Mapping Inundation of the Amazon Basin

Anthony Freeman [1], Bruce Chapman [1], Laura Hess [2], Paul Siqueira [1], John Holt [1], Luciano Dutra [3]

[1] Radar Science and Engineering Section, Jet Propulsion Laboratory, NASA
M/S 300-227 4800 Oak Grove Drive, Pasadena, CA 91109 USA
Tel : 1-818-354-3603, Fax : 1-818-393-5285, E-mail : anthony.freeman@jpl.nasa.gov
[2] Institute for Computational Earth System Science, University of California, Santa Barbara CA 93106 USA
[3] Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, SP 12227-010, Brasil

Abstract

In late September 1995 NASDA began a new phase of operations for the JERS-1 SAR - the Global Rain Forest Mapping project. The first rain forest area to be mapped was the Amazon Basin, during October and November of that year (the low flood season for much of the region). This data acquisition was repeated nearly six months later to acquire a second map of the Amazon, during the high flood season in May/June 0f 1996. The main objective of this project is to map inundation over the Amazon basin by comparing data from the high- and low-flood seasons. Most of the data collected during these two phases of this project, a total of ~5000 frames of data, was received and processed by the Alaska SAR Facility, then sent to JPL and NASDA for post-processing and analysis.

The quality of the data processed by ASF for this project is exceptional. It has proved to be very well calibrated, easy to mosaic together, and to have the level of information content required. We have been able to complete a 100 m mosaic from 1500 frames of the entire Amazon basin from the 'low-flood season' and to perform a simple land-cover classification on the resulting 3 GByte data set using a supercomputer at JPL. First results from our analysis of the multi-season data appear to confirm that inundation beneath the canopy can be mapped using this data.

A simple classification scheme was used to classify different forest types (i.e., forest, hill forest, flooded forest), disturbed areas such as clear cuts and urban areas, and river courses in the Amazon basin. The algorithm used is a standard maximum-likelihood classifier, using the radar image local mean and standard deviation of texture as input. Rivers and clear cuts are detected using edge detection and region-growing algorithms. Results on forest/non-forest classifications are reported by geographic region.

1. Introduction

The National Space Development Agency of Japan (NASDA) initiated the Global Rain Forest Mapping Project (GRFM) in 1995. The objective of this project is to use the Japanese Earth Remote Sensing satellite (JERS-1) Synthetic Aperture Radar (SAR) to map the world's tropical rain forest regions at high resolution. This joint project between NASDA's Earth Observation Research Center (EORC), NASA's Jet Propulsion Laboratory (JPL), and the Space Applications Institute of the European Commission's Joint Research Center (JRC/SAI) has assembled a team of invited scientists to evaluate, analyze, and use the data.

The J-ERS-1 satellite travels in a 568 km altitude orbit with a payload that includes an L-band, HH-polarized SAR with a nominal 21m x 21m resolution, which images at incidence angles between 30 and 36 degrees. The J-ERS-1

SAR was the first polar-orbiting imaging radar system capable of monitoring the whole of the Earth's land surface, because of its on-board tape recorder system. Further, because of the ability of imaging radar to see through clouds and at night, and the sensitivity of L-Band backscatter measurements to different biomass levels, the J-ERS-1 SAR is well-suited to multi-temporal studies of the Earth's land surface.

2. Methods and Research Activities

The data used in this study are acquired by the JERS-1 SAR sensor under NASDA's direction. All data for the study was recorded on board the spacecraft and then downlinked for further processing at the Alaska SAR Facility (ASF). ASF processed the data to produce high resolution (~21 m), calibrated data products which were then transferred to JPL, where the data was first reduced, then a secondary radiometric correction applied. In the reduced format data, both mean backscatter (σ^0) and the standard deviation of texture were generated. Backscatter images provided in a full resolution (12.5m by 12.5m pixel spacing) 8 bit amplitude format by ASF were averaged down using 8 x 8 pixel boxes to produce low resolution (100m by 100m pixel spacing) byte amplitude images. The following equation (1) from [3] was used to calculate the standard deviation of the texture:

$$\sigma_T^2 = \frac{N(\sigma_p/\mu_p)^2 - 1}{N+1}\left(1 + \frac{1}{SNR}\right) \qquad (1)$$

where σ_p is the standard deviation of the pixel intensities over some area and μ_p is the mean pixel intensity. While other texture measures could be generated from the high-resolution data, we have found this one to be most useful. In practice, σ_p and μ_p for this study were calculated using 8x8 pixel boxes from full resolution backscatter images so that the resulting texture image pixel size was the same as that of the low resolution backscatter image.

Figure 1. *Area covered during the two Amazon Basin mappings of the GRFM project.*

After data reduction, the low-resolution byte images (2 x 1500 in all) from the low-flood season were mosaiced together both at JPL and at NASDA using two different approaches. A team from the University of California - Santa Barbara (UCSB), working with their colleagues from the Brazilian National Institute for Space Research (INPE) and from the National Institute for Research of the Amazon (INPA), have collected ground validation data from regions near the Amazon River to verify the geometric accuracy of the data and other characteristics.

2.1 Classification

2.1.1 Training Sites

The J-ERS-1 SAR images used for classification were taken from three separate areas of the Amazon rain forest:

(1) Manu National Park, Peru
(2) Manaus, Brazil
(3) Sena Madureira, Brazil.

A plant and wildlife preserve protecting areas of the Western Amazon, Manu National Park is a popular study site due to its biodiversity and pristine conditions. Manaus and Sena Madureira are also important sites studied by scientists worldwide. Data from all of these sites were compiled to form a set of training patterns for the classifier. These patterns then served as prototype vegetation models against which all J-ERS-1 images of the Amazon were compared.

From ground measurements and study of J-ERS-1 SAR images of Manu, two different forest patterns have been differentiated in the Western Amazon, upland forest and floodplain forest. Upland forest occurs in two varieties, the first occurring in hilly regions of Manu north of the Rio Madre de Dios. The hills are from 20-30m high and between 50-150m wide, thereby protecting much of the forest from river flooding. Hilly forests are composed of a mosaic of mature forests with a height ranging from 20-45m. On the hill tops the canopy is closed, while on the slopes and in rivulets between hills palm and bamboo becomes more prevalent. This forest type is also semi-deciduous, which implies that some tree species lose their leaves in the dry season. The second type of upland forest is found on a plateau between the base of the Andes and the Rio Manu. It is similar to the hilly forests, but contains different soil and trees only 20-30m high. The canopy is uneven and much of the upland forest is deciduous, implying that this forest type should experience the greatest amount of leaf loss in the dry season. It is not readily apparent that upland and floodplain forest types found in the Manu area are valid in all parts of Amazon, and from studying images of the Manaus area, hilly regions appear to be distinguishable from the combination upland/floodplain forest type (i.e. the forest type).

In the floodplain forest surrounding the rivers, regular flooding damage occurs and therefore, many stages of forest succession can be found. For example, areas far from the rivers and flooding are the most mature, with a very homogeneous, closed canopy and an average height of 50m. The understory of this mature forest contains a homogenous growth of palms. As one approaches the rivers, however, a less regular forest can be found. Stands of almost pure *Heliconia* banana plants from 2-3m high signal the first stage of regrowth. In areas where severe flooding has killed everything, the forest is very disturbed with a dense cover of *liana* vines. These areas generate a strong L-band response.

Other areas known as Aguajales are characterized by stands of *Mauritia* palms. These palm stands range from very wet to dry and also generate a strong L-band response.

2.2 Classification Method

2.2.1 Mean Backscatter and Texture Image Generation

Both mean backscatter and texture were essential measurements for this classifier. Mean backscatter separated flooded forest, open water, clear-cuts, and urban areas. Texture distinguished more subtle features, such as rivers hidden beneath the rain forest canopy.

2.2.2 River Masking / Image Processing

In J-ERS-1 SAR data open water appears dark, with a very low that is strongly affected by the noise floor. In addition, the texture changes depending upon whether we are in the center of a lake or near a river bank. As a result of these properties, it was found that the best way to classify open water was by applying a region growing operation to the data. In this method, we find regions of connected dark pixels and call them water. Since low vegetation also appears dark, a threshold is applied to distinguish between the two classes.

Our first attempts at classifying J-ERS-1 SAR image data resulted in many false alarms in hilly regions. The backscatter variations associated with the changing slopes of the terrain was often mistaken for water by the classifier. To correct this problem, image processing operations were performed to isolate rivers in hilly regions. This involved convolution with a 17 x 17 Difference of Gaussian edge detecting filter, global thresholding, and region growing of the resulting river areas.

Both of these operations occur before the Bayesian maximum-likelihood classification, creating image masks which denote certain pixels as water or low vegetation. These masks reduce the amount of confusion by effectively removing potentially troublesome pixels through the use of simple rules which are not based on probability measures. The information contained within these masks is ultimately combined with the results of the Bayesian classification through a series of logical rules.

2.2.3 Maximum-likelihood Classification

The classification scheme implemented in this algorithm is based on a supervised Bayesian maximum-likelihood classifier. Essential components of such a classifier are as follows:

(1) A feature vector containing the measurements of interest
(2) Training patterns representing distinct classes
(3) Statistical models of these classes
(4) Decision rules to determine which class the feature vector is most likely to belong.

For classification of J-ERS-1 images of the Amazon, the feature vector was comprised of two components, backscatter and the texture measurement discussed earlier. Each pixel in a J-ERS-1 image to be classified is represented by these two measurements.
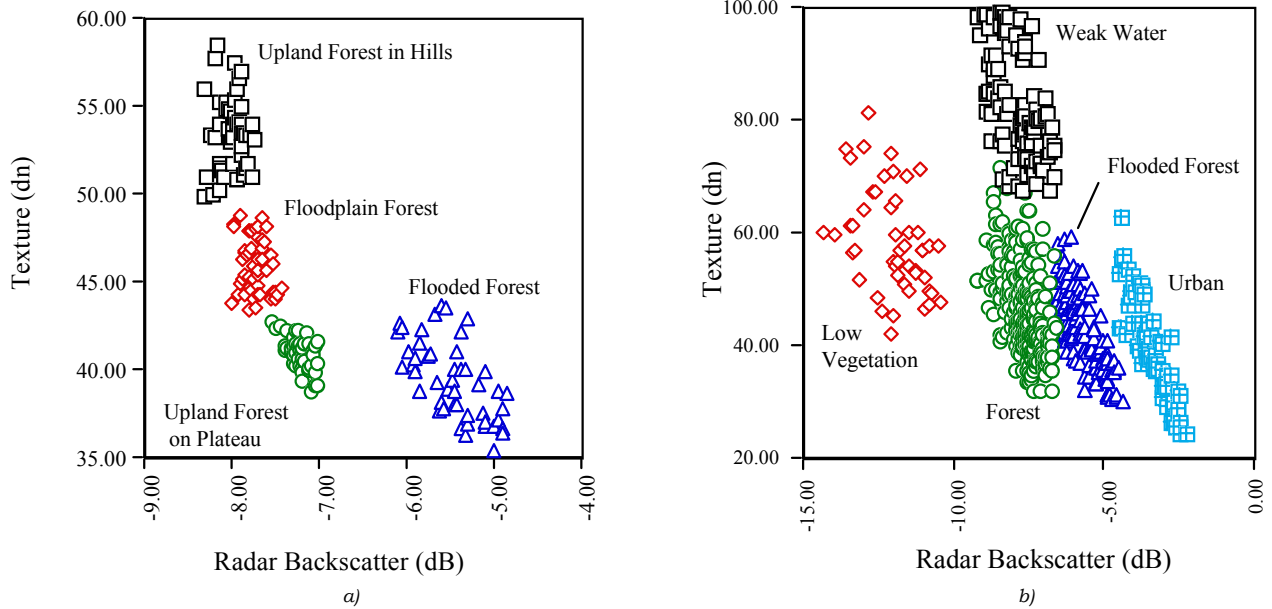
Figure 2. a) : Training patterns created by averaging given pixel.
b) : Training patterns created by 15x15 boxes around a averaging 3x3 boxes around a given pixel

Seven vegetation classes were distinguished for the Manu National Park area:

(1) Upland forest
(2) Floodplain forest
(3) Flooded forest
(4) Low vegetation, such as clear-cut
(5) Water
(6) Urban areas
(7) Unknown.

Once these were chosen, training patterns were created to provide a model of each class. Since no one image contained sufficient quantities of every class, the training patterns were extracted from several images, using box sizes of 3 x 3 or 15 x 15 pixels for the calculations. The patterns for floodplain forest, upland forest, and water were tabulated from data taken of Manu National Park. Flooded forest and urban areas were found in sufficient quantity in images of Manaus, Brazil. Finally, low vegetation in the form of extensive clear-cuts was found in an image of Sena Madureira, Brazil.

Bivariate Gaussian distributions were chosen to model each training pattern. In order to calculate these distributions, the mean, variance, and covariance of both and texture were calculated and are listed in Table I. For the purposes of decision making, the *a priori* probability of each class occurrence was assumed to be equal.

2.2.4 Classifying Hill Forest

Once the Bayesian classification was completed the classifier searched for areas to be classified as hill forest. First, 3 x 3 pixel box averages of backscatter and texture around each pixel were passed through an initial threshold filter to generate a first data set of hill forest candidate pixels. This first pass through the scene served to exclude much of the flooded forest, water and low vegetation using a box size capable of resolving small features. Then, for a second pass,

standard deviations of DN (backscatter image pixel data number) values for 15 x 15 boxes around the hill forest candidate pixels were passed through another thresholding operation generating a second data set containing updated hill forest candidate pixels. This pass eliminated much of the upland and floodplain forest regions. A third pass through the scene involved passing the standard deviations of DN values for 21 x 21 pixel boxes minus any pixels previously classified as flooded forest around the updated candidate pixels through yet another threshold filter in an attempt to eliminate some of the residual hill misclassifications from the previous passes, particularly on the borders of flooded forest regions or rivers with other forest types. The result was the final hill forest classification. Figure 2a), b) suggest that threshold operations can be effective to perform the hill forest classification as described above.

2.2.5 The Final Step

An ambiguity of great concern involved the upland and floodplain forest types. Since the training patterns were created from an image of the Western Amazon where ground observations were made, accurate forest classifications for the Manu images can be expected. Furthermore, it is assumed that the forests from the Western Amazon will generalize throughout the entire rain forest. However, some suspicious forest classifications have appeared throughout scenes near Manaus (e.g., upland forests occurring near flooding rivers while floodplain forest occurs further away). Two factors may be contributing to this problem:

(1) the existing training patterns lack the necessary information to generalize correctly throughout the Amazon
(2) very different forest types exist near Manaus which require entirely new training patterns or rules to correctly distinguish them.

This forest type ambiguity lead to the employment of a final step consisting of grouping the Upland and Floodplain Forest classes into a single forest class. In addition, the brightness of forest and hill forest pixels were made proportional to the backscatter to show features in these regions which could contribute to a better understanding of the nature of the terrain from the final classified image.

2.2.6 Generalizing the Classifier

In creating a classifier which operates on only one or two images, it is conceivable to account for most image oddities directly within the context of the classifier code. However, when faced with the prospect of producing vegetation maps for hundreds or thousands of images, there must be some way for the algorithm to adjust to a wide variety of circumstances. In our classifier, this is handled by a set of user-controlled options. At the present time, the following options are in place:

(1) Toggle on or off averaging filters which smooth the forest and low vegetation classifications.
(2) Permit pixels to be classified as urban. This is necessary since highly disturbed forest may look identical to some urban areas.
(3) Permit pixels to be classified as hill forest. Selecting this option only for scenes with hills present will avoid hill forest misclassification.
(4) Allow the user to manually set the threshold between water and low vegetation.
(5) Allow the user to exert greater control over the J-ERS-1 image correction process mentioned earlier.

3. Results

Figure 3 shows a small segment of the multi-season data set after the data from the two dates has been co-registered. This data was then used in a supervised classification (maximum likelihood), the result of which is also shown in the figure. Table I shows classification accuracies for the single-season data used separately, and the significant improvement obtained when the combined high-flood and low-flood data are used. Work is currently in progress to optimize the estimates of inundation extent from the multi-season data sets. The mosaic of the second (high flood) season data should be completed later this year.

Table I. *Classification accuracies for supervised classification of multi-season GRFM data for the area shown in Figure 3.*

Class	High + Low Flood	High Flood only	Low Flood only
Water	99.2	53.9	97.4
Forest	93.2	64.7	61.6
Flooded Forest	96.4	96.8	50.2
Flooded Shrubs	66	18	40
Clear Cut 1	80.6	50	22.2
Clear Cut 2	86.4	74.2	74.2
Sandbar	77.9	57.9	18.6
Total	92.2	65.9	63.8

JERS-1 HIGH/LOW FLOOD DATA

🟥 High Flood Data (7/93)
🟩 Low Flood Data (10/95)
🟦 Ratio High/Low Flood

INUNDATION MAP

🟥 FLOODED FOREST   🟨 UNFLOODED FIELDS
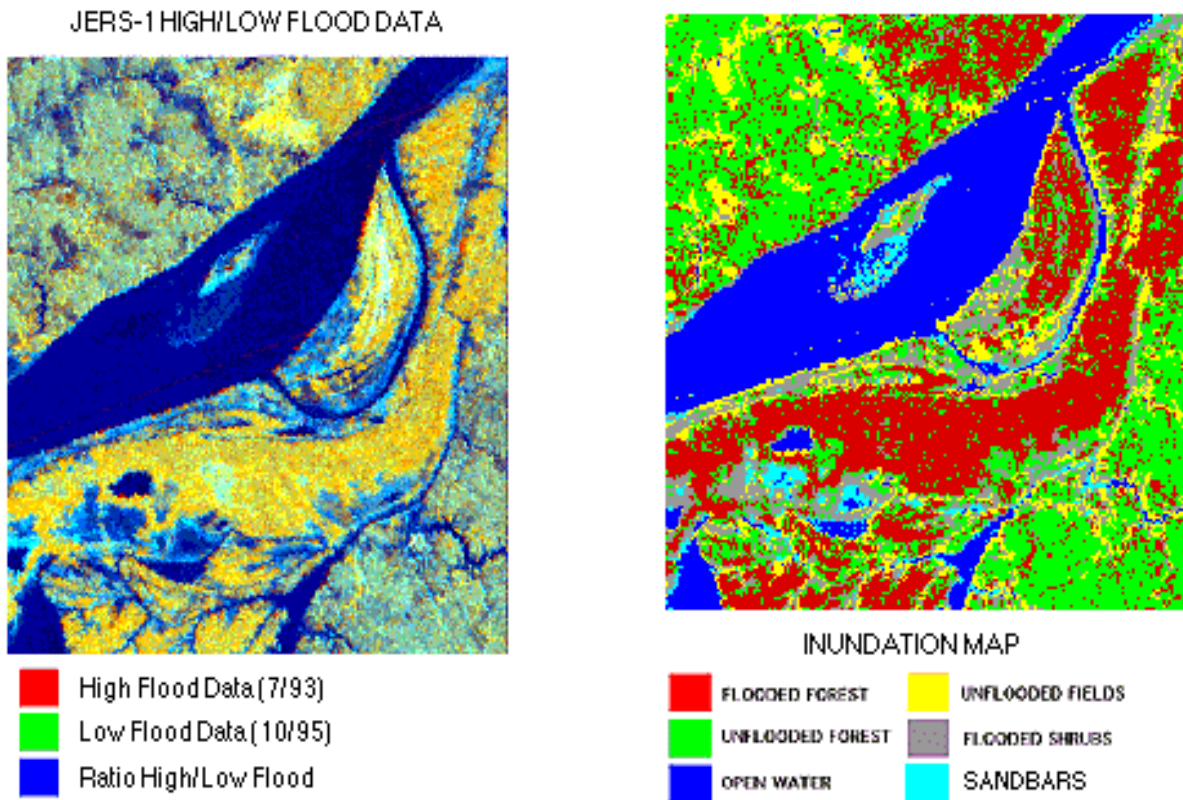🟩 UNFLOODED FOREST   ▦ FLOODED SHRUBS
🟦 OPEN WATER   🟦 SANDBARS

Figure 3. *Example of multi-seasonal data and resulting supervised classification for an area south-west of Manaus, Brazil.*

3.1 Classification Results

The classifier was first tested on the Manu National Park training image. Several noteworthy features were immediately apparent in the vegetation map. First, irregular yellow spots represented areas of bright L-band response, such as palm stands, inundated areas, and regions of highly disturbed forest. Several small rivers which were barely discernible in the original image showed a high texture, and therefore were easily distinguished by the classifier. Three examples are the Rio Blanco which empties from the southwest into the Rio Madre de Dios, and the Rio Los Amigos and Rio El Amiguillo in the hills north of the Rio Madre de Dios. The floodplain forest can be seen surrounding the Rio Madre de Dios and the small rivers in the hills.

Looking at a vegetation map of a J-ERS-1 SAR image taken near Manaus, Brazil, in July 1993, we saw additional features that can be discerned by the classifier. The effectiveness of the river masking procedure was evident, as displayed by the algorithm's ability to accurately classify large expanses of open water as well as follow narrow, convoluted waterways and inlets. Large tracts of flooded forest were detected in addition to inundation effects near small rivers. Clear cutting was also prevalent, and some areas of low vegetation were detected near river boundaries.

Not everything was perfect, however. Rather large occurrences of the Unknown class appeared at the boundaries between water and flooded forest. This can be attributed to the classifier attempting to average two very dissimilar pixel types and getting caught somewhere in between the two vegetation classes. Some of the clear-cuts also contained erroneous spots of water. These misclassifications were the result of darker than usual clear-cuts combined with an imperfect thresholding operation during the river masking procedure. Some of this can be remedied with the user options, but often at the expense of smaller waterways being mislabeled as low vegetation. There is just not enough separation between these classes in the data.

Regarding the hill forest classification, particularly troublesome areas appeared at the boundaries of flooded forest with forest classes. Some misclassification may remain in the final product, however since the brightness of the hill forest class is proportional to the backscatter, it is possible in many cases to pick out this form of misclassification by visual inspection.

An upper bound on the classifier performance (hill forest and urban regions not included) was estimated by calculating the classification accuracy in areas of the Manu training image for which ground truth was available. In Table II, the "# ground truth pixels" column represents the number of pixels known, from the ground truth data, to belong to a specific class while the "# matching pixels" column shows how many of those pixels were actually assigned to the correct class in the final classification image. The "upper bound on performance" column is simply the percentage of correctly classified pixels. The water measurements were made on small rivers hidden under the forest canopy, so for expanses of open water the performance of the classifier is expected to be better than what is shown here.

Perhaps more interesting than looking at the classification of one scene is comparing a sequence of vegetation maps created from data taken over an entire year. As the season shifted from dry to rainy, a definite seasonal effect was observed. Manu images taken in March showed a few large flooded forest regions that did not appear elsewhere throughout the year. These spots were regions in the original image which generated high L-band responses and only appeared flooded during the height of the wet season. More likely than not, these were spots that had experienced recent flooding damage or were sensitive palms which grew leaves only during the wettest months.

Table II. Upper bound on the classifier performance.

Class	# ground truth pixels	# matching pixels	upper bound on performance
Forest	300	287	96%
Flooded Forest	150	143	95%
Low Vegetation	50	49	98%
Water	100	82	82%

4. Conclusions

An algorithm has been presented that classifies vegetation types present in J-ERS-1 images of the Amazon Rain Forest. Radar backscatter and the standard deviation of texture are measurements used to distinguish the vegetation types. At the heart of the algorithm is a supervised, Bayesian maximum-likelihood classification capable of differentiating the following classes: Forest, Flooded Forest, Low Vegetation (clear-cut areas), Urban, and Water. A series of thresholding operations are employed to differentiate a class denominated Hill Forest. Image processing techniques such as edge-detection, thresholding, and region growing are implemented to reduce confusion by masking some water and low vegetation pixels before the Bayesian classification is actually run. An image scaling operation is also performed to reduce the effect of cross-track calibration errors. Realizing the difficulty in trying to create a generalized algorithm which can deal with hundreds of images, a set of user options to tailor the classification rules is under continual development.

Results indicate an excellent ability to classify water in all forms, including open water and rivers hidden beneath the rain forest canopy. Flooded forest, low vegetation, hill forest and urban areas are also distinguished with only occasional confusion. Upland and floodplain forest are classified accurately in images of Manu, but there appears to be some difficulty in generalizing these two forest classes beyond the Western Amazon, so they have been combined into a single forest class.

The ultimate goals of this classifier are to track deforestation, detect seasonal effects such as flooding, and accurately map the forest types throughout the Amazon. So far, we are well on our way to achieving all three goals. Clear cuts have been accurately detected in images of Manaus and Sena Madureira, and our vegetation maps have shown the ebb and flow of flood waters in seasonal sequences taken over Manu and Manaus. Evaluation of the classifiers performance could be enhanced by increased availability of ground truth data, however the results are very promising thus far, with our algorithm capable of classifying to a high degree of accuracy images ranging from the Western Amazon to Brazil.

5. References

[1] Rosenqvist, A., 1996, The global rain forest mapping project by JERS-1 SAR, Proc of XVIII ISPRS Cong. In Vienna, Austria, 1996.

[2] Freeman, A, Chapman, B., and Alves, M.,1996, The JERS-1 Amazon Multi-mission Mapping Study (JAMMS), Proc of 1996 Intnl. Geosci. Rem. Sens., pp830-833, Lincoln, Nebraska, 1996.

[3] P. H. Swain, Remote Sensing, the Quantitative Approach. United States of America : McGraw-Hill, Inc., 1978.

[4] Shimada, M., "J-ERS-1 SAR Calibration and Validation", presented at the CEOS SAR Calibration Workshop, Noordwijk, The Netherlands, September 1993.

[5] Hara., Y. and Ono, M., "Analysis of J-ERS-1 SAR Imagery", Proceedings IGARSS '93, pp. 1191-1193.

[6] Yamagata, Y. and Yasuoka, Y., "Classification of Wetland Vegetation by Texture Analysis Methods Using ERS-1 and J-ERS-1 Images", Proceedings IGARSS '93, pp. 1614-1616.

[7] Freeman, A., Alves, M. and Williams, J. "Calibration Results for J-ERS-1 SAR Data produced by the Alaska SAR Facility", Proceedings IGARSS '93, pp. 965-967.

[8] B. Chapman, M. Alves, A. Freeman, Validation and Calibration of J-ERS-1 SAR Imagery, Proceedings of the J-ERS-1 Results Reporting Meeting, 1995.