INTEGRATION OF REMOTE SENSING, GEOGRAPHIC INFORMATION SYSTEMS AND DATABASES TO IDENTIFY IRRIGATED WINTER CROPS

Manoel Jimenez Ortiz
Antonio Roberto Formaggio
José Carlos Neves Epiphanio

Instituto Nacional de Pesquisas Espaciais - São José dos Campos (SP)
formag@iltid.inpe.br

COMMISSION VII, WORKING GROUP 6

KEYWORDS: Geoprocessing, Agriculture, Classification

ABSTRACT

This work presents a methodology to identify irrigated winter crops, using a multitemporal dataset of imageries and ground auxiliary data referring to three consecutive years (1988, 89 and 90). An image analysis system (SITIM), developed by INPE, and a GIS as well as a Database Management System (dBASE IV) were used to make the integration of these "multifonts" and multitemporal data, to identify the winter crops. The area under study was digitally classified, using a maximum likelihood algorithm, in order to evaluate the results obtained with the proposed methodology. Afterwards both results were compared using Kappa statistics. The indexes to both the proposed methodology and to the digital classification method were 0.668974 (very good) and 0.472170 (good), respectively. The values obtained allowed the conclusion that, using the methodology proposed for areal classification, an improvement of 28.93% over the digital classification was achieved. Furthermore, along this study those problems related to multifont data integration were identified, suggesting the use of a reliable system for geoprocessing.

1. INTRODUCTION

The crop forecast is a task whose objective is to anticipate the results of those agricultural products to be planted in a defined area of interest. This information is delivered to those institutions that are responsible for the infrastructure, transportation, sales, storing and distribution of crops.

Being so, the development of methodologies that integrate crop forecasts with the "ground truth" out in the fields, would be of strategic importance, specially for continental countries with a large agricultural production like Brazil.

The development of new processing techniques, and specially of Geographic Information Techniques (GIS), is opening new possibilities for the use of spectral, spatial as well as temporal information from orbital sensor systems, stored in geocoded databases.

Furthermore, these databases can also store historical data from the areas of interest, thus allowing an integration with other available information of interest and consequently permitting an instantaneous retrieval of these information, taking into account manipulations made at the system.

Being so, the objective of this study is to develop a methodology to identify irrigated winter crops, integrating multitemporal TM-Landsat data by an image analysis system (SITIM-INPE) at the GIS-INPE, using historical ground truth data stored at a Database Management System (dBASE IV).

2. AGRICULTURAL STATISTICS VERSUS SATELLITE DATA

Historically crop forecasts have been made using statistical procedures, using samples obtained in the field as input data. This type of sampling procedure is both time consuming and expensive. In this frame, remote sensing has been an useful tool to optimize these procedures.

At several countries, such as at the USA (Chen, 1992), there are studies being performed to use information obtained from orbital remote sensing as input data for crop forecast models.

According to Allen & Hanuschak (1988) at the National Agricultural Statistics Service (NASS), since the fifties, remote sensing data has been used (initially aerial photographs and more recently Landsat data) to build up sampling panels from regions of interest.

In parallel to the experience developed at USDA, the General Directorate for the Agriculture of the Commission of European Communities (CEC) and the Statistical Service of the European Communities (EUROSTAT) at the Joint Research Centre (JRC), have developed the MARS ("Monitoring Agriculture with Remote Sensing") Program. According to Boissezon & Sharman (1992), the objective of this Program is to provide up-to-date agricultural information to those Governmental Institutions responsible to direct the European agricultural policies. In this study, the authors demonstrate the importance on the use of remote sensing integrated to a historical database to identify agricultural targets for crop forecast.

134
In Brazil, several researchers are working on the development of methodologies directed to crop forecast. The works from Chen (1990), Epiphanio (1988), Formaggio (1989) and Ortiz (1993) are noteworthy.

3. "KAPPA" STATISTICS

In order to use remote sensing data for crop forecast, the thematic classification of orbital imagery is a fundamental step. So it is necessary to evaluate the precision of this procedure. According to Story & Congalton (1986) the most common form to express the accuracy of a thematic classification is through the error matrix (also known as confusion matrix). Congalton & Mead (1983) and Congalton et al. (1983) used a method called "Kappa Statistics" to improve the evaluation of the results of a thematic classification from error matrices. Ortiz (1993) made a detailed description of this method. The objective of this statistical procedure is to evaluate the degree of similarity (dependency) between maps, such as between a classified and a reference map. The values resulting from the use of this method may vary from below "zero" until "one". Those values "lower than zero" indicate no similarity (total independence) and one indicates equality (total dependency). Hence so, these values can be grouped in sections, allowing a qualitative evaluation of the results obtained. Andis & Koch (1977) grouped the "KAPPA" value as it is shown in Table 1.

<table>
<thead>
<tr>
<th>&quot;KAPPA&quot; Value</th>
<th>Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;0.80</td>
<td>EXCELLENT</td>
</tr>
<tr>
<td>0.80 - 0.60</td>
<td>VERY GOOD</td>
</tr>
<tr>
<td>0.60 - 0.40</td>
<td>GOOD</td>
</tr>
<tr>
<td>0.40 - 0.20</td>
<td>REASONABLE</td>
</tr>
<tr>
<td>&lt; 0.20</td>
<td>POOR</td>
</tr>
</tbody>
</table>

4. DESCRIPTION OF THE AREA UNDER STUDY

The area under study is localized in the municipality of "Guaira", São Paulo State (Brazil), corresponding to a square of 15x15 km of the Topographic Sheet Guaira (IBGE, 1972), with geographical coordinates S 20°26'02" to S 20°17'24" and W 48°25'02" to W 48°14'58".

According to Oliveira & Prado (1991) the predominant soil type in this region is the "Latosolos Roxo" followed by "Latosolos Variação Uva", two Oxisols derived from basic rocks.

The main land use in the region under study is agriculture. There are dry cultivated fields as well as irrigated fields by central pivots and by self-propelled systems. At the irrigated fields the main cultures are: beans, tomato, soybeans, sorghum seeds, maize seeds, onion and potatoes as well as other cultures of secondary occurrence. In this square there are also secondary forest, some small reforestation stands and some areas planted with Hevea Brasiliensis (the rubber tree).

5. MATERIALS


The data manipulation and analysis was made using an Image Analysis System (STI for "Sistema de Tratamento de Imagens") a Geographic Information System (SIG for "Sistema de Informações Geográficas") and a Database Management System (SGBD for "Sistema de Gerenciamento de Banco de Dados"). The STI and SIG both are developments from INPE, and are known respectively as SITIM (for "Sistema Interativo de Tratamento de Imagens") and SIG (for "Sistema Geográfico de Informações") (Imagem, 1993).

The integration of these three systems was done by specific modules existing at the SIG. So it was possible to visualize an image analyzed by SITIM in a GIS environment, as well as to make questions to the database of dBASE IV, using the "software" Code-Base 4.2 (Sequiter Software Inc. 1990), that was available in the manipulation module, to visualize spatially its data.

6. METHODOLOGICAL STEPS

This work was made considering three main methodological steps. The first one was the acquisition of ground information, during field campaigns, to obtain those data of the lots to be studied and to generate a tabular database.

The second step was the manipulation in the SITIM/SIG/dBASE environments to create both a georeferenced database and a tabular database.

Finally the last step were the digital classifications in order to evaluate potential use of the multitemporal and georeferenced database. Afterwards a statistical analysis was made to obtain quantitative parameters to compare the results obtained with the use of the database versus the results obtained without the use of the database.

6.1. Procedures at SITIM/SIG and dBASE IV

At the SITIM environment the following preliminary procedures were done with the nine multi-temporal Landsat scenes available: atmospheric correction, gray level transformation to reflectance values, registration of image versus topographic map and registration of image versus image.

Afterwards the resultant images were transferred to the GIS. To make this operation possible, initially a project was created at UTM (Universal Transverse Mercator) projection and at 1:50,000 scale.

6.1.1. Elaboration of a Georeferenced Database

The georeferenced database was elaborated mainly over two categories of data to be tentatively used in a GIS: the category polygon and the category spectral images.

The category polygons included the information planes (IPs) whose information originated from the topographic map, the soils map, the image interpretations and the spectral classifications originated from SITIM. Table 2 shows the main IPs of the category polygons and its descriptions.
The category spectral images included those IPs whose information originated from the nine Landsat scenes, previously analyzed with SITIM and transferred to the GIS. So images were generated and named "recortadas" (cut), including only those lots selected for the classification while the other adjacent areas were cut off.

### TABLE 2 - Information Planes of the Category Polygons That Are Included in the Georeferenced Database

<table>
<thead>
<tr>
<th>NAME OF THE PLANE</th>
<th>AUXILIARY PLANES</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1A_500</td>
<td>IP with the soils module</td>
</tr>
<tr>
<td>G1A_501</td>
<td>IP with the change of module</td>
</tr>
<tr>
<td>G1A_502</td>
<td>IP with the module of G1A</td>
</tr>
<tr>
<td>G1A_503</td>
<td>IP with the module of G1A_S01</td>
</tr>
<tr>
<td>G1A_505</td>
<td>IP with the module of G1A_S02</td>
</tr>
</tbody>
</table>

6.1.2. Elaboration of a Tabular Database

The Tabular Database was generated using all information obtained in the field. This database was furthermore fed with the gray levels, the digital representation of reflectance from all spectral bands of the nine scenes used in this study. These gray levels were collected from samples of the fields of interest, using the SITIM function "Calculation of Statistical Parameters".

The Tabular Database was generated at D1ASE IV and connected to the GIS by a specific function called "Consultation to D1ASE." This function associates a record of the Tabular Database to a certain polygon, through a label. This identifier should be unique. So it is not possible to keep one record for two polygons and vice versa.

Initially an unique database was elaborated containing all information available for consultation from the years 1988 to 1990. This database had a structure of 120 fields and totally 70 records for each field. Those records are the information of the lots of IP G1A_PIVO (graphic base), which was used as a base because it had the smallest area element with an unique information, a necessary condition to perform the integration of databases.

6.2. Classification using a database

In order to test the joint use of the database and the digital image classification to identify irrigated cultures planted in the winter of 1991, three experienced image interpreters were invited (referred to as CK91, CF91 and CV91). The system was presented to these interpreters and they were informed on all the possibilities offered by the database during the classification. So three thematic images were generated from the module under study.

6.3. Conventional digital classification

To evaluate if the proposed methodology presented satisfactory results, three further image interpreters were asked (referred to as CO91, CT91 and CS91) to make a conventional digital classification, i.e. a classification without the use of a multitemporal database.

6.4. Crossing the result of classifications with ground truth

To evaluate quantitatively the method proposed, an analysis of the precision of classification was made, where the results obtained by classification using both methods (with and without database) were compared with ground truth data of 1991.

Using the results of these crossings, it was possible to verify how the interpreters made the interpretation, i.e. for each method used there were differences of classification. To perform this analysis both the error matrices and the "KAPPA" statistics were used.

7. RESULTS AND DISCUSSION

7.1. Georeferentiation of images - Registration

After the radiometric uniformization of the nine multitemporal images, the georeferentiation of an image was made with a known cartographic base, putting this image in a standard system of geographic coordinates. The precision indices obtained for this operation were the following: 0.321 pixel for those points used in mapping and 0.587 pixel for those points not used in mapping. Consequently a precision was obtained within the limits of cartographic precision (LME, 1976). The result of the second phase of registration was an average internal position error always below 0.5 pixel.

7.2. Tabular Database versus planting tradition

After the elaboration and manipulation of a tabular database it was possible to confirm that most of the farmers present a tradition in plantation, strengthening the hypothesis that multitemporal information may be used to improve the procedures of identification of agricultural targets.

This concept of tradition in plantation is complemented by the agricultural tradition of this region, that was useful for to identify the agricultural targets of interest. This identification was done by comparison among the areas of interest in different years, i.e., those areas planted with the same crop in successive years, created interpretation patterns for the interpreter, such as the spectral pattern that was helpful to identify cultures during interpretation.

Nevertheless we should emphasize that not all farmers behaved within what was expected, varying their cultures from year to year and being against the concept of tradition in plantation.

Table 3 summarizes the plantations made by farmers for the years 1988 to 1991 within the module under study. One can observe the predominance of beans and a decline of wheat cultures. Another aspect refers to lots with residuals of cultures or bare lands. Most of these lots were planted earlier or were planned for later plantations. This should be considered while analyzing Table 3, since it just presents an instantaneous picture of a phase from the winter period. These lots were called "residuals of cultures" and "bare soils" respectively. If the sampling that originated this table, would have been made at a different time, but during the same agricultural year (winter), these lots would look different (for instance with cultures).

One can consider that the data bank allowed an individual analysis of each farmer and the elaboration of the history of
each area. This history allowed also the study of the agricultural evolution of each lot, improving the performance of the identification of cultures from the module under study. This agrees with the findings of Boissezon and Sharman (1992) who emphasized on the utility of a database to identify plantations of crops, associated to the growing experience of photo-interpreters. At Project MARS the photo-interpreters are visual interpretation memory as a further element of analysis.

This history allowed also the study of the o• the identification of cultures from the module under study. This agrees with the findings of Boissezon. and Sharman (1992) wh() emphasized on the utility of a database to identify plantations of crops, associated to the growing experience of photo-interpreters. At Project MARS the photo-interpreters are visual ititerpretation memory as a further element of an analysis. 1-- 0(1 ,PACloN
developed, it became evident that a more complete database Wherei N'' Number of Lois and Arca* -

pluviontetri( data, changes of tnarkets doe to prices) would be more difficult, because many planes are needed to store lhe

methods but also emphasizes lhe benefits obtained with this

In spite of these problems of compatibility described above, the systems used had a good performance at the development of the study. This is a proof of the viability of the methodology proposed. We should emphasize that some conceptual changes should be introduced into the system so that it becomes more integrated and more user friendly, eliminating unnecessary operation steps. Another option would be the use another system that would be more adapted to the proposed methodology, with a more advanced concept, such as SPRING (Cimara et al., 1993) that has a relational database.

7.4. Classification assisted by a database

According to description of the Methodology, each interpreter generated an image containing the observed themes in the lots studied. These images were migrated to the SGI and transformed in Information Planes that were afterwards compared with ground truth. Figure 1, for example, presents the map of the module under study with the interpretation made by interpreter CK91.

The three interpreters used all the facilities available in the system, performing initially a spectral separation of the 1991 scene that was afterwards classified thematically through the SITIM classification algorithms.

Figure 1 - Map containing the classified plane by interpreter CK91.

Each area. This history allowed also the study of the agricultural evolution of each lot, improving the performance of the identification of cultures from the module under study. This agrees with the findings of Boissezon and Sharman (1992) who emphasized on the utility of a database to identify plantations of crops, associated to the growing experience of photo-interpreters. At Project MARS the photo-interpreters are always the same, because its objective is to keep a human visual interpretation memory as a further element of analysis.

That problems found during the integration of multitemporal data were reported by Wang (1991) who also mentions several problems but also emphasizes the benefits obtained with this integration.

In spite of these problems of compatibility described above, the systems used had a good performance at the development of the study. This is a proof of the viability of the methodology proposed. We should emphasize that some conceptual changes should be introduced into the system so that it becomes more integrated and more user friendly, eliminating unnecessary operation steps. Another option would be the use another system that would be more adapted to the proposed methodology, with a more advanced concept, such as SPRING (Cimara et al., 1993) that has a relational database.

7.4. Classification assisted by a database

According to description of the Methodology, each interpreter generated an image containing the observed themes in the lots studied. These images were migrated to the SGI and transformed in Information Planes that were afterwards compared with ground truth. Figure 1, for example, presents the map of the module under study with the interpretation made by interpreter CK91.

The three interpreters used all the facilities available in the system, performing initially a spectral separation of the 1991 scene that was afterwards classified thematically through the SITIM classification algorithms.

Figure 1 - Map containing the classified plane by interpreter CK91.

The three interpreters used all the facilities available in the system, performing initially a spectral separation of the 1991 scene that was afterwards classified thematically through the SITIM classification algorithms.

Figure 1 - Map containing the classified plane by interpreter CK91.
The spectral separation was important so that the interpreters had a feeling of the spectral differences that cannot be perceived by the human eye. Since the classifier used four spectral bands (TM3, TM4, TM5 and TM7), it was able to make the spectral separation of these crops in some cases. The next step was the use of the database.

In order to evaluate which of the information stored in the tabular database would be most useful for the user, those files from the database that are most frequently used were recorded. The first choice was those that contain all the crops planted each year. The second place was taken by those files that contain the gray levels of each area for all the years of interest. The other files available were used in case of specific doubts.

After each consultation, the interpreters listed the data considered to be of interest for the identification of targets and, at the end of the procedure, all lots were identified. During this part of the study, each interpreter generated an image containing the identification of the themes observed at the lots under study. These images were also transformed into Information Plaines that were afterwards compared with the ground truth. Figure 2 shows the map of the module under study with the classification made by the invited interpreter CO91.

Figure 2 - Map containing the Information Plane classified by the interpreter CO91.

The three interpreters used all the available facilities offered by SITIM, at the module responsible for the application of MAXVER (Maximum Likelihood Algorithm), besides inspecting the images of 1990 to acquire a spectral knowledge of the module under study. The procedure for the identification of crops, in this case, was much easier because the interpreters could use the database with the historical data of the module under study. This procedure made the classification faster. The classification took, in average, six hours but it also became less precise both from the point of view of the correct identification of the

At Table 4 one can observe that the interpreters classified all analyzed data and the class NON-CLASS. Had an area of zero. Another important observation refers to the tomato culture. In this case, the database influenced negatively the interpreters (mainly CV91) because the history of the tomato culture indicates a decrease of the tomato planted area of the former year (1990).

7.5. Conventional Classification

We are presenting below the results obtained by the three interpreters that used the method of digital classification.

Table 4 presents the list of the crops identified by the interpreters who were invited to test the methodology and its respective areas.

Analyzing Table 4 one can observe that the use of the database uniformed the classification of the interpreters for various targets. This observation becomes clear when areas identified as beans, bare soils and residuals of cultures are observed.

### TABLE 4 - LIST OF TARGETS IDENTIFIED BY INVITED INTERPRETERS AT THE LOTS OF THE MODULE UNDER STUDY AND ITS RESPECTIVE AREAS IN HECTARES LOTS AND AREAS IDENTIFIED BY INTERPRETER

<table>
<thead>
<tr>
<th>CLASSES</th>
<th>N°</th>
<th>AREA*</th>
<th>N°</th>
<th>AREA*</th>
<th>N°</th>
<th>AREA*</th>
</tr>
</thead>
<tbody>
<tr>
<td>BARE S.W.</td>
<td>06</td>
<td>138.14</td>
<td>06</td>
<td>137.90</td>
<td>06</td>
<td>131.96</td>
</tr>
<tr>
<td>BEANS</td>
<td>20</td>
<td>317.85</td>
<td>20</td>
<td>251.42</td>
<td>20</td>
<td>291.55</td>
</tr>
<tr>
<td>BARE S.</td>
<td>13</td>
<td>337.90</td>
<td>11</td>
<td>265.66</td>
<td>13</td>
<td>317.75</td>
</tr>
<tr>
<td>MAIZE</td>
<td>07</td>
<td>165.69</td>
<td>09</td>
<td>256.03</td>
<td>07</td>
<td>119.53</td>
</tr>
<tr>
<td>D-MAIZE</td>
<td>08</td>
<td>33.85</td>
<td>03</td>
<td>14.90</td>
<td>08</td>
<td>124.52</td>
</tr>
<tr>
<td>S-W-ITBM</td>
<td>05</td>
<td>87.05</td>
<td>02</td>
<td>46.56</td>
<td>00</td>
<td>00</td>
</tr>
<tr>
<td>ARCHONIA</td>
<td>00</td>
<td>-</td>
<td>02</td>
<td>7.76</td>
<td>00</td>
<td>-</td>
</tr>
<tr>
<td>S-W-TRANS</td>
<td>00</td>
<td>7.76</td>
<td>00</td>
<td>-</td>
<td>00</td>
<td>00</td>
</tr>
<tr>
<td>TOT-CLASS</td>
<td>00</td>
<td>-</td>
<td>00</td>
<td>-</td>
<td>00</td>
<td>00</td>
</tr>
<tr>
<td>TOTAL</td>
<td>44</td>
<td>1165.8</td>
<td>39</td>
<td>1165.9</td>
<td>48</td>
<td>1164.9</td>
</tr>
</tbody>
</table>

Where: N° = Number of lots and Area* = Area in Hectares

At Table 4 we can see that class "SOYBEANS" even though it was planted in the agricultural year of 1991, it was not listed at Ground Truth (Table 3), because it is always planted very late and it does not appear very clearly on the classified image.

Nevertheless two interpreters, while analyzing the history of the soybean culture using the database, concluded that some areas that were fallow or had residues of vegetation should be classified as soybeans. This phenomenon explains in part how this class appeared. One must consider also the phenology of maize and sorghum which, at some lots, caused spectral confusion for the interpretation.

The spectral separation was important so that the interpreters had a feeling of the spectral differences that cannot be perceived by the human eye. Since the classifier used four spectral bands (TM3, TM4, TM5 and TM7), it was able to make the spectral separation of these crops in some cases. The next step was the use of the database.

In order to evaluate which of the information stored in the tabular database would be most useful for the user, those files from the database that are most frequently used were recorded. The first choice was those that contain all the crops planted each year. The second place was taken by those files that contain the gray levels of each area for all the years of interest. The other files available were used in case of specific doubts.

After each consultation, the interpreters listed the data considered to be of interest for the identification of targets and, at the end of the procedure, all lots were identified. During this part of the study, each interpreter generated an image containing the identification of the themes observed at the lots under study. These images were also transformed into Information Plaines that were afterwards compared with the ground truth. Figure 2 shows the map of the module under study with the classification made by the invited interpreter CO91.

Figure 2 - Map containing the Information Plane classified by the interpreter CO91.

The three interpreters used all the available facilities offered by SITIM, at the module responsible for the application of MAXVER (Maximum Likelihood Algorithm), besides inspecting the images of 1990 to acquire a spectral knowledge of the module under study. The procedure for the identification of crops, in this case, was much easier because the interpreters could use the database with the historical data of the module under study. This procedure made the classification faster. The classification took, in average, six hours but it also became less precise both from the point of view of the correct identification of the
targets analyzed (see next section) but also from the spatial point of view (forms of the lots).

The irregularity observed at the format of the polygons becomes evident while analyzing Figure 2, where interpreter C091, because of not knowing the historical behavior of plantation practices of the farmers in the module under study, divided some lots on a non-typical format (see arrows in Figure).

This type of behavior by this interpreter indicated that it is not enough to make the identification of agricultural crops or even a classification or to have just the spectral knowledge of the area. It is necessary to have further knowledge of the area which can be obtained from the database or from lengthy experience, like it was done at Project MAKS.

Table 5 contains the list of crops identified by the invited interpreters to perform the conventional classification.

At this table one can observe that the interpreters did not present an uniform behavior, since a variation of the sizes of each class occurred from interpreter to interpreter.

### TABLE 6 - List of Targets Identified by the Interpreters on the Lots of the Module Under Study and Its Areas in Hectares (Conventional Classification)

<table>
<thead>
<tr>
<th>CLASS</th>
<th>Areaa</th>
<th>INTERPRETERS</th>
<th>% of Global Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOMATO</td>
<td>63.81%</td>
<td>C091</td>
<td>63.81%</td>
</tr>
<tr>
<td>SORGHUM</td>
<td>57.32%</td>
<td>C091</td>
<td>57.32%</td>
</tr>
<tr>
<td>MAIZE</td>
<td>47.69%</td>
<td>C091</td>
<td>47.69%</td>
</tr>
<tr>
<td>BEANS</td>
<td>31.56%</td>
<td>C091</td>
<td>31.56%</td>
</tr>
<tr>
<td>SUGAR CANE</td>
<td>21.12%</td>
<td>C091</td>
<td>21.12%</td>
</tr>
<tr>
<td>韞</td>
<td>0%</td>
<td>C091</td>
<td>0%</td>
</tr>
<tr>
<td>TOTAL</td>
<td>100%</td>
<td>C091</td>
<td>100%</td>
</tr>
</tbody>
</table>

Where N° = Number of lots and Areaa = Area in hectares

We could verify also that the two interpreters were not able to classify some areas, that's why there are NON-CLASS (Non-classified) lots. Furthermore the class TOMATO was not identified, even though it occurs in the module. All these aspects show again that the use of the multitemporal database is of great help for the identification of agricultural targets on satellite images.

7.6. Statistical Analysis of Classifications

Table 6 presents the percentages of global performance for each interpreter from both the group that was helped by the database as well as from the group that made just the conventional classification.

Analyzing the classifications made by all interpreters, two different groups were found out: those interpreters of the first group (C591, C991 and C191) who obtained a global performance (precision) of 75.69%, 73.43% and 71.84% respectively (average of 73.65 and standard deviation of 1.57) and the interpreters of the 2nd group (C091, C591 and C191), who obtained a global performance of 63.81%, 51.73% and 55.82% respectively (average of 57.12 and standard deviation of 5.00).

These results show that the first group had a better performance than the second one. The improvement of performance from the latter group is probably due to the use of the multitemporal database, because this was the sole factor to differentiate among both groups of interpreters.

### TABLE 6 - Performance of Two Groups of Interpreters Who Tested the Methodology Proposed

<table>
<thead>
<tr>
<th>INTERPRETERS USING DATABASE</th>
<th>% OF GLOBAL PERFORMANCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>C091</td>
<td>75.69</td>
</tr>
<tr>
<td>C591</td>
<td>73.43</td>
</tr>
<tr>
<td>C191</td>
<td>71.84</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>INTERPRETERS WITHOUT DATABASE</th>
<th>% OF GLOBAL PERFORMANCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>C091</td>
<td>63.81</td>
</tr>
<tr>
<td>C591</td>
<td>57.32</td>
</tr>
<tr>
<td>C191</td>
<td>55.82</td>
</tr>
</tbody>
</table>

Table 8 shows the results generated from the error matrices (statistical comparison among each interpretation and the ground truth).

### TABLE 7 - Results of "KAPPA" Statistics

<table>
<thead>
<tr>
<th>MATRICES</th>
<th>LOWER LIMIT</th>
<th>UPPER LIMIT</th>
<th>LANDIS &amp; KOCHE (1977)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CK91</td>
<td>0.693289</td>
<td>0.696249</td>
<td>&quot;VERY GOOD&quot;</td>
</tr>
<tr>
<td>CV91</td>
<td>0.691734</td>
<td>0.694754</td>
<td>&quot;VERY GOOD&quot;</td>
</tr>
<tr>
<td>C091</td>
<td>0.689219</td>
<td>0.692175</td>
<td>&quot;VERY GOOD&quot;</td>
</tr>
<tr>
<td>CK91</td>
<td>0.657277</td>
<td>0.660234</td>
<td>&quot;GOOD&quot;</td>
</tr>
<tr>
<td>CV91</td>
<td>0.649555</td>
<td>0.652511</td>
<td>&quot;GOOD&quot;</td>
</tr>
<tr>
<td>C091</td>
<td>0.640033</td>
<td>0.643089</td>
<td>&quot;GOOD&quot;</td>
</tr>
</tbody>
</table>

To improve the evaluation of Table 7 the scale of values proposed by Landsis & Koch (1977) was used, where the classifications of the first group obtained the quality index "VERY GOOD" and the classifications of the second group obtained index "GOOD".

We could verify that the indices of correct percentage per category of each matrix, obtained by both groups, presents the summary of the behavior of the interpreters. Being so, we noticed that, at group 1, classes bare soil, residuals of cultures and beans obtained each a good performance, whereas at group 2 the good performance occurred at classes bare soil, residuals of cultures and maize, except for interpreter CT91 who obtained an index of 11.99% for the class maize.

Classes bare soil and residuals of cultures had a quite different spectral reflectance of all others. This difference was a contribution for its identification. As for the beans fields, the high index of correctness of classification at group 1 was due to consultations of the database by the interpreters.

Referring to what was discussed at Section 7.4., related to the negative contribution of the database to the identification of tomato fields, mainly by two of the interpreters, it is to say that this can be confirmed by the indices of correctness for this class, where we verified 39.17% (CK91), 31.56% (CV91) and 61.71% (C991).

Even though group 1 obtained a lower performance than group 2 for the class maize, at the other classes (tomato, sorghum and "abobrinha"), at least one of the interpreters of group 1 obtained a correct identification. This did not happen at group 2. In this case interpreters of group 2 were not able to localize any lot of classes tomato and "abobrinha", identified erroneously as sorghum.
It was verified also that interpreters of Group 2 obtained higher error indices than group 1 for the classes bare soil, cultural residuals, sorghum and non-classified (NON-CLASS.) areas. For classes tomato and "abobrinha" there was no error percentage since they were not identified in the classification. Classes sorghum and NON-CLASS. were identified, but within areas of other classes, reaching an inclusion error of 100%.

Referring to the isolated good performance of the interpreters of group 1 at the identification of sorghum and "abobrinha" fields, this can be credited to the different approaches of each interpreter on the use of the database. It should be emphasized that a longer contact of the interpreters to the system developed in this work, would strongly increase the results already obtained.

5. CONCLUSIONS

The results obtained with this methodology presented an average global performance index of 73.65 % with a standard deviation of 1.57. We can conclude that this index was higher than the index obtained by conventional classification (57.12%), with a standard deviation of 3.00 due to the good performance of the database.

The statistical analysis of results showed a superiority of 28.93% of the proposed methodology in comparison to the conventional classification method.

Referring to "KAPPA" statistics, group 1 obtained an average value of 0.668974 and group 2 of 0.472170 and according to the classification of Koch & Landis (1977) this would mean that these classifications are "very good" and "good" respectively.

The use of the tabular database associated to spatial databases. We suggest for the use of this methodology in larger areas, to use a system with a relational database concept working together with a module for image analysis. This would avoid all the problems related to data integration.

Referring to the establishment of a database, the systems used were efficient, allowing to store and manipulate a large number of information of the area studied.

As for the methodological steps performed, the pre-processing phase was important to uniform the images used, taking care specially with the atmospheric correction, in order to avoid an overestimation of gray levels that should be extracted from the rough images. Being so, we recommend the use of methods of atmospheric correction that would minimize this effect. In spite of a small temporal sampling within the database (3 years), almost all areas maintained a tradition of planting, even taking into account the dynamics of planting typical for winter crops, which was a help for those interpreters that used the database. It would be interesting to test this methodology in a database that contains a larger temporal series and that would include other useful information for the interpreter, such as rainfall and agricultural statistics data.

REFERENCES


Boissezon, H.; Sharrnon, M. Rapid estimates of crop production using remote sensing: European experience. Italy, Joint Research Centre, 1992 (Manuscrito).


Ortiz M.J. Integração do Sensoriamento Remoto, Sistema de Informações Geográficas e Banco de Dados na Identificação de Culturas Agrícolas de Inverno. INPE, São José dos Campos, SP. 1993. 140p. (Dissertação de Mestrado)


