F3 - Fire 3 module: accuracy assessment using the Pareto Boundary

Task 1 - Degrade burnt area classifications with different resolutions and apply the Pareto Boundary concepts for accuracy assessment of burnt area maps

Dataset description |working directory: FIRE3

Southern Sudan (Large fires) Landsat 7 ETM+ Path/Row – 173/055 Pre Fire date – 01/12/2000 Pos fire date – 24/12/2000 ***s_sudan*** – RGB color composite - 7,4,3 ***s_sudan_fire*** – burnt area classification (vector file) ***s_sudan30m*** – burnt area classification (grid file)	
Brazil (Small fires) Landsat 5 TM Path/Row – 231/068 Pre Fire date – 16/08/2000 Pos fire date – 17/09/2000 ***brazil*** – RGB color composite – 5,4,3 ***brazil_fire*** – burnt area classification (vector file) ***brazil30m*** – burnt area classification (grid file)	

Russia (Large fires)

Landsat 7 ETM+
Path/Row – 124/022
Pre Fire date – 10/04/2000
Pos fire date – 01/09/2000

russia – RGB color composite - 7,4,3
russia_fire – burnt area classification (vector file)
russia30m – burnt area classification (grid file)

Colombia (Small fires)

Landsat 7 ETM+
Path/Row – 006/057
Pre Fire date – 20/12/1999
Pos fire date – 05/01/2000

colombia – RGB color composite – 7,4,3
colombia_fire – burnt area classification (vector file)
colombia30m – burnt area classification (grid file)

Southeastern Sudan (Large fires)

Landsat 7 ETM+
Path/Row – 171/056
Pre Fire date – 01/01/2000
Pos fire date – 25/01/2000

se_sudan - RGB color composite – 7,4,3
se_sudan_fire – burnt area classification (vector file)
se_sudan30m – burnt area classification (grid file)

Eastern Colombia (Small fires)

Landsat 7 ETM+
Path/Row – 005/056
Pre Fire date – 01/01/2000
Pos fire date – 14/01/2000

e_colombia – RGB color composite – 7,4,3
e_colombia_fire –burnt area classification (vector file)
e_colombia30m – burnt area classification (grid file)

Mongolia (Large fires) Landsat 7 ETM+ Path/Row – 127/027 Pre Fire date – 10/03/2000 Pos fire date – 10/04/2000 ***mongolia -*** RGB color composite – 7,4,3 ***mongolia_fire*** – burnt area classification (vector file) ***mongolia30m*** – burnt area classification (grid file)	
Zimbabwe (Small fires) Landsat 7 ETM+ Path/Row – 172/073 Pre Fire date – 01/08/2000 Pos fire date – 28/09/2000 ***zimbabwe*** – RGB color composite -7,4,3 ***zimbabwe_fire*** – burnt area classification (vector file) ***zimbabwe30m*** – burnt area classification (grid file)	

objective: Compare the different resolutions and visualize the effects of pixel size degradation, in terms of lost detail and burnt area mapping/estimation error (choose 1 of the 4 possible (A-D) pairs of images to work on).

1. Load the composites corresponding to the chosen dataset, using one data frame for each one of the composites. Right-click on the data frame, and choose A*ctivate* to add or visualize data in the frame. The following steps will be applied to each one of the images separately;

2. Load the vector file that represents the burnt areas in each composite; Open the attribute table – the field Value identifies the polygons as fire scars (value 1) or non fire (value 0). Change the vector layer symbology, to C*lassified*, *Unique Values*, using the field Value. Click *Add all Values* and change the symbology of both values to no fill. Change the outline color to yellow for value 1. This allows the fire scars to be distinct, when overlaying the composite. If necessary, increase the outline width and maintain the vector file on top of contents table;

3. Right-click on the layer and select *Zoom to layer*; Zoom in and explore the classification and check the detail in the fire perimeter borders. Click *Full extent* to display the entire dataset;

4. Load the original 30 meters pixel size classification grid and zoom in to explore the grid detail;

5. In the ArcToolbox select *Aggregate.* With this tool degrade the classification to 250 meters resolution. **Which cell factor should you use? As aggregation method you will use Mean. Why?**

6. Open the attribute table of the grid you created. Right-click over the field value and sort ascending. **The pixel values in the 250-m grid displays floating values between 0 and 1, while the 30-m grid assumed integer values 0 and 1. Why?**

7. Close the attribute table and right-click the grid, select *Properties*, *Symbology, Classified*. Click *Classify*, and choose *Equal Interval* as classification method, with 20 classes; Zoom in through the map and vector file, switching between the 250-m and 30-m classification, to visually assess the differences in detail between both maps. **There is a smoothing effect on the edges of the fire scars. Why?**

8. Degrade the classification grid to 500, 1000 and 8000 meters resolution, and apply the same symbology as in 7. Compare the different resolution, switching the grids on and off. Zoom in to raster resolution in the over a small fire polygon. You should observe how the degradation affects the detail and size of mapped fires (fig.1). Compare the different resolutions and different datasets with distinct fire patterns, i.e. small fires vs large fires.

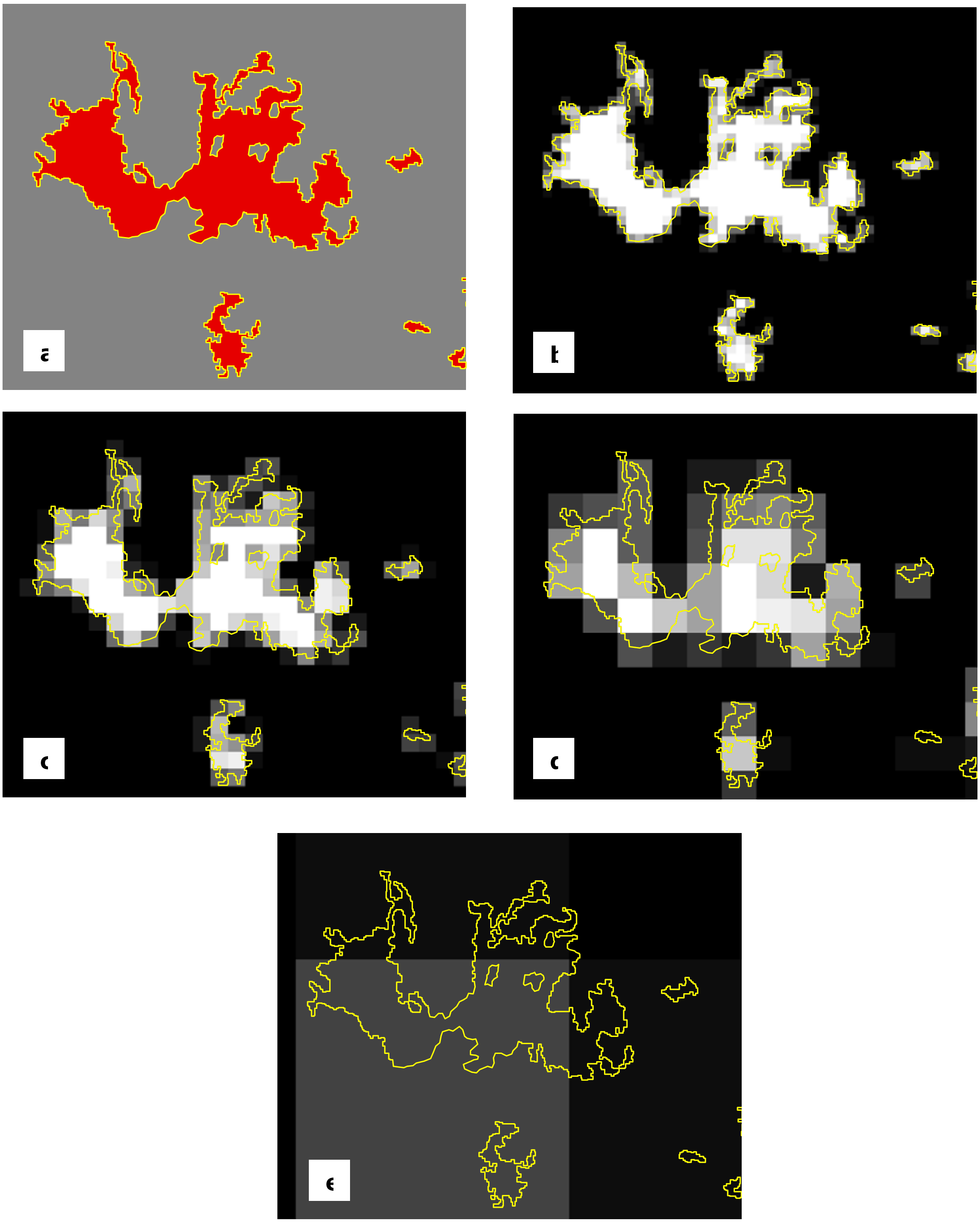


Figure 1 – Fire scars in 30-m Landsat derived classification (a) and burnt area maps degradated (using pixel aggregation) to 250 (b), 500 (c), 1000 (d) and 8000-m (e) pixel size.

9. **The increase in pixel size implies coarser classifications as previously observed. Which of the images in your dataset is more affected by spatial degradation? Justify.**

10. Minimize ArcMap. Browse the data directory to the *PBcalc* folder. It contains input files for the program that calculates the Pareto Boundary – **P**areto **B**oundary **calc**ulator (*PBcalc).* These include a binary file and a header file (.hdr) and replicate the ones generated in the previous steps, converted to a binary format compatible with *PBcalc*. The Pareto Boundary will be calculated for the **250, 1000** and **8000-m** resolutions, and for each image separately; therefore the process will be completed after 6 runs. Figure 3 represents the construction of the Pareto Boundary.

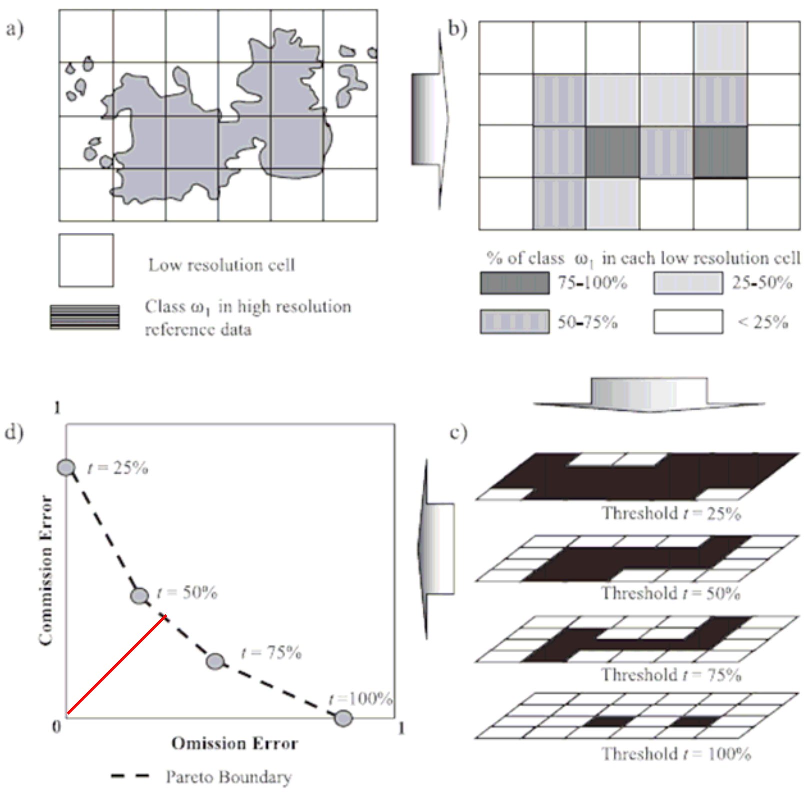


Figure 3. Procedure for generating a discrete set of points belonging to the Pareto Boundary, starting from the high-resolution map, for a desired low spatial resolution: (a) a low-resolution grid is overlaid on the high-resolution reference map; (b) the percentage of class ω1 is computed for each cell; (c) a set of low-resolution products is generated by thresholding the percentage of class ω1; (d) the confusion matrix is produced for each one of these maps. Omission error and commission error are derived and plotted in the omission error/commission error space (δ is the minimum distance between the Pareto boundary and the origin of the coordinate system, or perfect classification (from Boschetti et al. 2004).

11. Run *Pbcalc* (fig.4). You will be asked to input the name and location (complete path) of the input file, as well as the number of columns (or samples) and rows. This last information is included in the header file (.hdr), which can be opened with Notepad. The program will also need a name and destination (complete path) of the output Ascii (.txt) file. This output consists of a table with the following information:
    1. Fire/no fire classification threshold (0.05 increments)
    2. Omission error
    3. Comission error
    4. Distance between (omission error, commission error) pairs of values and the origin of the graph (see Fig. 3c - δ)

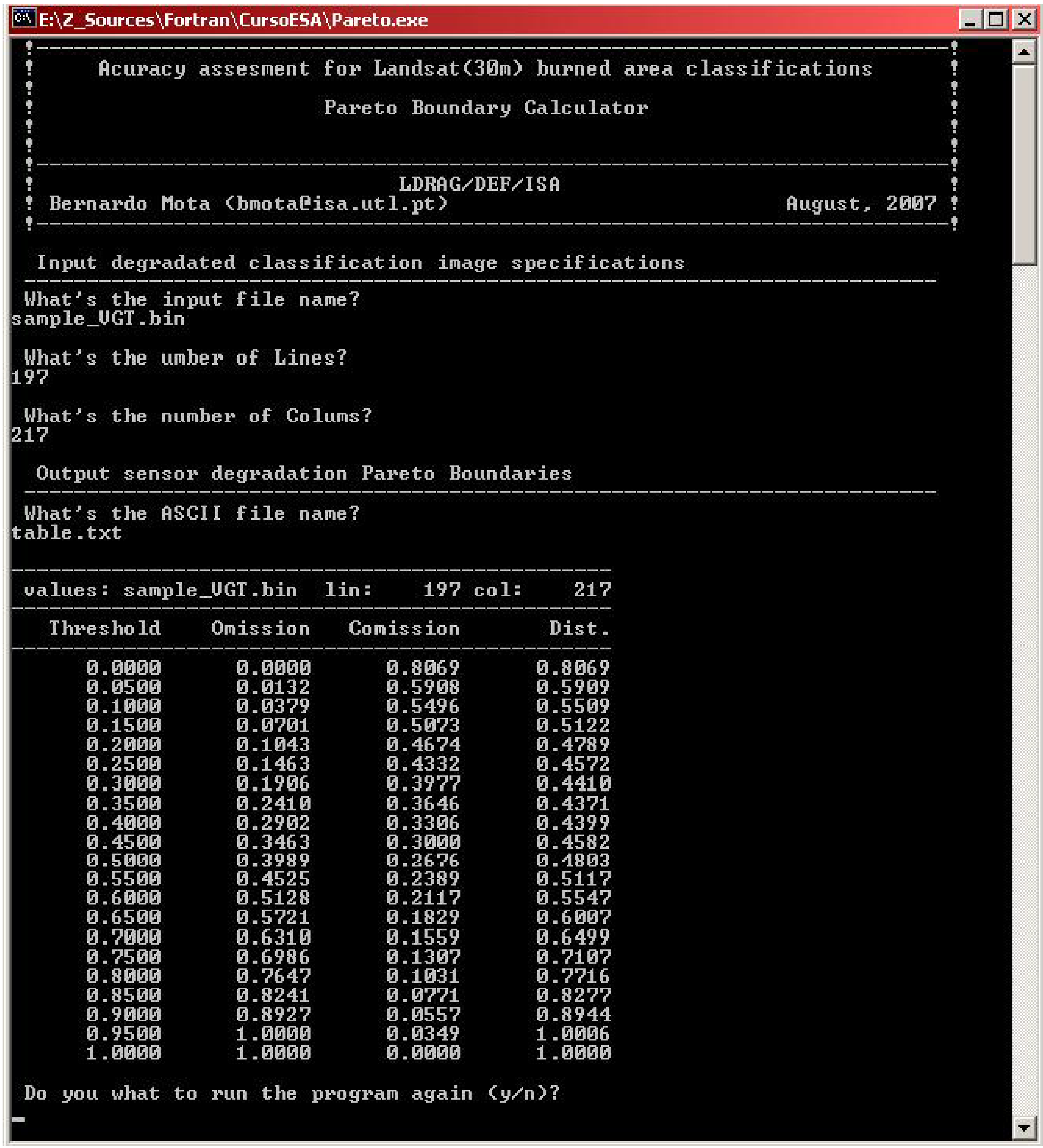
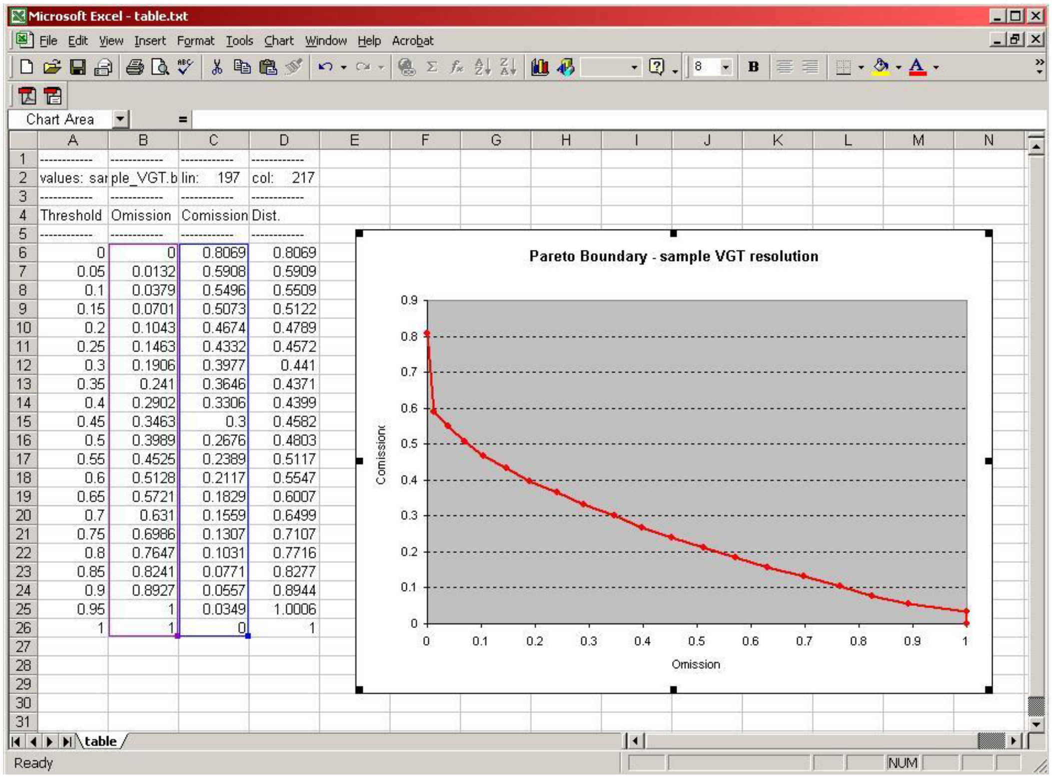


Figure 4. Sample run of the Pareto Boundary Calculator.

12. Import the output Ascii file to MS EXCEL (fig. 5) and plot the following: 1-The curves from the 3 resolutions for the same image; 2-The curves from the different images of the same pair, for the same resolution. **What can you conclude?**



	A	B	C	D
1	------------	------------	------------	------------
2	values: sar	ple_VGT.b	lin: 197	col: 217
3	------------	------------	------------	------------
4	Threshold	Omission	Comission	Dist.
5	------------	------------	------------	------------
6	0	0	0.8069	0.8069
7	0.05	0.0132	0.5908	0.5909
8	0.1	0.0379	0.5496	0.5509
9	0.15	0.0701	0.5073	0.5122
10	0.2	0.1043	0.4674	0.4789
11	0.25	0.1463	0.4332	0.4572
12	0.3	0.1906	0.3977	0.441
13	0.35	0.241	0.3646	0.4371
14	0.4	0.2902	0.3306	0.4399
15	0.45	0.3463	0.3	0.4582
16	0.5	0.3989	0.2676	0.4803
17	0.55	0.4525	0.2389	0.5117
18	0.6	0.5128	0.2117	0.5547
19	0.65	0.5721	0.1829	0.6007
20	0.7	0.631	0.1559	0.6499
21	0.75	0.6986	0.1307	0.7107
22	0.8	0.7647	0.1031	0.7716
23	0.85	0.8241	0.0771	0.8277
24	0.9	0.8927	0.0557	0.8944
25	0.95	1	0.0349	1.0006
26	1	1	0	1

Figure 5. Sample Pareto curve plot in MS EXCELL.

13. For each image, choose the optimal fire/no fire classification threshold, i.e., **what should be the best compromise between omission and commission errors?**

14. Return to ArcMap. Remove the 500 meter grid.Right-click on the 250-m grid copy and paste it above itself. Using the classified symbology, create 2 classes using as breaking value the Pareto optimum treshold.

15. In the ArcToolbox select *Reclassify* and choose the one included in the Spatial Analyst Toolbox. Choose as input raster the result from step 15 and as reclass field, Value. The values bellow the breaking point will be given the value 0 (non-fire) and values above breaking point will be given value 1 (fire).

16. Repeat the same procedure to the 8000-m grid. Compare the 20-classes map and the optimum 2-classes map for both tested resolutions and images (fire patterns) **Why does this 2-classes map represent the best classification?**

17. To fill the table below, double-click the grid, *Properties*, *Source, Raster information*, and fill the pixel size column. Use the information on the attribute table to complete the number of pixels column, and calculate the burnt area for the 3 resolutions and 2 images. Complete the table using the same procedure for the reclassified grids.

	Large fires			Small fires		
	pixel size	**nº pixel**	**area**	**pixel size**	**nº pixel**	**area**
30-m						
250-m						
1000-m						
8000-m						

18. What are the main conclusions you can derive from this exercise?

REFERENCES

Boschetti L., Flasse S., Brivio P., 2004, Analysis of the conflict between omission and commission in low spatial resolution dichotomic thematic products: The Pareto Boundary. *Remote Sensing of Environment* 91, 280–292.