Classification of SST Quality Using a Combined Forest of Weak and Strong Classifiers

Katherine Kilpatrick, Susan Walsh, Elizabeth Williams, Peter Minnett
University of Miami Rosenstiel School for Marine and Atmospheric Science
4600 Rickenbacker Causeway, Miami, Fl 33149 USE email:kkilpatrick@rsmas.miami.edu

ABSTRACT

Presented are results from a cloud/quality classification model based on boosted Alternating Decision Trees for VIIRS SST imagery. Identification and exclusion of clouds from satellite based Infrared measurements of sea surface temperature (SST) is critical to achieve accurate retrievals. Historically identification of clouds has been driven primarily by a few uniformity tests within a small box, brightness temperature minimums, and comparisons to low-resolution gap-free reference fields. These tests do an adequate job of identifying large cold clouds, and uniformity tests identify moderate sized patchy clouds, but the efficacy of these tests often decreases at cloud edges, small wispy clouds, and low level clouds and fog, when cloud temperature can be closely equilibrated with the sea surface, particularly at high latitudes. The heavy reliance on just a few uniformity thresholds is often overly conservative near cloud edges, and in strong geophysical SST frontal regions.

The use of a majority vote from an ensemble of both weak and strong cloud classifiers offers the potential to identify more cloud types and improve the retention of SST gradients in the pool of best quality SST retrievals.

An ensemble of 4 Alternating Decision Trees classifiers were trained to classify VIIRS SST retrievals as either clear or cloudy, using 10 fold-cross validation and boosting. The training sets consisted of a subset of randomly selected records in the VIIRS buoy Matchup Database (MUDB).

model cases:

- Night
- Day non glint
- Day moderate
- Day high glint

Only the classification model for day non-glint is shown below

Each ADtree model was validated on a set of independent records not used in testing or training set.

Non- glint classification model MUDB validation data set:

Correctly Classified Instances	29732	91.0015 %
Incorrectly Classified Instances	2940	8.9985 %
Kappa statistic	0.82	
Mean absolute error	0.2073	
Root mean squared error	0.2723	
Coverage of cases (0.95 level)	99.9847 %	
Total Number of Instances	32672	

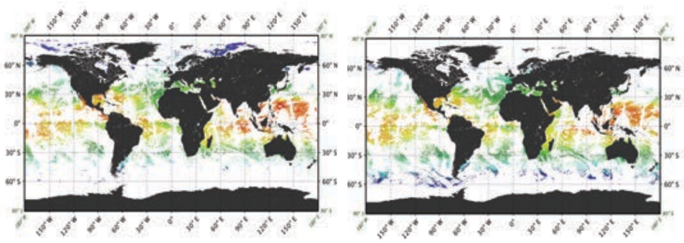
```
confusion matrix      a     b   <-- classified as
              22123  2380 |   a = cloud
               1668 22835 |   b = clear
```

Alternating decision tree Model for day non glint:

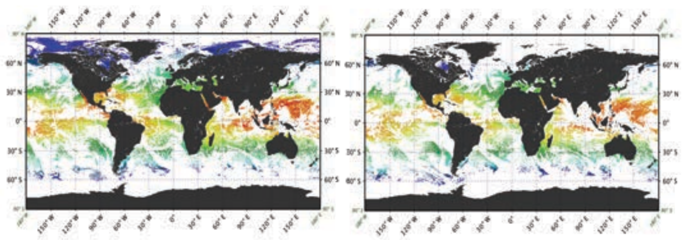
Instance where the Glint coef < 0.005
Decision node:vote
+= confidence good clear -= confidence bad cloud
Final sum votes all TRUE nodes <0 flag as cloud
rho= visible band reflectance BT= brightness temperature Kº

```
: 0
|  (1) rho 1610 < 0.16: 0.805
|  |  (2) rho 748 < 0.062: 0.393
|  |  |  (3)rho 1380 < 0.004: 0.287
|  |  |  |  (9)BT deficit 11um < 0.002: -0.681
|  |  |  |  (9)BT deficit 11um >= 0.002: 0.026
|  |  |  |  |  (13)rho 748 < 0.039: 0.364
|  |  |  |  |  (13)rho 748 >= 0.039: -0.21
|  |  |  (3)rho 1380 >= 0.004: -1.244
|  |  (2)rho 748 >= 0.062: -0.572
|  |  |  (5)min rho 610 5x5 box < 0.032: 0.455
|  |  |  (5)min rho 610 5x5 box>= 0.032: -0.395
|  |  (4)sensor zenith angle < 64.994: 0.216
|  |  |  (8)rho 1380 < 0.007: 0.065
|  |  |  (8)rho 1380 >= 0.007: -1.077
|  |  (4)sensor zenith angle >= 64.994: -0.708
|  (1)rho 1610 >= 0.16: -1.755
|  |  (6)rho 1610 < 0.266: 0.642
|  |  (6)rho 1610 >= 0.266: -0.19
|  |  |  (14)max-min rho 678 5x5 box < 0.103: 0.425
|  |  |  (14) max -min rho 678 5x5 box >= 0.103: -0.195
|  |  (10)11um-12um BT < 0.235: -0.189
|  |  (10) 11um-12um BT  >= 0.235: 0.411
|  |  (15)water vapor NCEP Kg/m2 < 2.946: 0.038
|  |  (15) water vapor NCEP Kg/m2  >= 2.946: -1.137
|  (7)max -min 11um BT 5x5 box < 0.762: 0.156
|  (7) max -min 11um BT 5x5 box  >= 0.762: -0.188
|  (11) water vapor NCEP Kg/m2  < 1.315: 0.327
|  (11) water vapor NCEP  Kg/m2 >= 1.315: -0.054
|  |  (12)sst < 278.171 Kº: -0.679
|  |  (12)sst >= 278.171 Kº: 0.05
Legend: -ve = Bad, +ve = Good
Tree size (total number of nodes): 46
Leaves (number of predictor nodes): 31
```

MODIS AQUA L3 4km



Day SST VIIRS L3 4km Night SST



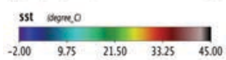
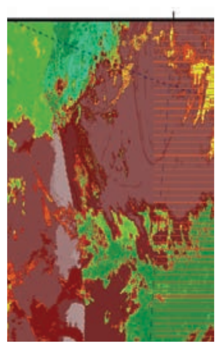
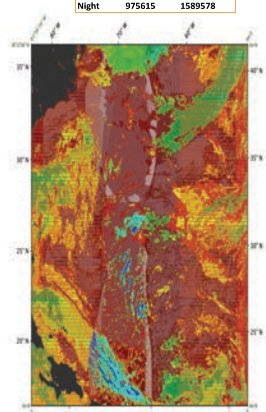


Table 1 Clear pixel count L3 Global Image

Pixel Quality level	MODIS A SST L3 4km count	VIIRS SST L3 4km count
Best		
Day	2286952	3998602
Night	2512379	3156603
Good		
Day	1245048	2711723
Night	975615	1589578





Level-2 1km pixel ensemble majority Vote Day time

enlarged area same gulf stream area as SST frontal region image to the right

Concept of Alternating Decision Trees

The classification method of Alternating decision trees (ADTree) (Freund and Mason 1999, Pfahringer et. al . 2001) trains a collection of binary decision nodes each ending with a prediction node. Nodes contain a vote that is scaled to the predictive power of the test. When combined with boosting algorithms, where at each training iteration instances that were previously misclassified are given a larger weight, a very accurate ensemble classification model can be developed.

To predict cloudiness a pixel to be classified transits all decision nodes that are true, and the prediction values from all true nodes are summed to form the final vote. For VIIRS imagery a positive sum is clear and a negative vote is cloud. The magnitude of the vote provides an indication of the confidence of the classification for a given pixel.

In some instances **the combined vote from a collection of weak prediction nodes when voting together as a block can modify or over ride the vote of a single strong prediction node.** This ensemble method is a very different classification strategy from the single decision tree method currently used for NASA/NOAA Pathfinder and NASA MODIS SST products.

Reference:
Freund, Y., Mason, L.: The alternating decision tree learning algorithm. In: Proceeding of the Sixteenth International Conference on Machine Learning, Bled, Slovenia, 124-133, 1999.s
Bernhard Pfahringer, Geoffrey Holmes and Richard Kirkby. Optimizing the Induction of Alternating Decision Trees. Proceedings of the Fifth Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining. 2001, pp. 477-487

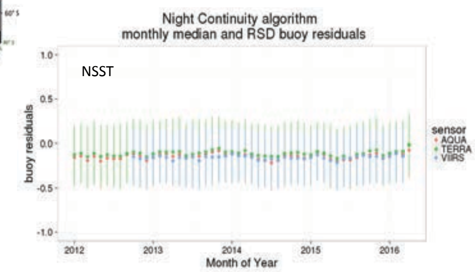
Improved global coverage L3 4km for VIIRS using ensemble of ADTrees

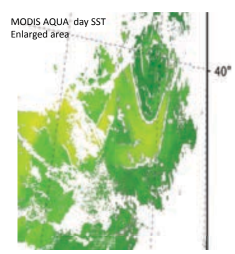
The increased coverage in the VIIRS June 19, 2014 L3 4 km maps compared to MODIS AQUA v2014.0 is very striking, 20% at night and 40% day time.

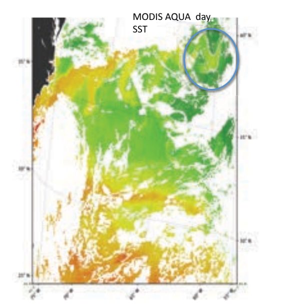
Table 1 shows the count of Best and Good quality pixels. Some of the increase for VIIRS can be attributed to the larger swath, but even at the Best quality level where the range of zenith angles are comparable between sensors the increased coverage remains.

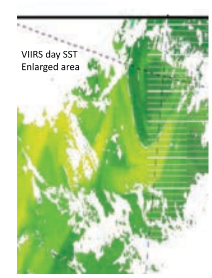
Visual inspection of the images indicates that the gain in the day time VIIRS coverage occurs primarily in polar regions, and for both day and night images in areas around cloud edges.

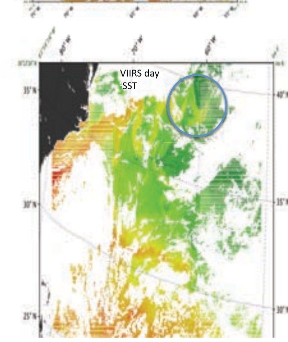
Any increase in coverage must not sacrifice product accuracy. The plot shown below, of monthly median error statistics based on skin SST minus sub-surface buoy SST, for VIIRS and MODIS sensors indicates **the increase in VIIRS coverage is not at the expense of product accuracy.**











Improved coverage at SST frontal regions: L2 day time image (satellite perspective) over the Gulf Stream on June 19th 2014. The VIIRS image using the ensemble of ADTree classifier shows **improved retention of good quality pixels at frontal boundaries compared to a MODIS AQUA image 20 minutes later using a standard decision tree.** Note: the horizontal white lines on the right side of the VIIRS image are missing data related to on-board along scan pixel aggregation/deletion.

UNIVERSITY OF MIAMI
ROSENSTIEL SCHOOL of MARINE & ATMOSPHERIC SCIENCE