REMOTE SENSING TECHNIQUES TO THE DETECTION AND MAPPING OF FLOODING DYNAMICS WITHIN THE PANTANAL, MATO GROSSO DO SUL STATE, BRAZIL: PRELIMINARY RESULTS

João dos Santos Vila da Silva
Hermann J.H. Kux

INPE
SÃO JOSÉ DOS CAMPOS
Junho de 1991
REMOTE SENSING TECHNIQUES TO THE DETECTION AND MAPPING OF FLOODING DYNAMICS WITHIN THE PANTANAL, MATO GROSSO DO SUL STATE, BRAZIL: PRELIMINARY RESULTS

João dos Santos Villa da Silva
Hermann J.H. Kux

REMOTE SENSING TECHNIQUES TO THE DETECTION AND MAPPING OF FLOODING DYNAMICS WITHIN THE PANTANAL, MATO GROSSO DO SUL STATE, BRAZIL: PRELIMINARY RESULTS*

João dos Santos Vila da Silva
EMBRAPA/CPAP - Centro de Pesquisa Agropecuária do Pantanal
Corumbá, MS, Brasil

Hermann J. H. Kux
INPE - Instituto Nacional de Pesquisas Espaciais
Caixa Postal 515 - 12201 - São José dos Campos, SP, Brasil

ABSTRACT

The Pantanal, an extensive floodplain region in Central-Western Brazil, is annually flooded. TM-scenes (bands 3, 4 and 5) from the dry and flood seasons, were used in this study. Its objective is to evaluate the performance of classification algorithms (Maximum Likelihood, K-means, Parallelepiped, Slicing) in order to discriminate four thematic classes (water, flooded areas, wet areas and dry areas) of interest. An analysis from the results of these classifications is made. The classifier on Maximum Likelihood (MAXVER) showed that in the scene of the dry season, the results were the most consistent, followed by the Density Slicing classifier at band. This study concludes with recommendations for further works on flooding dynamics within this huge ecosystem.

1.0 INTRODUCTION

The Pantanal, located in Central-Western Brazil, is one of those - on a worldwide basis - few large floodplains, where inundations are not catastrophic events, but temporary and recurrent phenomena on a yearly basis, that are necessary for man and environment.

This floodplain is drained by numerous rivers, brooks and creeks which, during the rainy season, determine periodical flooding variable in intensity, extension and water residence time. Within this ecosystem where agricultural land use is almost impossible, extensive cattle raising based mostly on native grasses, is the main economic activity. Therefore, the detection and mapping of flooded areas is a relevant task of interest to the management of cattle raising and wildlife in the Pantanal.

Data obtained by TM and MSS-Landsat have been utilized to evaluate the extent of inundated areas in several countries, using both visual and digital classification techniques. Those techniques on superposition of images by multispectral, multitemporal (or both) color composites were used, among others, by Deutsch & Ruggles (1974), Kruss et al. (1981), Pinto et al. (1985), Pinto & Florenzano (1988) and Florenzano et al. (1990). This allowed the identification of changes that occurred in the area of interest before and after inundation. Other techniques, such as linear contrast stretch, ratio between bands, conventional and nonconventional filtering can also help in these classifications (Darch, 1979; Philipson & Hafker, 1981; Rose & Rosendahl, 1982).

In order to perform these studies, the following algorithms, among others, have been used for digital supervised and nonsupervised classification: Maximum Likelihood, K-means, Slicing and Parallelepiped.

Among the few studies using remote sensing techniques in the Pantanal, the following should be mentioned: Darch (1979) analysed the moisture status and drainage conditions of the Pantanal by using MSS images from bands 4, 5 and 7, for visual and digital classification. Two test sites, respectively to the North and Southeast of Corumbá, were selected. Color images were obtained by the projection of the following combination of filters over B&W diapositives: WRATTEN gelatin filters 47B (blue), 58 (green) and 29 (red) for the bands 4, 5 and 7, respectively. The results indicate that the combination of bands 4 and 7 is most adequate to discriminate between dry and wet areas. The combination of band 4 with a red filter and of band 7 with a blue filter allowed an excellent discrimination between dry (color: pink/purple) and wet (color: orange/red) areas. This study shows that the combination of nonconventional filters with MSS-bands is helpful for information extraction and that the use of conventional color filters could increase the difficulty to discriminate targets of interest.

Rooy (1982), using B&W transparencies of MSS-data at 1:500.000, mapped visually flooded areas during six different periods. He recommended to use bands 5 and 7 for the determination of these areas within the Pantanal. Furthermore, he emphasized that the images present several reflectance levels, defining a sole type of surface (water, wetlands, dry land, open vegetation, etc) or a combination where one type may predominate or not. According to the same author it is theoretically possible to discriminate between dry and flooded areas. Nevertheless, in practical terms this discrimination is frequently quite difficult due to a variety of possible water-soil-vegetation interactions.

Silva (1985), using color infrared aerial photographs,
identified four vegetation types. Utilizing MSS band 5 and 7 digital data, he classified three vegetation types due to similar spectral responses of semideciduous aluvial forests and transition vegetation. As far as the delimitation of open water is concerned, flooded areas did not show a separability among water and other land cover classes and, owing to the limitation of spatial resolution (80m), water bodies with small dimensions were not registered. The author recommends to use TM sensor data, since its higher spatial resolution of 30m would permit a more detailed analysis on the variations of the open water surface.

Kux et al. (1986) evaluated the applicability of MOMS data, band 1, for hydrological studies, as compared to TM bands 3 and 4, in a section of the southern Pantanal. The TM/MOMS classification matrix presented a high percentage of inclusion errors among the dense vegetation classes (specially the Carandá-palm formations) and moist areas (I and II). The soil moisture condition was regarded as the responsible factor for the confusion.

Ponzoni et al. (1989), performed the physiognomic mapping of the vegetation from the Pantanal National Park by using both TM images and 35mm aerial photographs. They evaluated the influence of the variations of the water level on the physiognomies identified. They concluded that the data takes from the dry season would permit a higher contrast of vegetation features. These authors found out a high visual correlation among band TM-5 and different moisture levels. Furthermore this band is of fundamental importance to identify the physiognomic vegetation units during the dry season. As for the flood season, this band presented low contrast with the predominance of dark tones, indicating the presence of water in the whole area. The same happened with band 3, that did not present different tones of water surfaces because, during this season, the waterflows and lakes within this region presented high transparency. TM-4 band is recommended for this period since it permits a better discrimination between vegetation and moisture classes.

Probably due to the reduced number of studies and to the complexity of this region, there is no consensus as to the best procedure to be followed with remote sensing techniques. The objective of this study is to evaluate the performance of the classification algorithms Maximum Likelihood, K-means, Slicing and Parallelepiped, implemented at INPE's image analysing system (SITIM 150) to identify the flooded sections from a test site within the Southern Pantanal.
2.0 MATERIALS AND METHODS

A test site (Figure 1) with an area of 216.5 km² in the municipality of Corumbá, Mato Grosso do Sul State (Geographical coordinates: Lat. 19°28'49"S, Long. 57°18'19"W and Lat. 19°37'29"S, Long. 57°09'09"W) with characteristic cattle raising activities, was selected. The native vegetation consists mainly of grass and herbs, open and closed woods and bushes.

Two TM data takes, bands 3 (visible), 4 and 5 (both infrared) from orbit 226, point 74, quadrant A were used. The first data take (May 11th, 1989) corresponds to the flood season, whereas the second (September 16th, 1989) to the dry season.

Initially the digital numbers (DN) of both datasets were transformed into apparent reflectance and afterwards into reflectance imagery according to the equations proposed by Eosat (1986) and Markham & Barker (1987).

This transformation is necessary mainly for comparisons of images taken at different times (Bowker et al. 1985; Epiphanio & Formaggio, 1988; Ahern & Sirois, 1989). After this procedure the images were registered to the corresponding topographic map at 1:100,000. During the first data take a field check was made along rio Miranda and by road.

The classifications obtained using algorithms MAXVER (Maximum Likelihood), K-means, Slicing, Parallelepiped were qualitatively evaluated. The following grades were assigned: G (good), F (fair), P (poor) according to the coherence of the classification with ground truth. Following, each algorithm is briefly described below:

- MAXVER is a supervised classifier and the user must have a previous knowledge of the area to be classified and obtain training samples so that the classifier can learn the spectral characteristics of the thematic classes. It is based on stochastic models that are formalized within the statistical decision theory. It permits to associate a priori each class to a probability density function. Based on the Bayes rule, it provides the error probability when classifying a vector in a certain class, but that belongs to another one, i.e. it presents an inclusion or omission error.
- Slicing is also a supervised classification method. Its function is to divide the histogram of a band - defined by the area of the cursor or the theme - into gray level intervals, each point of the image being associated to these intervals. It can be executed in the manual, equidistribution and normal mode. In this case, the manual mode was used, where the user determines the lower and the upper limits of the DNs belonging to each class to be mapped.
- The Parallelepiped is a supervised classifier whose premise is
Figure 1 - Location of test site within Pantanal.
to establish an interval of maximum digital values for each class and each band, using training areas indicated by the user. An unknown part of the image is classified as belonging to a known class if its digital value is within this interval.

- K-means is a nonsupervised classifier where the user identifies the most representative areas of the scene under study, and establishes the K number of classes and its initial centers. The spectral data of the image are clustered around these centers and each point is associated to the closest class, i.e. to a smaller distance: the Euclidean distance.

Further information on these classifiers can be found in Duda & Hart (1973), Mather (1987) and Mascarenhas & Velasco (1989).

3.0 RESULTS AND DISCUSSION

3.1. EVALUATION OF TM - LANDSAT BANDS

At band 3 the flooded sections presented dark color, in contrast to higher nonflooded areas. In those areas where green leaves occur in great quantity, such as in the gallery forest, the tones are similar to those of flooded areas due to a high chlorophyll absorption.

Band 4 permits a high separability between vegetation and water, since vegetation reflects much in this spectral band, whereas water absorbs. This indicates that it is the most adequate band for the study proposed, since the reductions in reflectance (dark gray tones in the image) are due to open water. Nevertheless, this separability is extremely difficult because vegetation presents different densities where radiation does not reach the bottom.

At band 5 radiation is absorbed primarily by foliar water, followed by open water. At this band it is not possible to discriminate between flooded or wet areas and dry areas, when the former ones were caused only by the existence of open water, since it is not possible to identify where energy absorption is taking place. However, the reflectance of vegetation is very high. So, the clear gray tones in the image are mostly grasses and bushes, without inundation.

At the color composite of bands 3(blue), 4(red) and 5(green), the darker tones with lesser brightness and uniform texture were associated to open water (rivers, creeks, lakes). Intermediate dark tones (brown/black) were associated to totally flooded areas, with herbaceous vegetation, bushes and tall to medium woods. The orange color was associated to dense woods (gallery forest and "caapões", i.e. forest residuals on higher spots), without discriminating between dry and wet. The green color, with clear tones, was associated to dry areas, made up of herbs and
bushes. Notwithstanding, this scheme is not totally trustworthy, since areas that certainly were flooded also present this color, specially where high pasture grasses (*Paspalum hidrophilum* spp.) and aquatic macrophytes occur. The green color with dark tones was associated to humid areas.

3.2. **DESCRIPTION OF THE CLASSES MAPPED**

1 - **Water**: refers to the open water surface of channels (rivers, brooks and creeks) and lakes. Within this class, an underestimation of area may have occurred, because due to a proliferation of macrophytes the water surface is frequently hidden underneath. Ponzoni et al. (1989) frequently found this problem in areas with occurrence of macrophytes.

2 - **Flooded area**: refers to sections where the soil is completely covered by water surface. There is always occurrence of vegetation.

3 - **Humid area**: refers to those areas where it is not possible to infer whether there is water below the foliage or not. A strong indicator of moisture is a lighter coloration than class 2 above, what made possible this differentiation. Vegetation is also present here.

4 - **Dry area**: refers to sections where no water was detected. It is covered by herbs, bushes and trees, besides the roads that were built above the yearly inundation level.

It is worth informing that residuals of woods ("caapões"), identifiable at TM-images, are frequently associated to small elevations of 1-2m typical for the Pantanal, and can be aggregated to class 4. These sections are very important for this region, because they are refuges for wildlife and for bovines during the peak of flooding since they are hardly reached by inundations.

Some sections within both the wet and dry areas, used as natural pasture for cattle raising, present at TM-images different darker and lighter tones, separated only by a straight line, namely a fence! This is due to different pasture management, and acts as a further confusion factor, specially for the interpretation of the scene from the flood season.

3.3. **EVALUATION OF CLASSIFICATION RESULTS**

Taking into account the four themes, 12 classifications were performed, 5 and 7 for the dry and flood season respectively. Table 1 presents the results from these classifications. As far as the number four above is concerned, 4 classes can be observed
### TABLE 1: RESULTS FROM DIGITAL CLASSIFICATIONS (AREA IN km²) FROM DRY AND FLOOD SEASON

#### 1 - THEMATIC CLASSES: DRY SEASON

<table>
<thead>
<tr>
<th>CLASSIFIERS</th>
<th>Water area</th>
<th>flooded area</th>
<th>humid area</th>
<th>dry area I</th>
<th>dry area II</th>
<th>woods area</th>
<th>non classified</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maxvar, 3 bands</td>
<td>1.4(G)</td>
<td>10.1(G)</td>
<td>39.9(G)</td>
<td>156.2(G)</td>
<td>-</td>
<td>-</td>
<td>29.0</td>
</tr>
<tr>
<td>4 classes</td>
<td>1.3(G)</td>
<td>8.1(G)</td>
<td>32.0(F)</td>
<td>175.2(G)</td>
<td>-</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>Slicing band 4</td>
<td>1.7(G)</td>
<td>20.4(P)</td>
<td>18.3(P)</td>
<td>127.9(F)</td>
<td>-</td>
<td>-</td>
<td>23.8(G) 24.4</td>
</tr>
<tr>
<td>4 classes</td>
<td>17.8(P)</td>
<td>-</td>
<td>63.8(P)</td>
<td>104.2(P)</td>
<td>30.8(P)</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>Parallelepiped, 3 bands</td>
<td>9.5(P)</td>
<td>-</td>
<td>35.7(P)</td>
<td>70.7(P)</td>
<td>52.6(P)</td>
<td>16.3(P)</td>
<td>31(P) 0</td>
</tr>
<tr>
<td>6 classes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### 2 - THEMATIC CLASSES: FLOOD SEASON

<table>
<thead>
<tr>
<th>CLASSIFIERS</th>
<th>Water area</th>
<th>flooded area I</th>
<th>flooded area II</th>
<th>humid area</th>
<th>dry area</th>
<th>woods area</th>
<th>non classified</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maxvar, 3 bands</td>
<td>4.4(P)</td>
<td>13.3(P)</td>
<td>-</td>
<td>29.4(P)</td>
<td>18.7(P)</td>
<td>-</td>
<td>31.0</td>
</tr>
<tr>
<td>4 classes</td>
<td>5.6(F)</td>
<td>72.5(G)</td>
<td>25.2(G)</td>
<td>40.4(G)</td>
<td>11.2(P)</td>
<td>24.9(G)</td>
<td>37.8</td>
</tr>
<tr>
<td>Slicing band 5</td>
<td>14.2(G)</td>
<td>77.0(P)</td>
<td>-</td>
<td>93.4(P)</td>
<td>32.0(P)</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>4 classes</td>
<td>4.5(P)</td>
<td>74.8(P)</td>
<td>-</td>
<td>73.0(P)</td>
<td>45.8(P)</td>
<td>16.5(F)</td>
<td>0</td>
</tr>
<tr>
<td>Slicing band 4</td>
<td>14.6(G)</td>
<td>134.4(P)</td>
<td>-</td>
<td>7.4(P)</td>
<td>14.7(F)</td>
<td>45.4</td>
<td></td>
</tr>
<tr>
<td>Parallelepiped, 3 bands</td>
<td>119.8(P)</td>
<td>80.1(P)</td>
<td>-</td>
<td>15.5(P)</td>
<td>1.1(P)</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>4 classes</td>
<td>28.2(F)</td>
<td>56.2(P)</td>
<td>48.1(F)</td>
<td>54.0(P)</td>
<td>1.1(P)</td>
<td>29.2(G)</td>
<td>0</td>
</tr>
<tr>
<td>K-means, 3 bands</td>
<td>4 classes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 classes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Ø) = Good, (F) = Fair, (P) = Poor
in some classifications, due to the great spectral contrast found in the field. These themes at the end of classification were summed up to its corresponding class. The evaluations were made considering the two distinct periods of the hydrologic cycle of this region.

**Dry season** - As for the scene from this season, the spectral response defines best terrestrial targets, with little interference from water. So, only five classifications were needed. The MAXVER algorithm classified 87% of the trained area for 4 classes. The dry area was underestimated and included as non-classified.

The confusion between inundated and wet areas is almost imperceptible, and doesn't interfere in the result. It was not necessary to increase the number of classes because the non-classified sections can be included perfectly in the class "dry area". Altogether the result was considered good. Following, and with the same number of themes, comes the Slicing algorithm of band 4. This band allowed a better separability of the themes of the scene from this season, due to a good delineation of features in the image. Some confusion arose with the inclusion of classes "woods" and "humid area" close to creeks and flooded areas around the main road BR-262, into class "dry area". Notwithstanding this classifier can be used successfully during this season, specially when the speed of classification is considered. The same doesn't account for the Parallelepiped classifier, since, besides the need to increase the number of classes (woods), it obtained a much lower performance than expected. The discrimination between themes was not good, causing confusion among flooded, humid and dry areas. The K-means classifier had the worst performance of all. Even when increasing the number of centers (6 classes), the discrimination among them did not improve. So, for example, it overestimated the class "water" and confounded it with "flooded area". This classifier is not recommended for the study.

**Flood season** - The digital classification of data from this season is extremely difficult, as already emphasized by Rose & Rosendahl (1982) and Ponzoni et al. (1989), due to the different types of interactions between water and vegetation cover, hiding the spectral responses. In order to improve the performance of the classifications, the number of themes was increased in the MAXVER and Slicing classifications.

The classifiers MAXVER and K-means, both with 4 classes, and Slicing at band 4 with 5 classes had a poor performance (Table 1). The first underestimated the class water and confounded dry area with woods (which were flooded) and with humid area. The second included 92% of the area within the themes "water" and "flooded area", highly overestimating water, while regarding just
0.5% as a dry area and the remains as humid area. The classifiers with 4 classes (Slicing at band 5 and Parallelepiped) did not a good discrimination among the themes, but were the best to separate the class water. However, the confusion rate was high for all other classes.

To perform Slicing at band 4, it was necessary to include the class "woods", because it presents the highest DNs, suggesting to classify it as dry area, but it is flooded. Nevertheless this doesn't occur at band 5, where the highest DNs belong to the dry area. Contrarily to what is stated in the literature, band 4 was not the best one to discriminate water from other classes during this season.

The classifier K-means (6 classes) also had a poor performance. The class dry area was underestimated and included partially within classes "woods" and "humid area", whereas "water" included part of the "flooded area".

In spite of classifying only 83% of the area, MAXVER was also the most adequate for the flood season, although some minor omission errors occurred such as non-classification of roads, landing strips, and some dry farmyard of Fazenda Acurizal. Also smaller features of water and flooded areas, or the inclusion of areas covered by vegetation (woods, pasture and macrophytes) above the water surface within the class "dry area", are some of the problems that can be solved in a classification when increasing the thematic classes. This is only possible when excellent ground truth information is available.

4.0 FINAL REMARKS AND RECOMMENDATIONS

This work is the first phase of a study aiming to evaluate the dynamics of flooding in the Pantanal using, besides conventional hydrological techniques (evaluation of the height of hydrographs along rivers), advanced remote sensing and data processing techniques (GIS-Geographic Information Systems).

In order to perform successfully studies on extremely dynamic phenomena such as the flooding in the Pantanal, there is a demand for remote sensing data in a high temporal frequency and for information that would permit to define precisely, i.e. with the least possible error, the limits of dry, wet and flooded areas.

The demand for high temporal frequency of remote sensing data could be solved in part with the utilization of NOAA/AVHRR images. The studies then would be conducted at small scales (1:1,000,000), but with daily information encompassing the entire Pantanal.

As far as more detailed information from remote sensing data
is concerned, it is the intention of the authors, to use geocoded SAR scenes from the new orbital system ESA/ERS-1, merged with TM scenes to further improve thematic classifications and also to enlight the quite complex relationship among vegetation-soils-flooding levels within this ecosystem.

5.0 REFERENCES


EDIÇÃO SOBRE "New Look-up Tables".


SILVA, D. [1985]. Mapeamento das formações vegetais e da variação da lâmina d'água em parte do Parque Nacional do