GDA Corp.

CASA-AWiFS
(Version 1.7)

White Paper:

Validation of the Cloud and Cloud Shadow Assessment System for AWiFS Imagery

GDA Corp.

Copyright © 2007 GDA Corp. www.gdacorp.com
All Rights Reserved

Algorithm Description

GDA Corp. has developed an innovative system for **C**loud **A**nd cloud **S**hadow **A**ssessment (**CASA**) in Advanced Wide Field Sensor (AWiFS) imagery from the Indian Remote Sensing (IRS)-P6 satellite. The CASA-AWiFS system is an extension of previously developed CASA versions for Landsat, Ikonos, and QuickBird imagery. The system performs per-pixel detection of clouds and cloud shadows with minimal intervention from the software operator. All cloud / cloud shadow masks are generated at nominal, 56 meters pixel resolution.

The GDA algorithm is based on a novel approach to highly automated cloud detection. The major difference of the GDA approach from other cloud detection algorithms lies in its reliance on multiple lines of evidence for labeling a given feature as a cloud or cloud shadow. These multiple lines of evidence are extracted from the scene, the scene metadata, and an ancillary knowledge base. CASA utilizes a set of spectral, spatial, textual, and pattern evidences and uses hierarchical self-learning / self-calibration routines to strengthen and boost the detection and extraction of features of interest. By utilizing an object-based feature extraction approach, the CASA algorithm allows individual spectral, spatial, and/or contextual parameters to be relaxed due to their simultaneous and iterative use. Further, CASA utilizes invariant pattern recognition techniques which are independent of the rotation and scale of the object to identify features of interest. The algorithm is essentially a highly automated expert system that can be used for operational extraction of more or less any predefined class of features. A CASA graphical flowchart is presented in Figure 1.

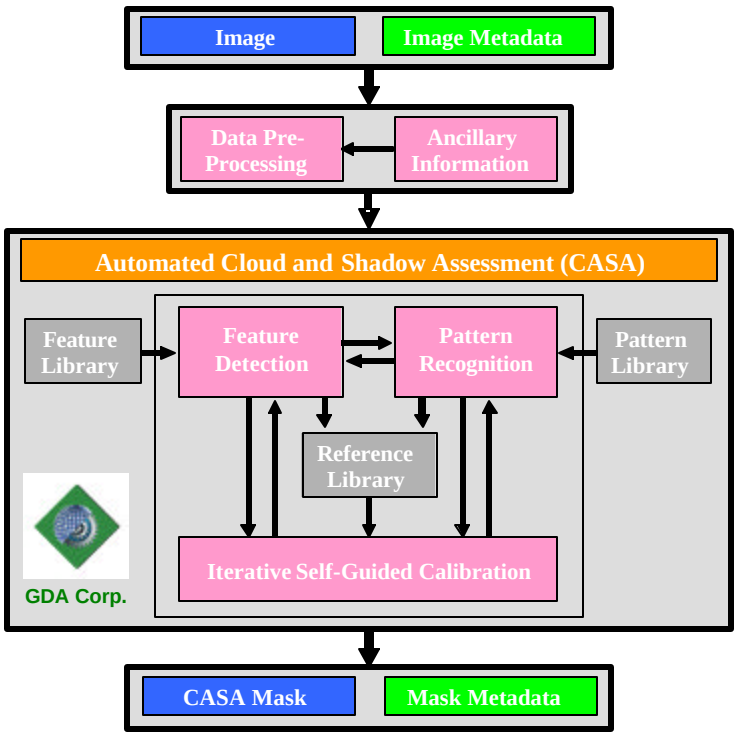

Figure 1: A Flowchart of the GDA CASA System

Performance Analysis

The current version of CASA-AWiFS (version 1.7) has been validated on 63 AWiFS scenes. AWiFS imagery was provided by the USDA Satellite Imagery Archive. The dataset primarily encompassed 2006 imagery for the USA, as this was the data most prominently represented in the USDA Satellite Imagery Archive. Additionally, some scenes were collected for Southern Africa and Europe. The aim of the collection was to obtain scenes covering different seasons and various atmospheric, cloud, haze, and ground conditions. Each scene was visually inspected to assess, separately, percent dense cloud cover and percent light, transparent cloud and haze cover. Results of the individual scene assessments were compiled into the "truth" cloud cover dataset (Appendix 1). The distribution of cloudy and light cloud / hazy scenes within the dataset is presented in Table 1.

% Cloud Cover	Percent of Scenes		
	Dense Cloud Cover	Light Cloud / Haze Cover	Total Cloud / Haze Cover
0 to 10%	44%	92%	38%
0 to 30%	78%	100%	76%
0 to 50%	90%	100%	87%
0 to 70%	100%	100%	97%
0 to 90%	100%	100%	100%
Max Cover	65%	25%	90%

Table 1: Distribution of cloud and light cloud / haze contaminated scenes in the validation dataset

The results of the validation comparing CASA-AWiFS outputs to the visual assessments consistently show a high accuracy of CASA-AWiFS with respect to dense cloud detection (Figure 2). CASA-AWiFS results have also shown a relatively good identification of light, transparent clouds and haze (Figure 3), even though CASA was not specifically designed for light cloud / haze detection. Consequently, the overall assessment of combined, cloud and haze contamination of the imagery has shown very good results (Figure 4).

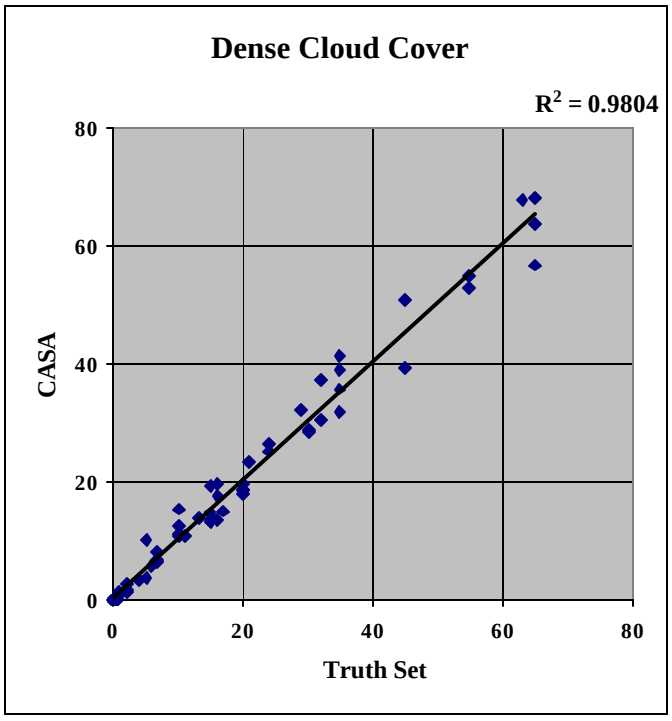

Figure 2: Dense Cloud Detection: CASA *vs.* Truth Dataset

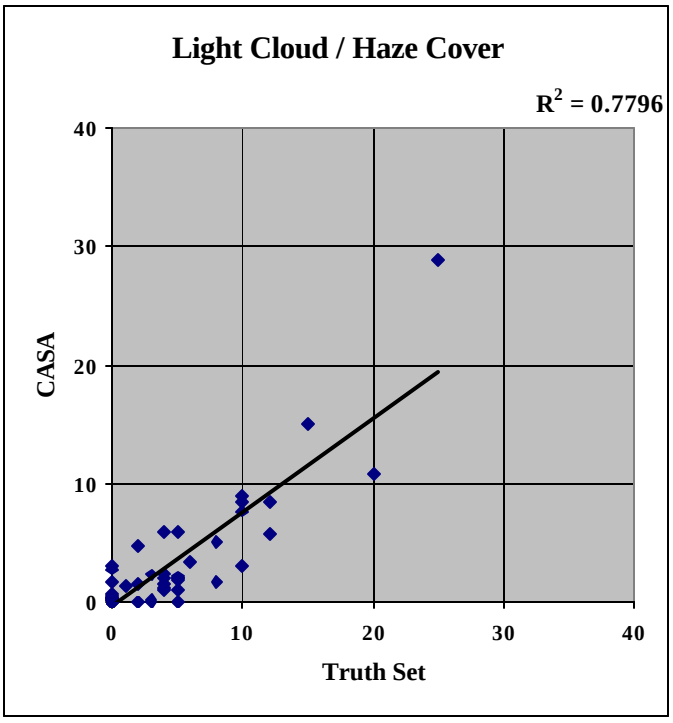

Figure 3: Detection of Light, Transparent Clouds and Haze: CASA *vs.* Truth Dataset

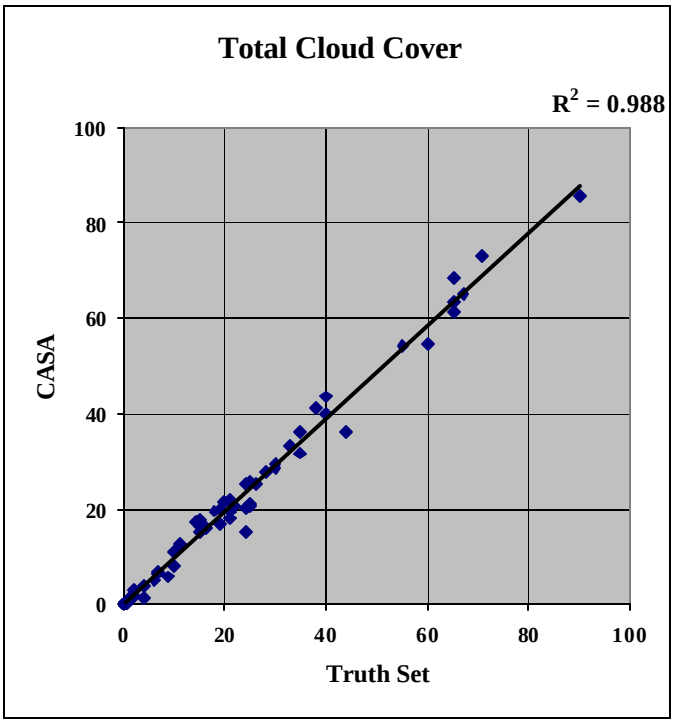

Figure 4: Combined Dense Cloud and Light Cloud / Haze Detection: CASA *vs*. Truth Dataset

The correlation values between the CASA-AWiFS results and the "truth" dataset are as follows:

- Dense clouds: 99.0%
- Light, transparent clouds and haze: 88.3%
- Combined, clouds and haze: 99.4%

As can be seen, CASA-AWiFS statistically correlates with visual cloud estimates very closely. Figure 5 displays a summary of the CASA-AWiFS *vs*. visual assessment differences (error level) for the validation dataset. Overall, CASA is found to be within 5% of actual total cloud cover for 95% of all scenes tested. CASA error level for dense cloud cover is 8% for 100% of all images tested. CASA estimates for light cloud / haze as well as for combined, dense and light, cloud cover, were within 9% for 100% of all scenes (Table 2).

Error Level	Percent of Scenes		
	Dense Cloud Cover	Light Cloud / Haze Cover	Total Cloud / Haze Cover
0 to 1%	46%	52%	48%
0 to 2%	70%	63%	70%
0 to 5%	89%	94%	95%
0 to 10%	100%	100%	100%
Max Error	8%	9%	9%

Table 2: CASA *vs*. visual (truth) estimate of cloud cover: Differences by level of error

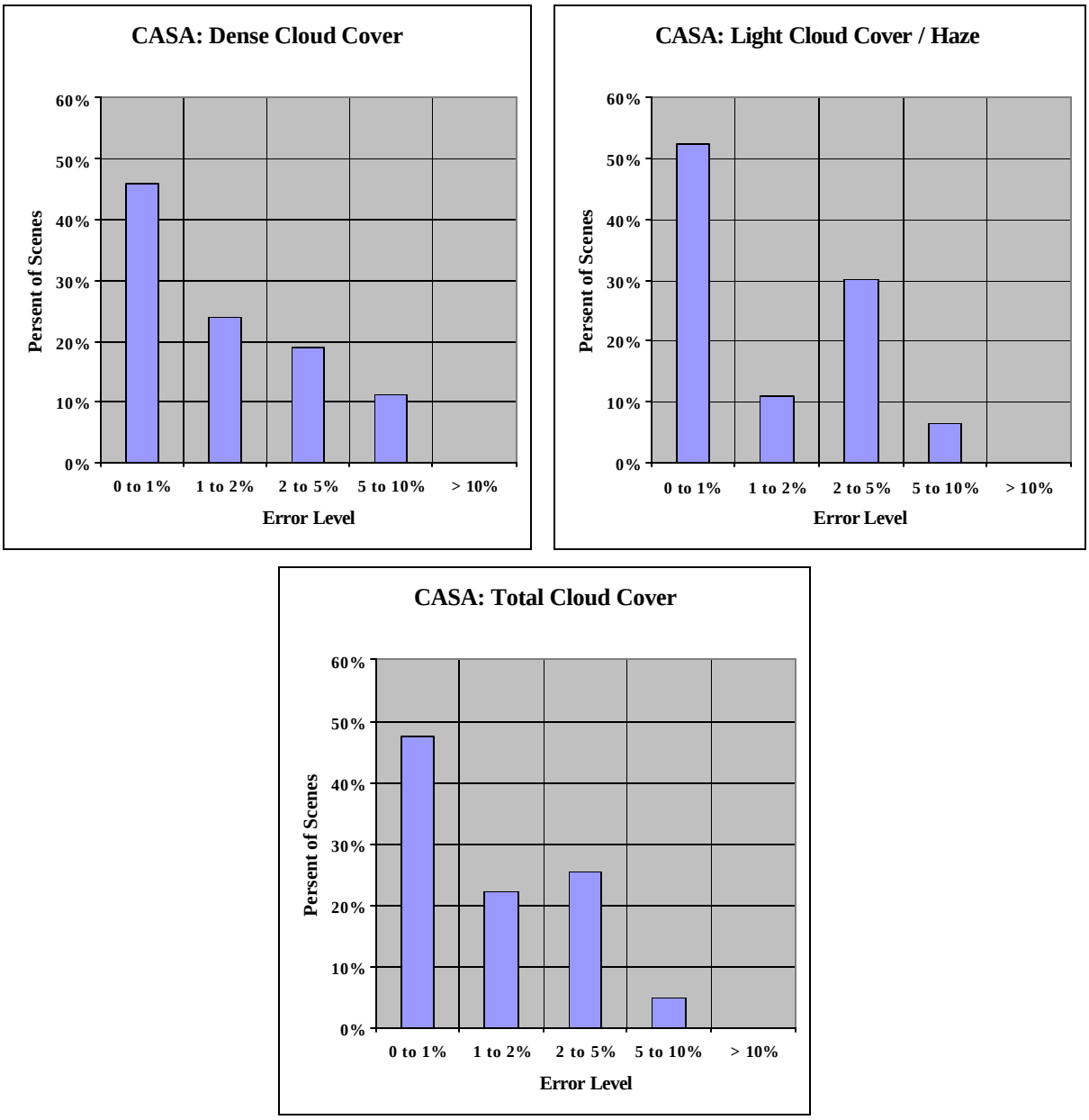

Figure 5: CASA *vs.* visual (truth) estimate of cloud cover: Differences by level of error

The next step of the validation effort moved away from purely statistical, scene-wide correlation of the CASA results against the "truth" dataset. Here, the spatial accuracy of the CASA-AWiFS performance was evaluated by comparing each cloud / cloud shadow mask to the original scene. This comparison involved linking mask and source imagery in a GIS software package and an interactively zooming into areas of interest for close-up assessment of mask quality. Although qualitative, this visual assessment confirmed the high spatial accuracy of the CASA-AWiFS masks. Pictures of all validation scenes and derived CASA masks are presented in Appendix 2.

Direct visual comparison of the CASA-AWiFS masks and the AWiFS scenes illustrates the ability of CASA to differentiate between clouds and other, spectrally similar to cloud, features, such as bare ground, snow, and ice as well as between cloud shadows and water on a consistent basis with a high level of accuracy. A few examples of the CASA-AWiFS performance are presented in Figures 6 and 7. A complete set of CASA results for the validation imagery can be found in Appendix 2.

Figure 6 provides an example of CASA differentiation between clouds and ice and snow on the ground. Spectral confusion between clouds and snow / ice is a known problem in remote sensing. This confusion is particularly severe in the visible bands and can persist in NIR. To resolve the confusion CASA evaluates the reflectance values of the SWIR band and uses this as an extra line of evidence in the separation of clouds from snow / ice.

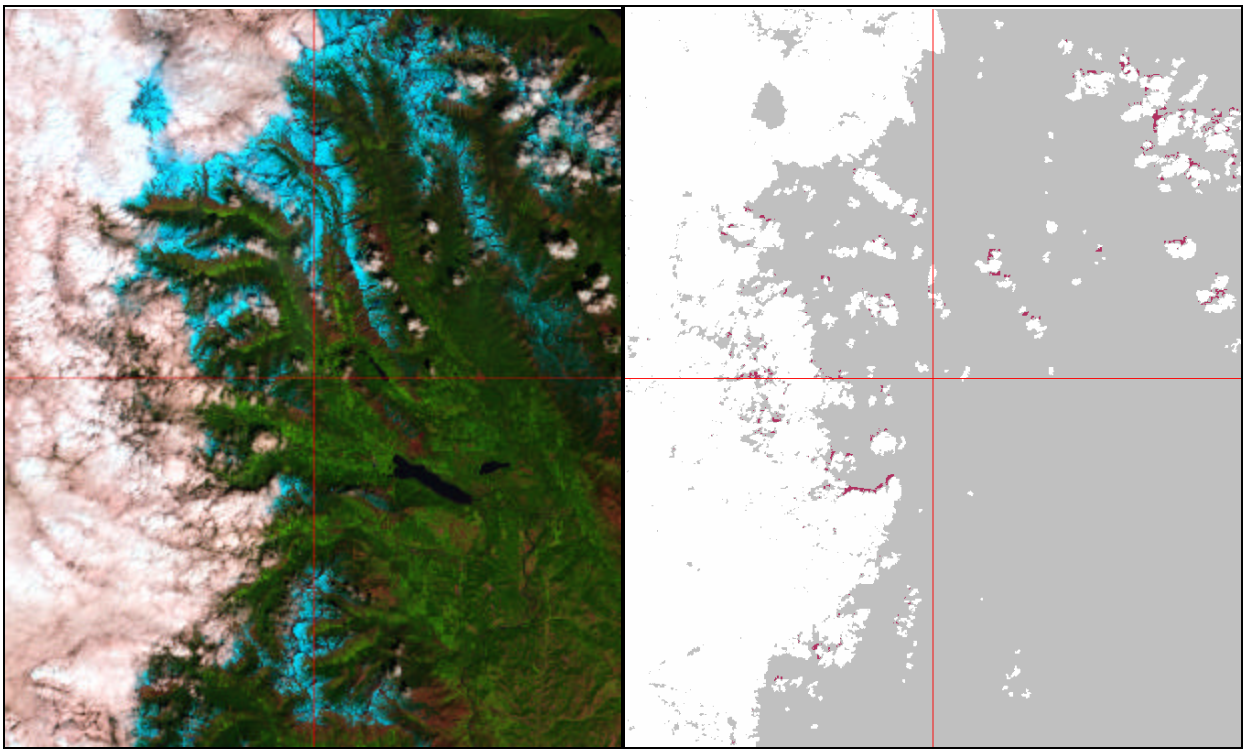

Figure 6: An example of CASA-AWiFS separation of clouds from ground snow and ice. Left: Subset of an AWiFS scene. SWIR-NIR-R band combination. Whitish areas are clouds. Bluish areas are snow and ice. Right: CASA mask. White color represents clouds, brown color represents shadows.

Figure 7 presents an example of CASA mapping of clouds and light clouds / haze in the presence of bright bare ground. It can be seen that CASA is rather successful in identifying the majority of areas of light cloud / haze contamination. Moreover, CASA is able to differentiate between haze and areas of bright background, which are spectrally similar to haze.

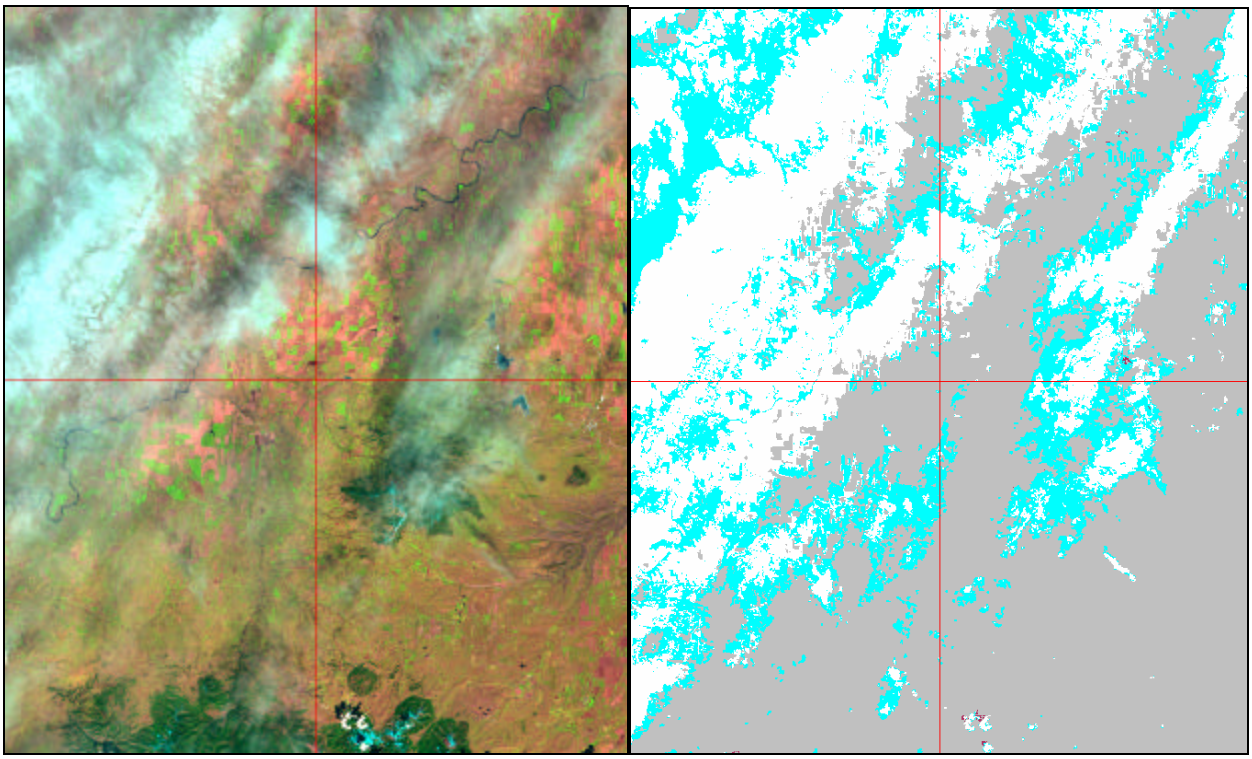

Figure 7: An example of CASA-AWiFS mapping of clouds and light clouds / haze in the presence of a bright background.
Left: Subset of an AWiFS scene. SWIR-NIR-R band combination. Whitish areas are clouds and light clouds / haze. Right: CASA mask. White color represents clouds, bluish color represents light clouds / haze.

For further details please contact:

GDA Corp.
Innovation Park at Penn State University
200 Innovation Blvd., Suite 234
State College, PA 16803
tel: 814-237-4060
fax: 814-689-3375
email: casa@gdacorp.com
web: http://www.gdacorp.com