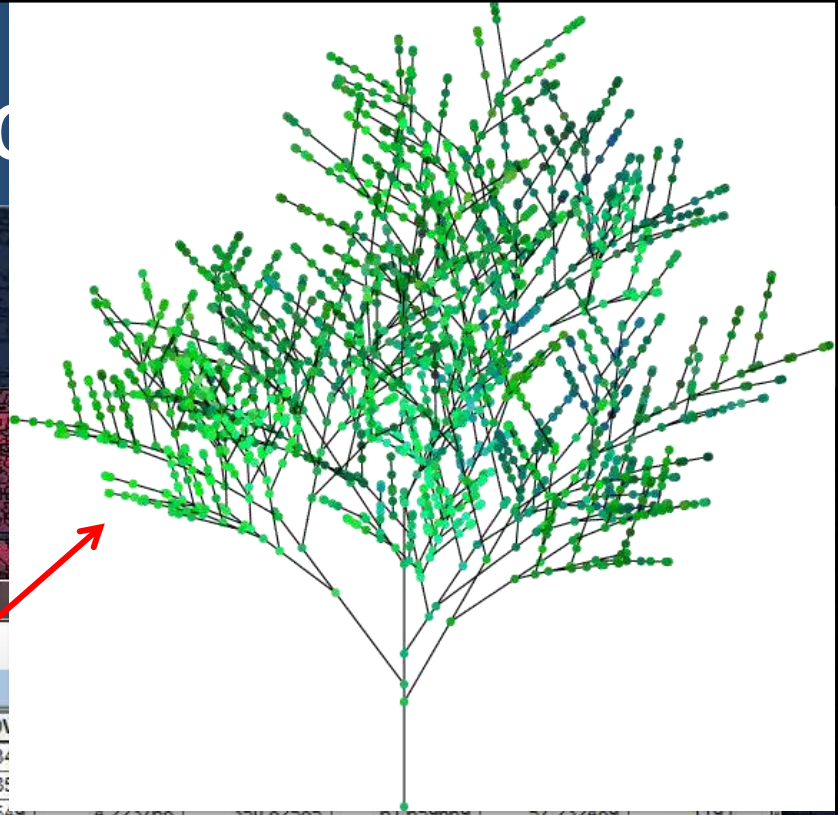
Random Forest advantages

- It is an ensemble classifier, i.e. more robust than a single classifier.
- It is a non-parametric classifier, i.e. no (incorrect) assumptions about statistical distribution of classes
- It handles a very large number of input variables
- It estimates the *importance of variables,* i.e. contribution of bands to the classification (for each class and for the classification as a whole)
- It calculates class probabilities for each sample, providing information on classification uncertainty
- It only has two parameters (number of trees and number of variables per tree) and it is insensitive to their values (compared to SVMs with several parameter)
- It outperforms most if not all state-of-the art classifiers


Skua Beach
Legend
ROI
Vegetation
NV
FF
DCC
AI
PCC
MF
T
H
Kilometers
0
0.25
0.5
1

Training random fo





Table

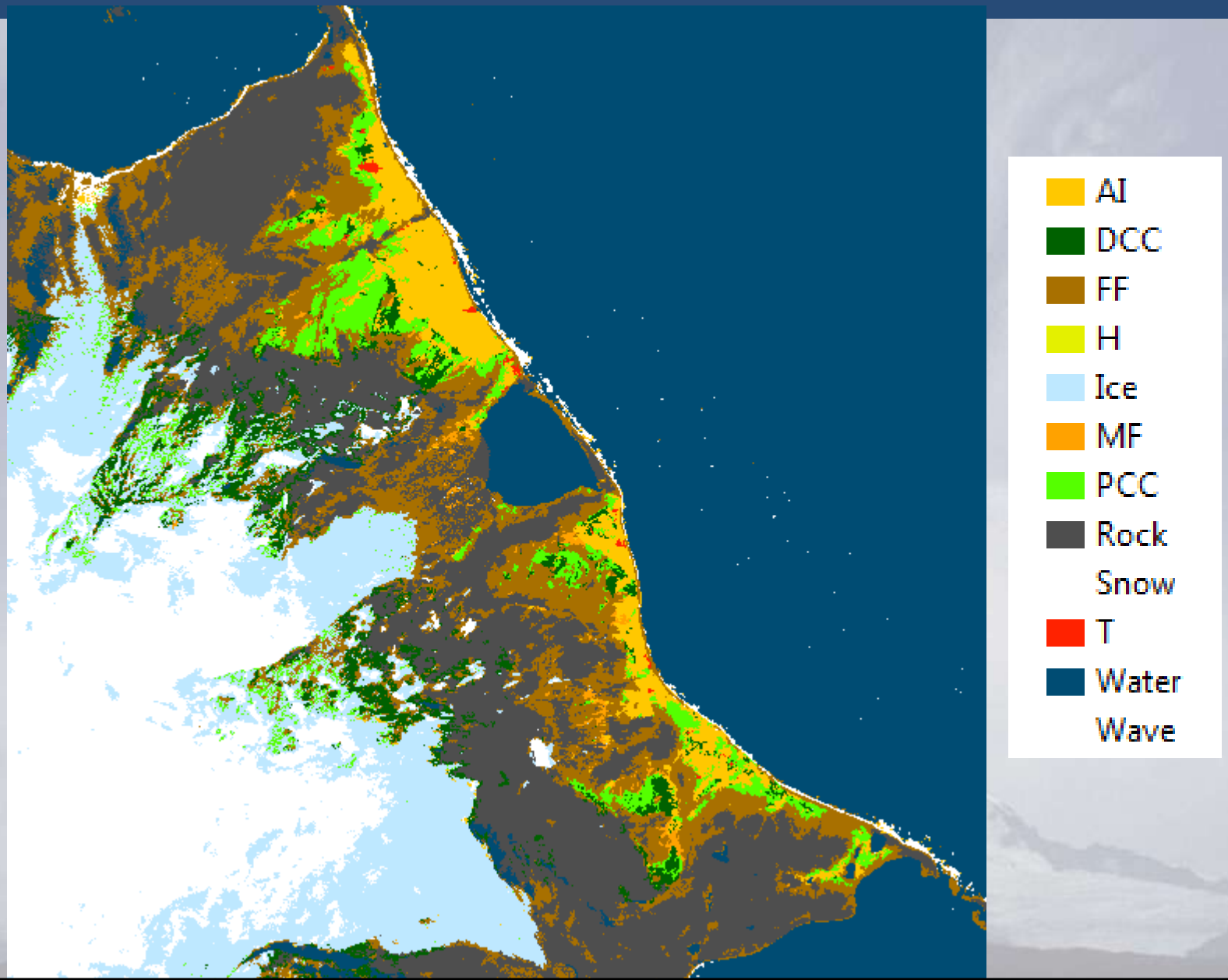
Objects_WV2_specterstack

Mean_WISAG	Mean_WV2_3	Mean_WV2_4	Mean_WV2_5	Mean_WV2_6	Mean_WV2_7	Mean_NDV					
10.67911	626.518987	598.772152	732.481013	769.860759	759.177215	0.1173					
8.590388	741.806452	695.5	1035.854839	1198.685484	1171.596774	0.2538					
9.004178	724.204545	664.25	1252.409091	1602.136364	1576.375	0.399549					
13.387769	653.135593	624.661017	631.364407	585.771186	540.508475	-0.03217	1.147136	58.54296	61.610204	61.695257	196.
8.286549	677.515152	618.818182	1023.030303	1198.848485	1167.727273	0.309073	4.646999	234.089483	39.924516	38.237718	72.
12.615414	569.041971	546.518248	541.919708	497.425182	483.20438	-0.047087	1.95292	21.000084	20.995516	23.580686	483
8.41762	485.662252	398.291391	531.655629	446.880795	336.490066	0.036008	5.967167	145.295175	48.387452	43.24484	159
8.782277	564.916667	486.416667	746.708333	702.583333	680.708333	0.173124	5.422416	164.521001	71.065062	64.733092	64.
8.230524	773.375	721.25	1621.125	2196.25	2257.875	0.504859	6.888322	130.373262	24.671593	50.982227	24.
11.591211	544.937736	523.703774	520.067925	476.833962	462.867925	-0.04715	2.805351	29.006095	24.816705	24.827385	700.
8.26834	732.615385	674.769231	1304.769231	1791.076923	1759.307692	0.439007	7.017577	409.197237	53.619964	38.590246	28.
7.188455	1102.5	1102.5	1208.25	1197.75	1193.25	0.0396	8.053459	117.680447	63.48425	64.101482	
7.831995	770.64	703.02	1481.88	1998.91	2001.06	0.458604	6.34496	549.435803	55.84317	60.737303	186.
7.999056	680.934307	651.208029	872.452555	960.427007	925.416058	0.176179	7.326483	251.227592	73.604118	71.341506	407.
8.16789	747.590164	698.786885	1130.47541	1321.065574	1328.491803	0.296241	7.304586	277.60072	50.51444	51.101603	138.
8.010256	582.362007	561.795699	561.946237	517.189964	501.215054	-0.042874	5.072204	53.994787	35.841915	34.392407	334.
8.282082	799.25	732.083333	1527.25	2029.25	2024.5	0.463539	7.118183	325.873474	59.059939	52.0888	30.
10.750884	857.2	804.4	858	836.6	815.4	0.017327	4.476801	164.933441	134.884543	120.876797	18.
8.006118	854.6	802.5	1435.4	1984.6	2028.9	0.4138	7.625832	400.308431	105.146802	63.459909	28

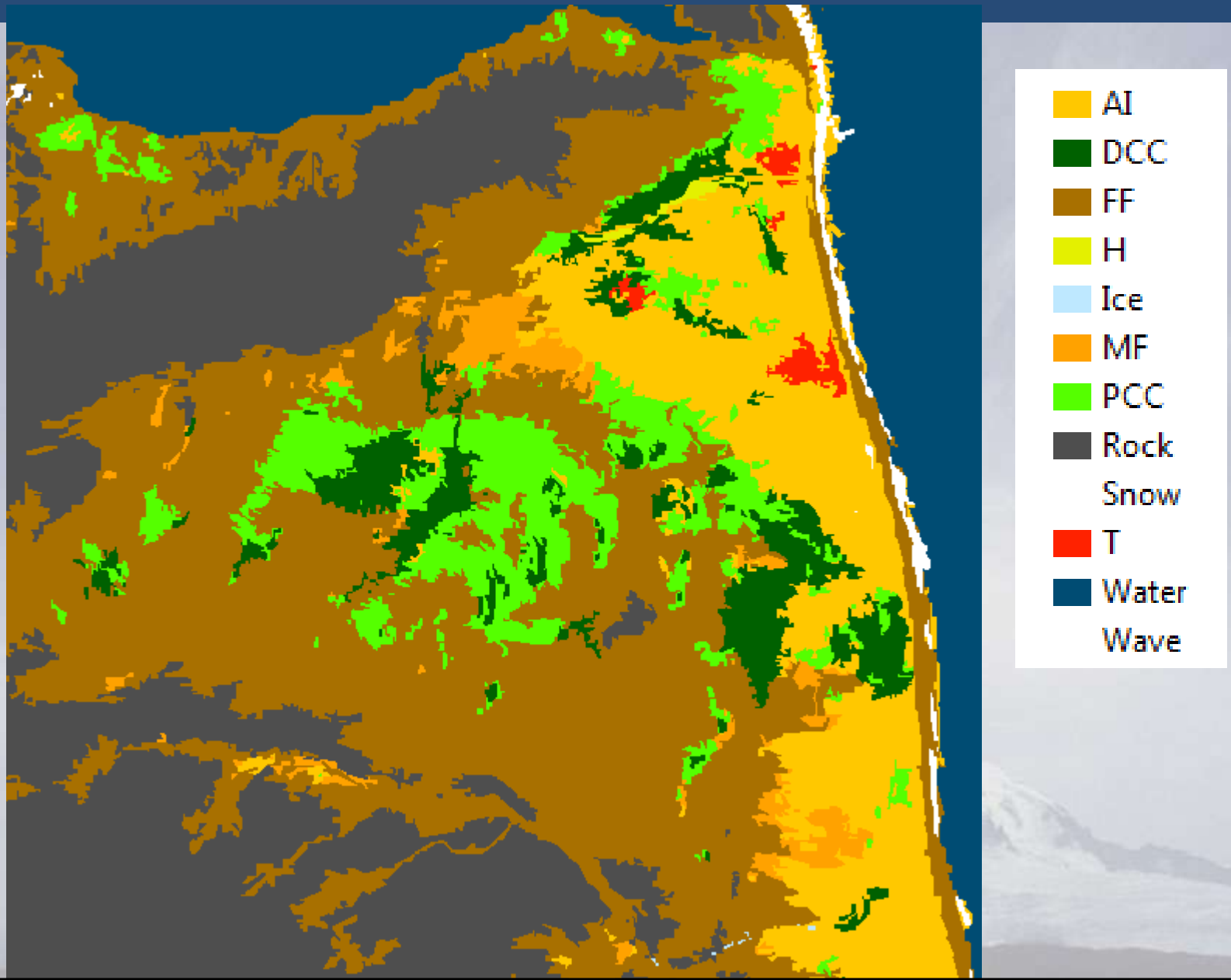
1 (1 out of 198536 Selected)

Objects_WV2_specterstack

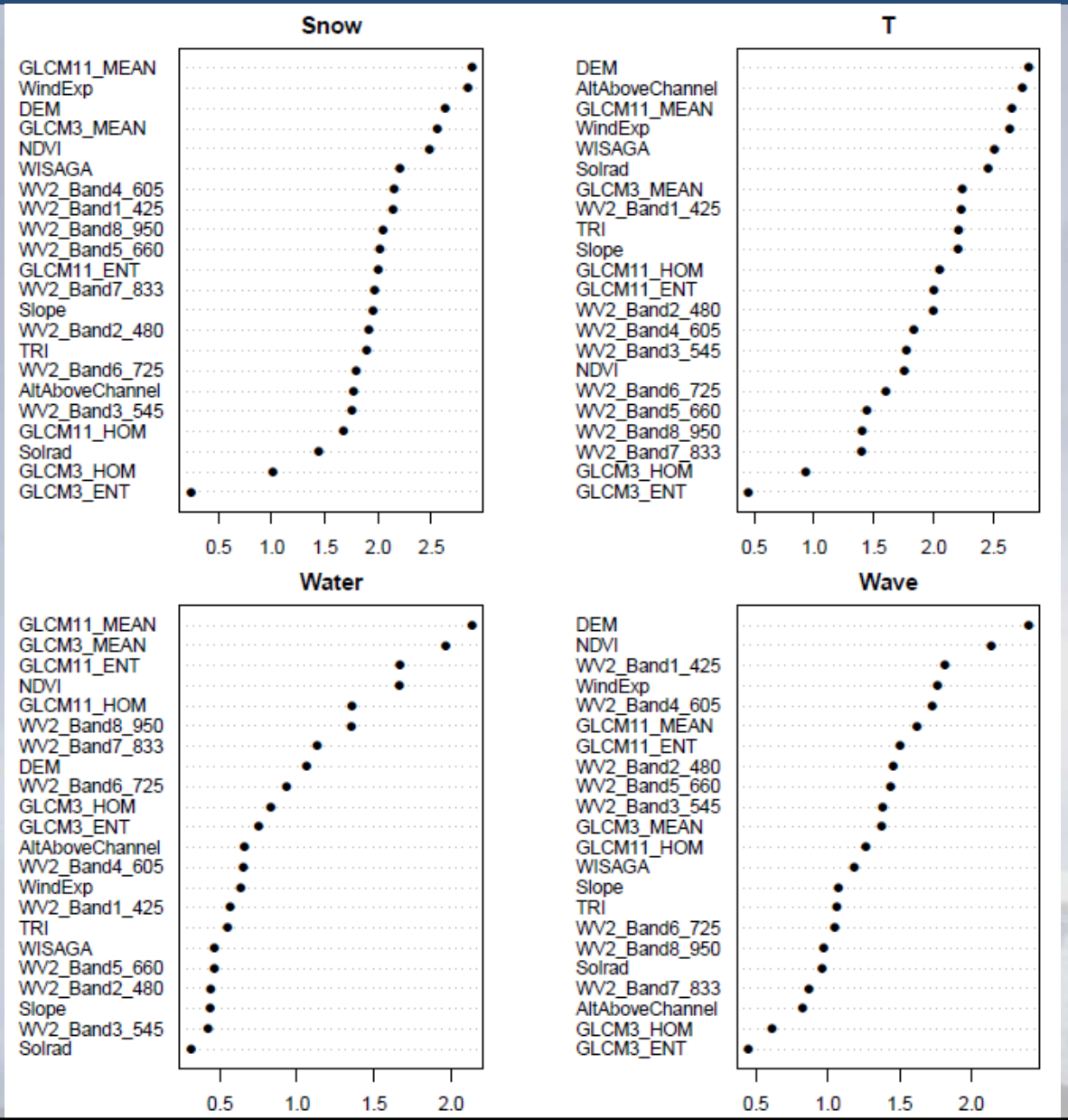
RF classification results



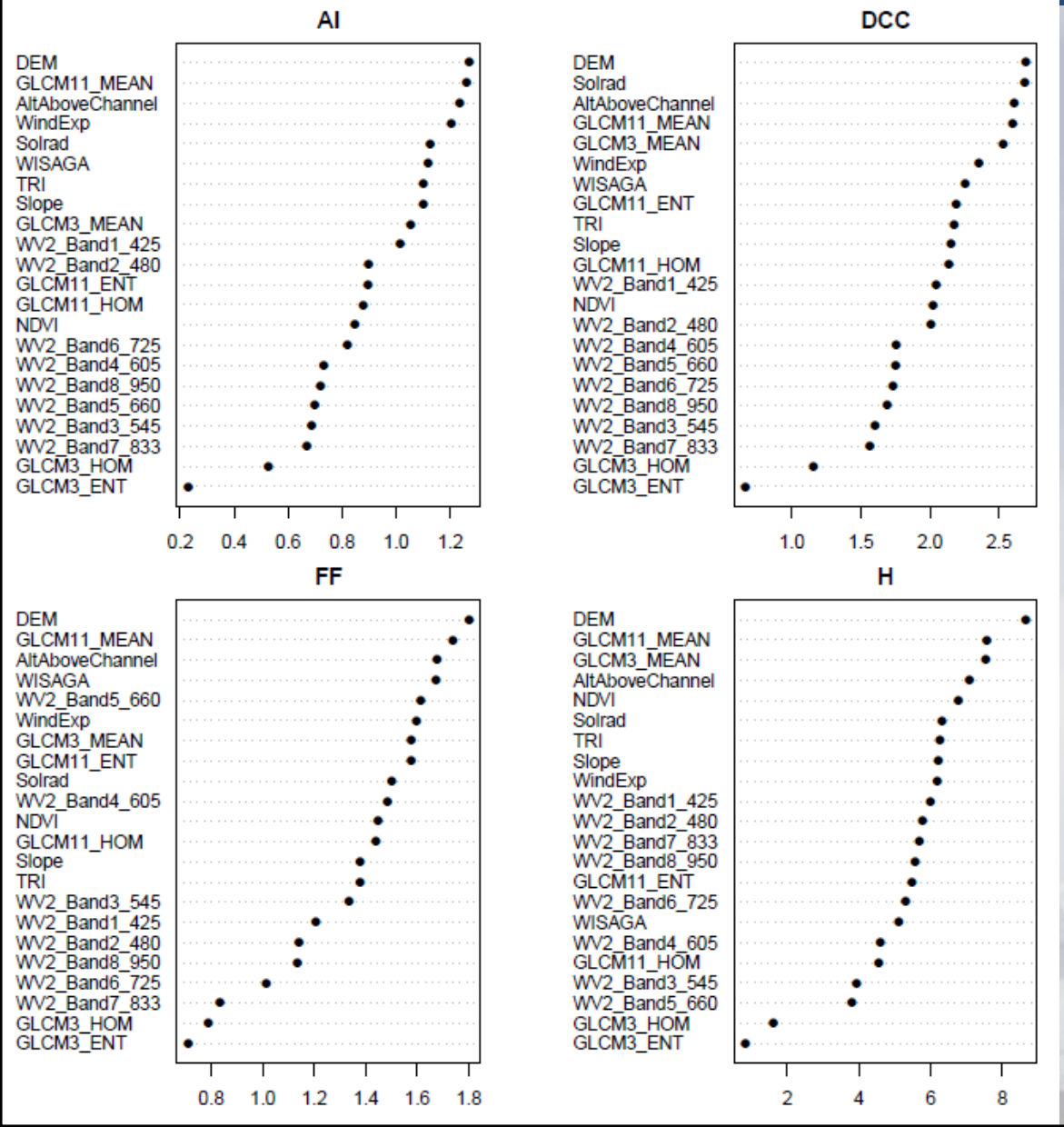
RF classification results

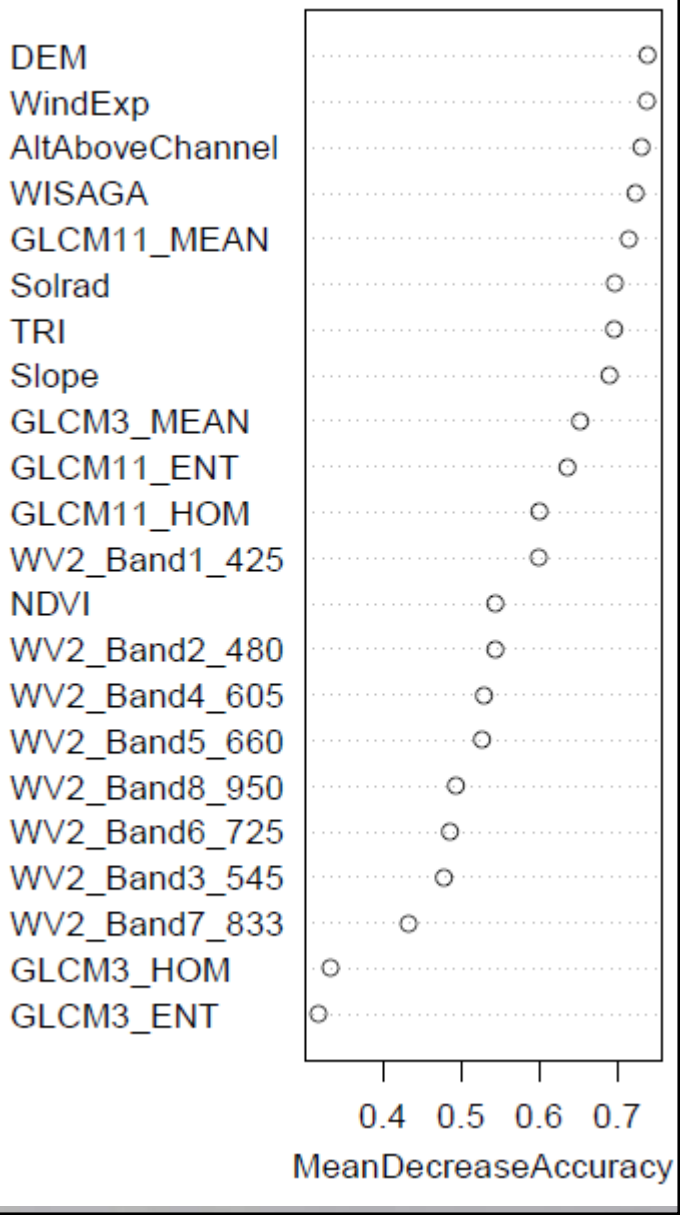


Variable importance I

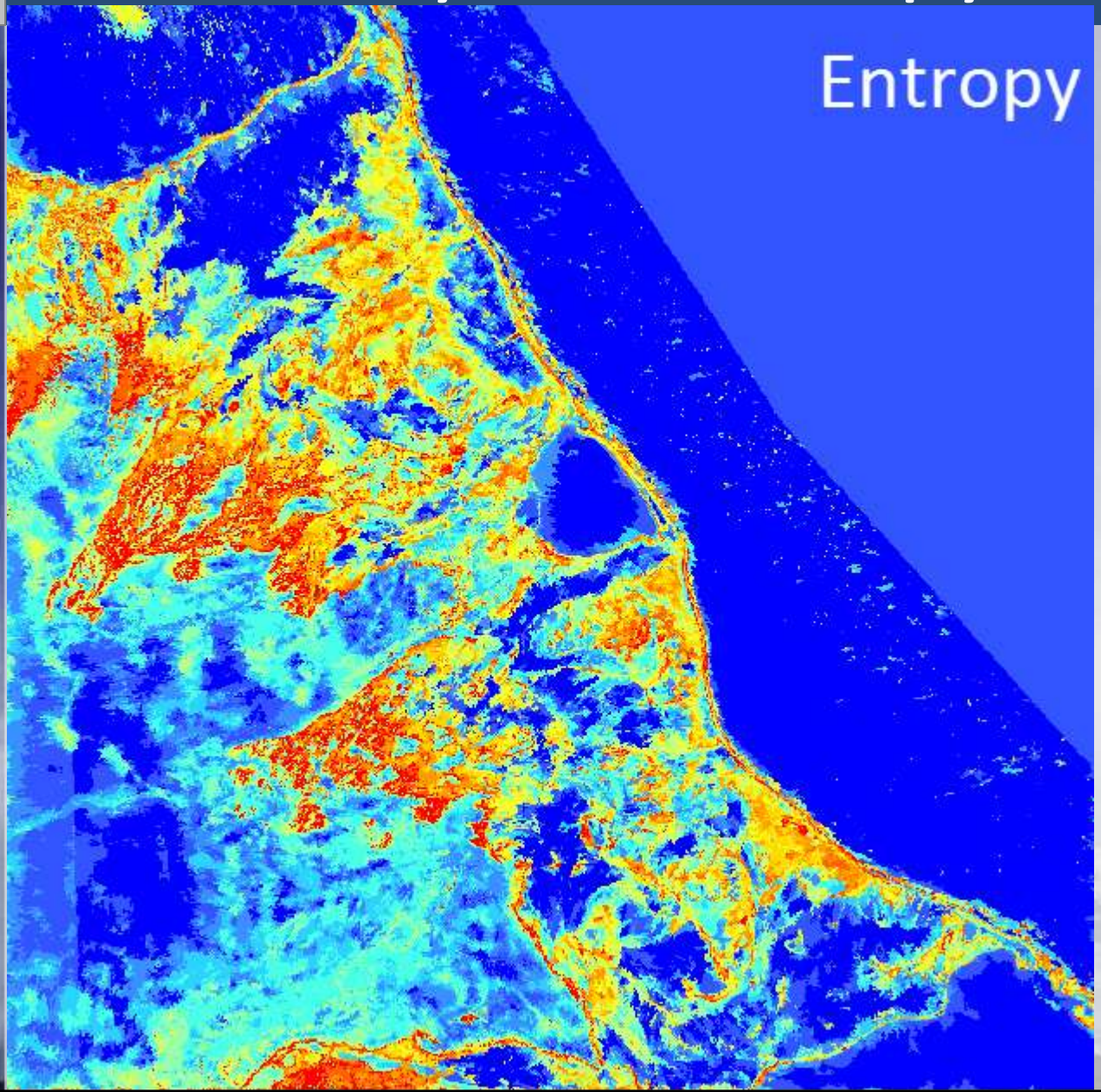


Variable importance II





Probability and Entropy



0.944001 - 1.000000
0.832001 - 0.944000
0.695501 - 0.832000
0.560001 - 0.695500
0.415501 - 0.560000
0.309001 - 0.415500
0.227501 - 0.309000
0.163001 - 0.227500
0.111001 - 0.163000
0.063001 - 0.111000
0.021501 - 0.063000
0.000000 - 0.021500

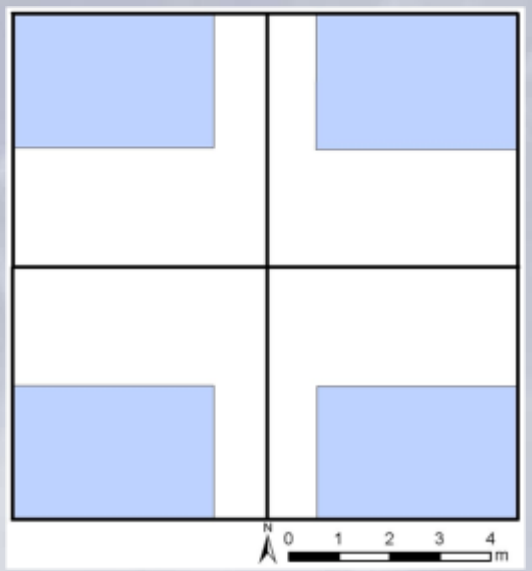
Entropy

1.717299 - 2.037054
1.560340 - 1.717298
1.421244 - 1.560339
1.290048 - 1.421243
1.156342 - 1.290047
1.014095 - 1.156341
0.844705 - 1.014094
0.653595 - 0.844704
0.453074 - 0.653594
0.259888 - 0.453073
0.094825 - 0.259887
0.000000 - 0.094824

Accuracy assessment – really?!

```
 OOB estimate of  error rate: 0.27%
Confusion matrix:
        AI DCC   FF  H Ice  MF  PCC Rock Snow   T Water Wave class.error
AI    2860   0    0  0   0   0    0    0    0   0     0    0 0.000000000
DCC      0 540    2  0   1   0    1    0    0   0     0    0 0.007352941
FF       3   0 1267  0   0   1    0    1    0   0     0    0 0.003930818
H        1   1    0 44   0   0    0    0    0   0     0    0 0.043478261
Ice      0   0    0  0 391   0    0    0    0   0     0    0 0.000000000
MF       0   0    4  0   0 168    1    1    0   0     0    0 0.034482759
PCC      2   1    1  0   0   0 1004    0    0   0     0    0 0.003968254
Rock     0   0    0  0   0   0    0  660    0   0     0    0 0.000000000
Snow     0   0    0  0   0   0    0    0  313   0     0    0 0.000000000
T        2   0    0  0   0   0    0    0    0 602     0    0 0.003311258
Water    0   0    0  0   0   0    0    0    0   0   556    0 0.000000000
Wave     0   0    1  0   0   0    0    0    0   0     0  453 0.002202643
```

Fieldwork needed - Multiscale sampling



Source: Phillippa Bricher

Deglaciation - colonisation

Source: Jenny Scott

Conclusions

- GEOBIA provides a powerful framework for image classification
- Spectral and terrain variables characterise the landscape
- Random forest classifier can be used with GEOBIA to classify objects into vegetation and land cover classes
- Robust multi-scale field sampling is required for training and validation

Questions?

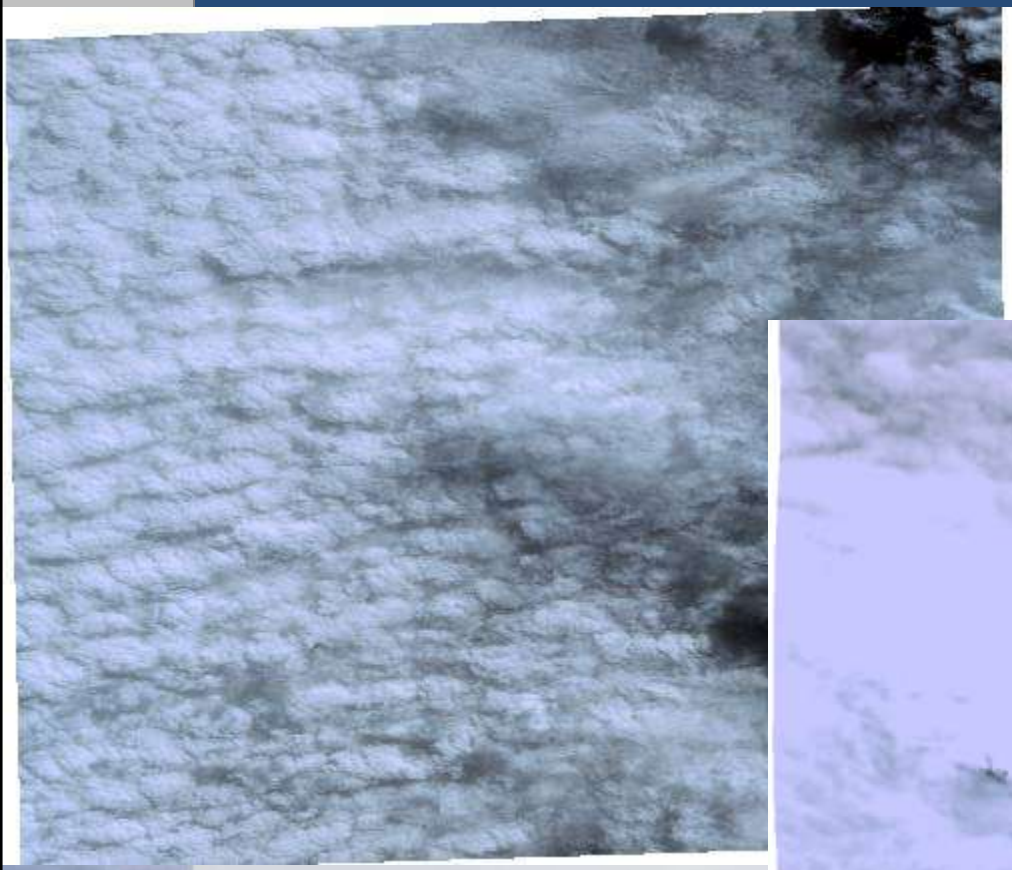
Contact details:

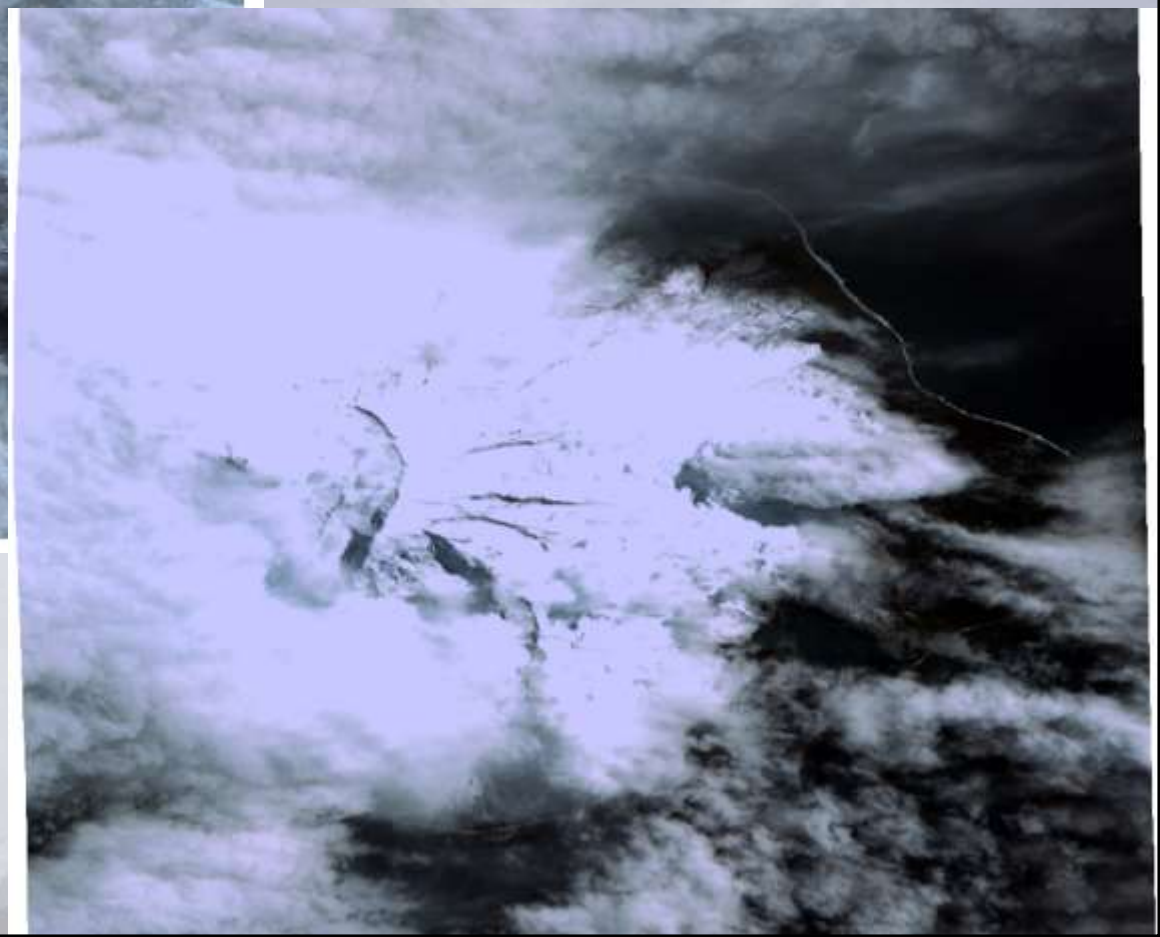
Dr Arko Lucieer
University of Tasmania
School of Geography & Environmental Studies
Private Bag 76, Hobart, Tasmania 7001
Arko.Lucieer@utas.edu.au
http://www.lucieer.net

Acknowledgements:

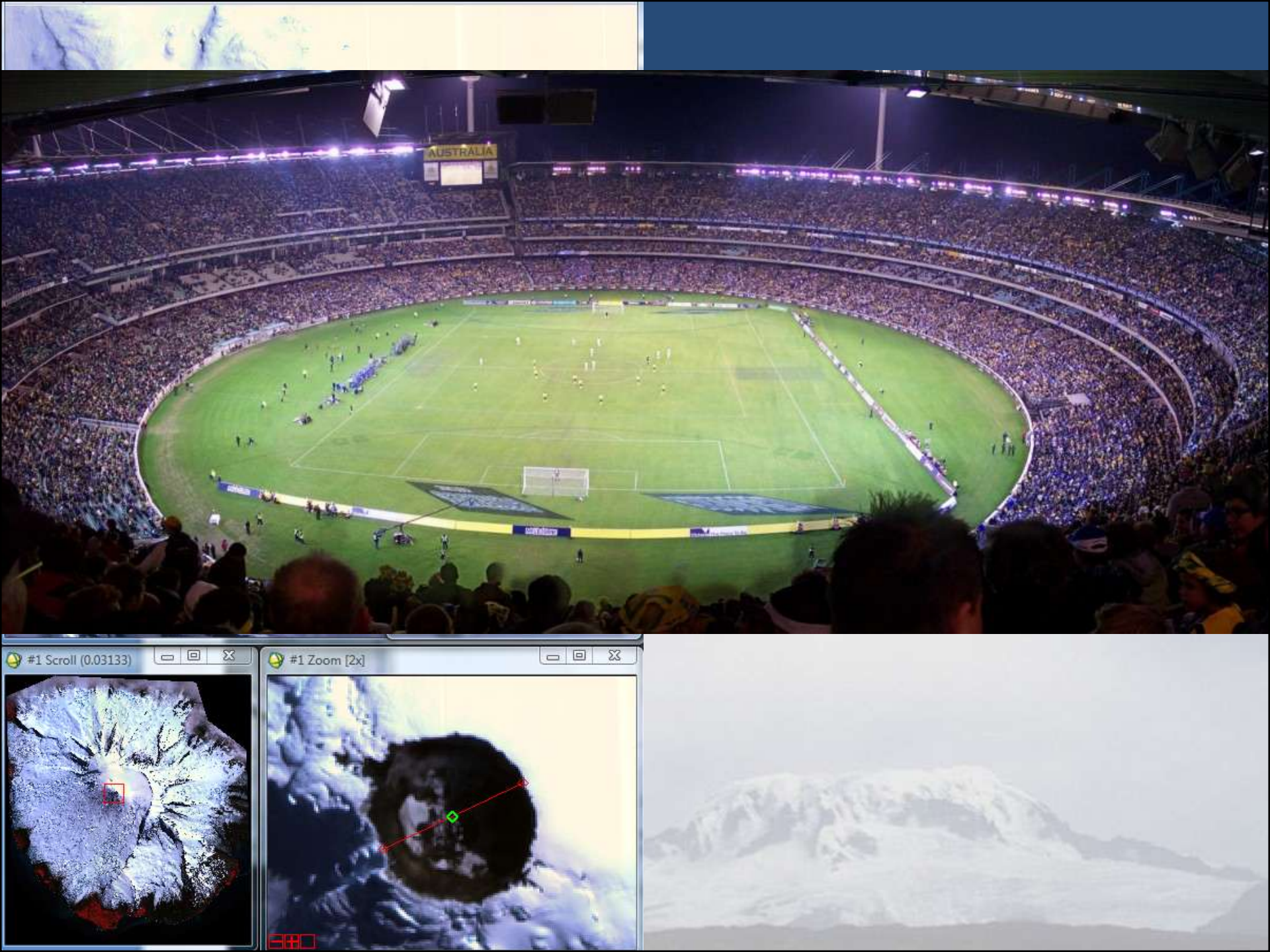
- Australian Antarctic Division: Science Grant (AAS2939)
- Australian Antarctic Data Centre (AADC): Imagery
- Theresa Adams: image pre-processing and change detection
- Jess Benjamin: ground truth identification and classification
- Iain Clarke: WorldView-2 geometric rectification
- Desiree Treichler: WorldView-2 atmospheric correction

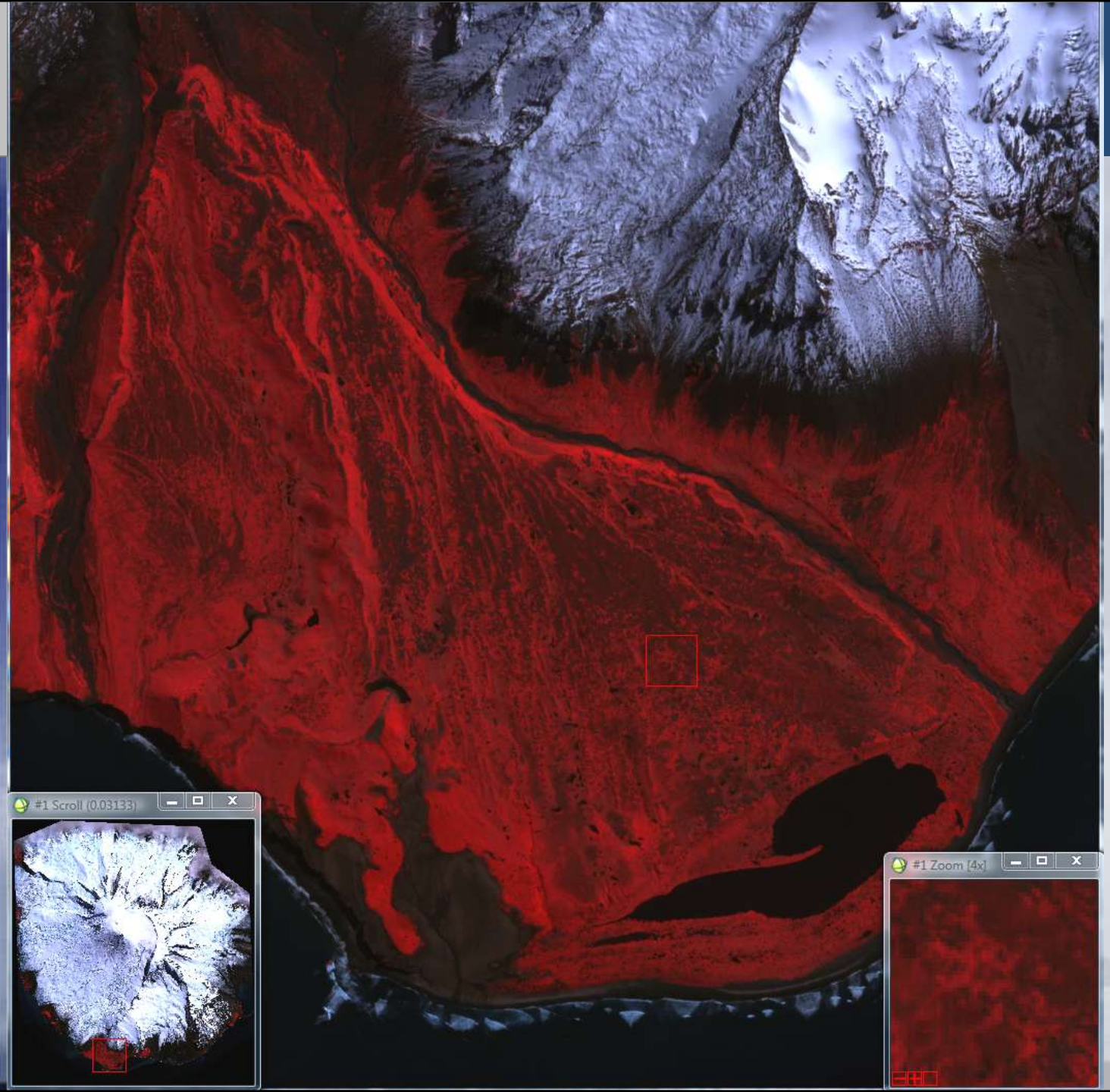
New acquisitions



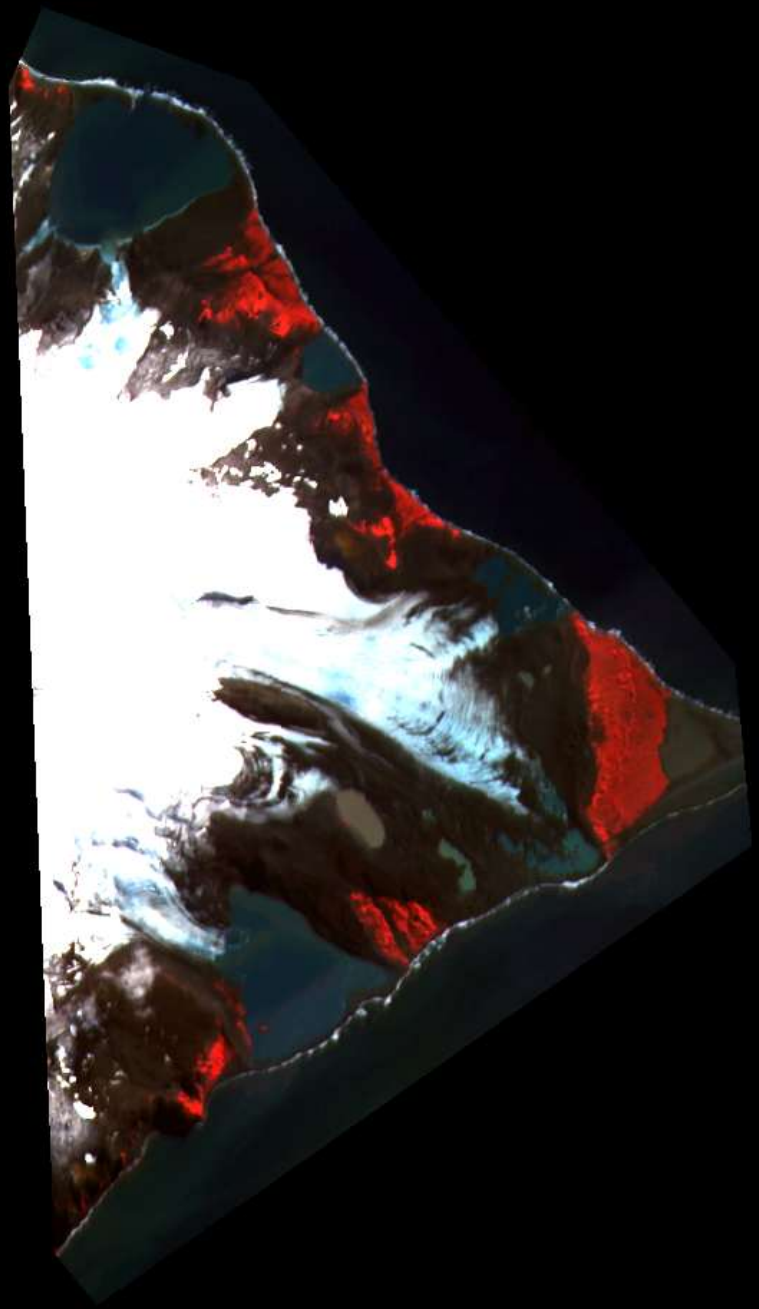



QuickBird 4 Feb. 2009

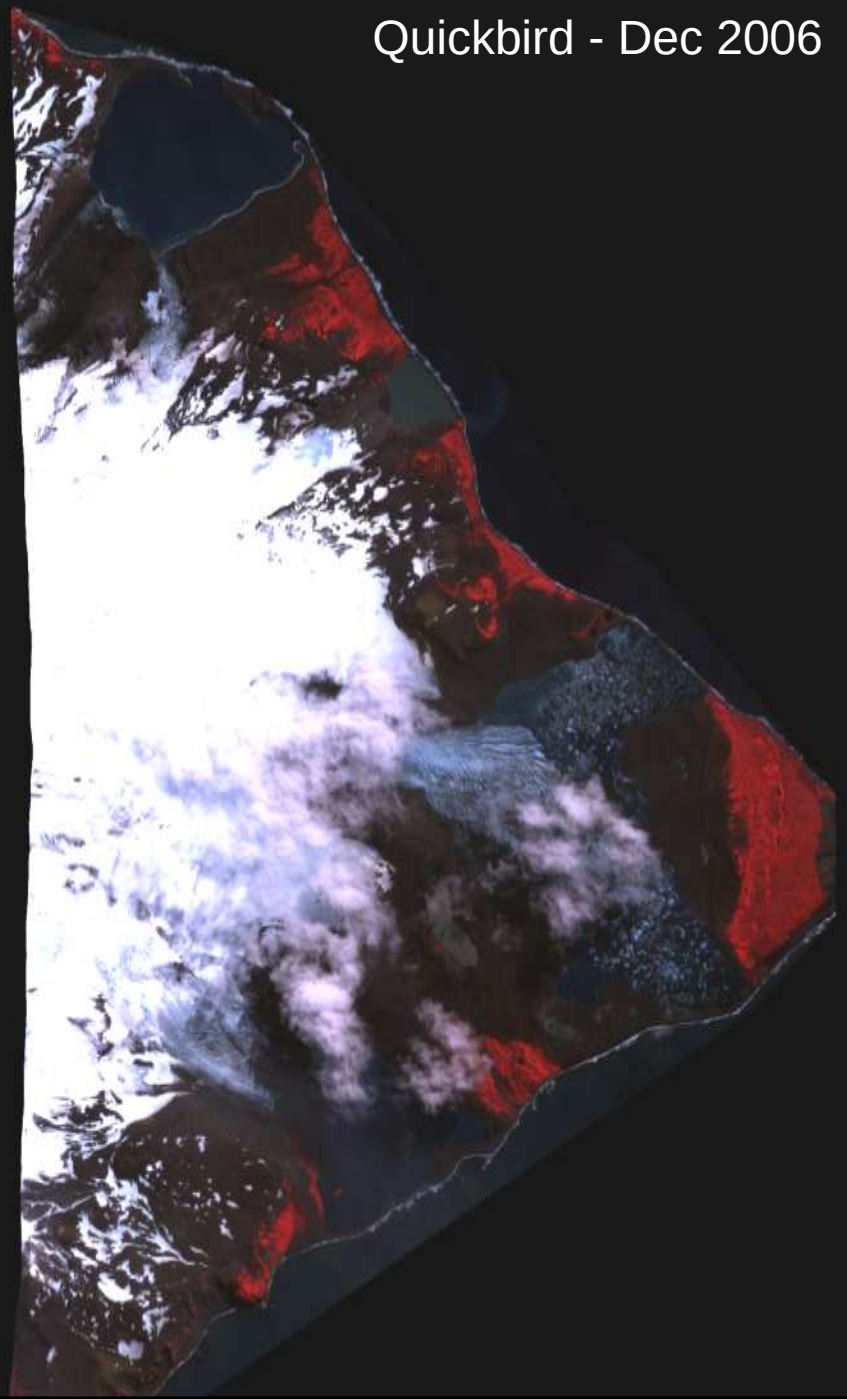

AUSTRALIA
#1 Scroll (0.03133)
#1 Zoom [2x]


#1 Scroll (0.03133)
#1 Zoom [4x]

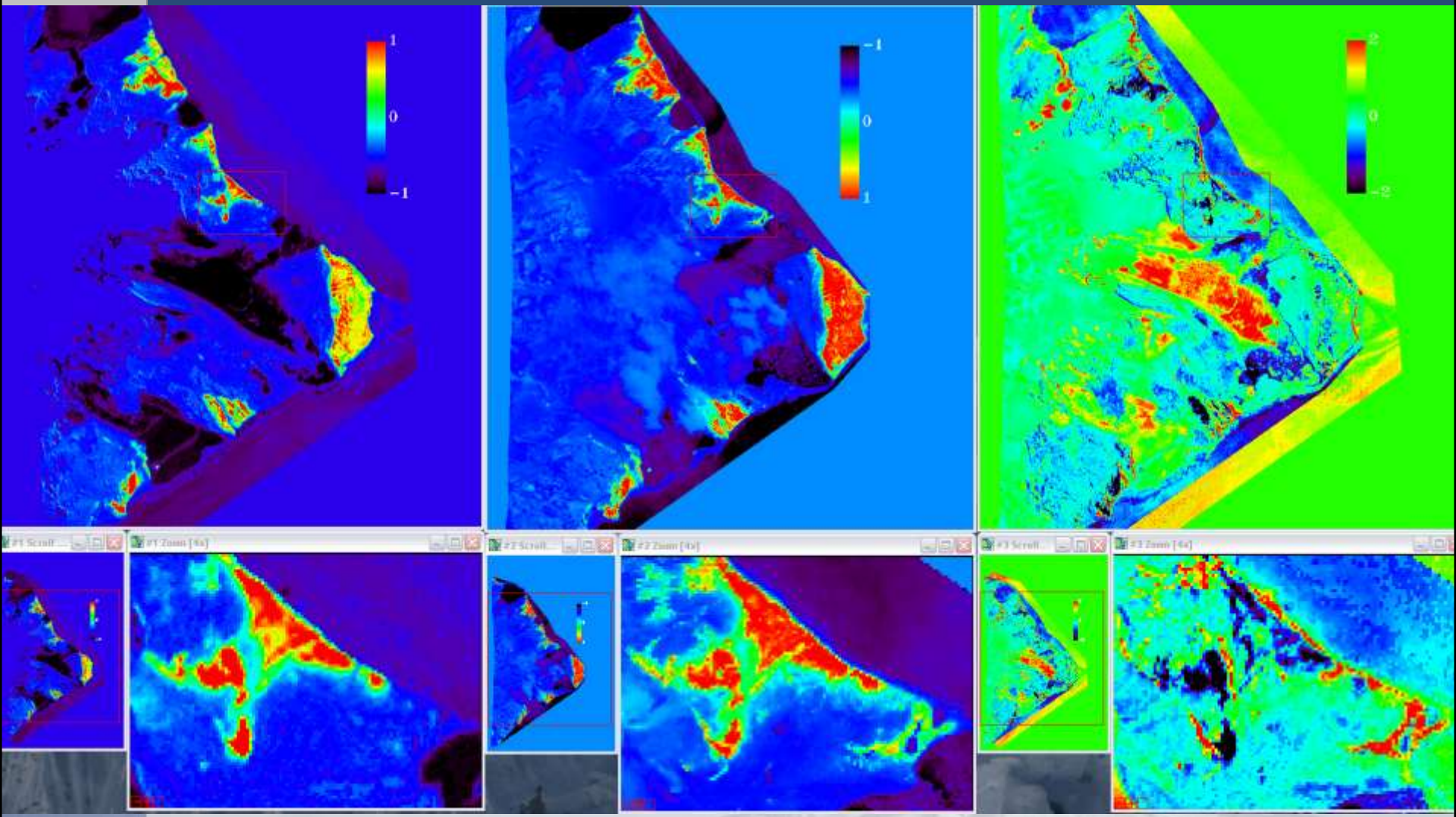
SPOT - Mar 1991



Quickbird - Dec 2006



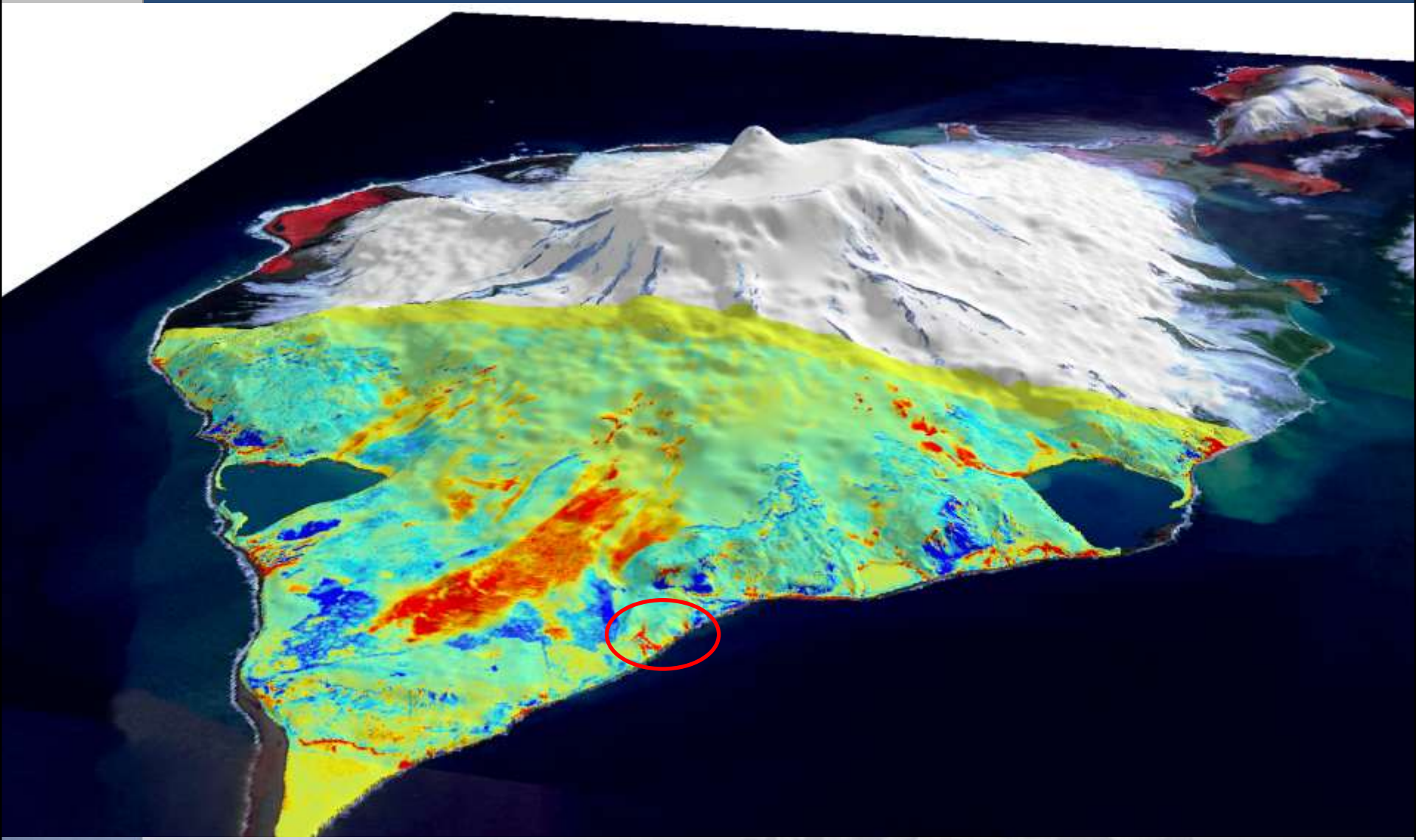
NDVI difference 1991 - 2006

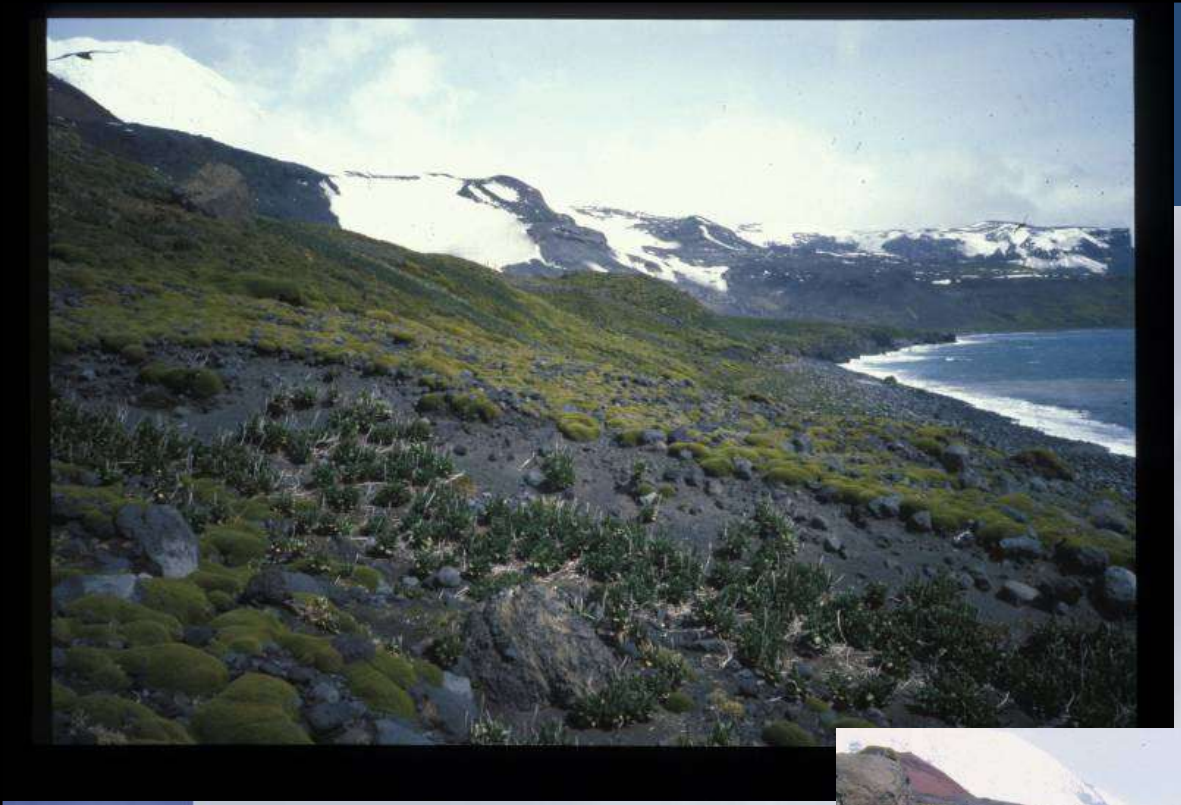


NDVI 1991

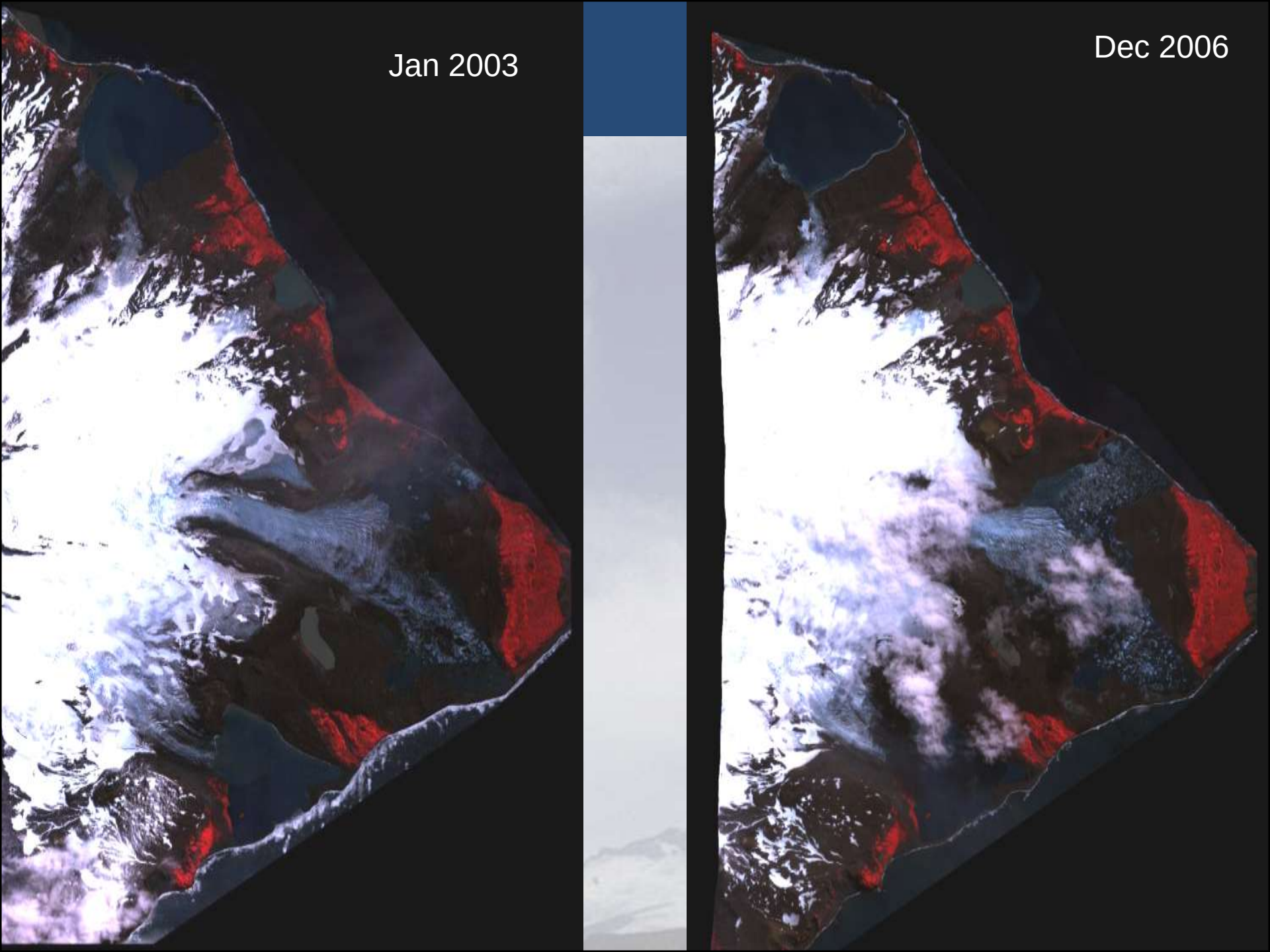
NDVI 2006

NDVI difference





Photos by Jenny Scott


Jan 2003
Dec 2006