A processing and software system for rice crop inventory using multi-date RADARSAT ScanSAR data

M. Chakraborty *, S. Panigrahy
Space Applications Centre, ISRO, SAC PO, Ahmedabad 380053, India

Abstract

An operational crop survey program requires standardised procedures and software packages to meet the specified targets of timeliness and accuracy of estimates. Currently, the focus is to include Synthetic Aperture Radar (SAR) data in such a program, as these data are available from a number of sensors. A procedure has been developed to use multi-date SAR data for rice crop inventory. The steps were packaged together for ease-of-use and with minimal user interaction. The package, SARCROPS, is built around the EASI/PACE software. It is, at present, tuned for RADARSAT ScanSAR data. The package was used during the 1998–1999 and 1999–2000 seasons to estimate the rice area at state level in India with the participation of a number of interdisciplinary users. Around 45 and 89 scenes of ScanSAR data were used during these two seasons, respectively. This paper reports the details of the SARCROPS processing chain. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: RADARSAT; ScanSAR; procedure standardisation; software package, SARCROPS; rice crop monitoring; multi-temporal classification

1. Introduction

Standardised procedures and appropriate software packages are critical requirements for remote-sensing-based operational crop survey programs, which involve a number of interdepartment and interdiscipline officials. Synthetic Aperture Radar (SAR) data have a significant role in remote-sensing-based crop surveys, particularly for rice-growing regions. SAR data from ERS and RADARSAT showed good results for rice crop identification in India (ESA, 1995; Chakraborty et al., 1997; Panigrahy et al., 1997, 1999). Investigations using ERS and RADARSAT SAR data in other major rice-growing countries have also shown promising results for rice crop monitoring, e.g., in China (RSI, 2000b), Java (Ribbes and Le Toan, 1999), Japan (Kurosu and Fujita, 1997), Indonesia (Le Toan et al., 1997) and Vietnam (Liew et al., 1998). It has also been established that multi-date data acquired at critical crop growth stages are an essential requirement to obtain desired classification accuracy.

RADARSAT ScanSAR Narrow B Beam (SNB) data having a wide swath and shallow incidence angle range are of particular interest for large area

Corresponding author. Fax: +91-79-674-8813.
E-mail address: manabsac@yahoo.com (M. Chakraborty).
rice crop monitoring (Table 1). In India, the median rice crop field area is about 0.25 ha. However, rice is grown in larger contiguous areas measuring 2 ha (ca. 5.6 × 5.6 pixels of SNB data) or more. SNB data have been found suitable for large area monitoring for this reason. Theoretically, ScanSAR Narrow data can monitor fields of 0.56 ha or larger as compared to 0.2 ha for Standard beam. However, it covers nine times the area of the Standard beam. Preliminary investigations using SNB data for the 1997–1998 season showed promising results for rice crop monitoring (Panigrahy et al., 2000). Based on this experience, a pilot project was launched during the 1998–1999 season to estimate rice area using multi-date ScanSAR data for four major rice-growing states of India. Subsequently, a major project was launched for rice crop inventory at the national level during the 1999–2000 season.

A software package, named SARCROPS, was developed, encompassing the necessary processing steps and enabling analysis of multi-date SAR data for rice crop inventory in a uniform and easy manner. It was developed around the EASI/PACE software and included preprocessing steps required for SAR data analysis like calibration, speckle removal and creation of multi-date datasets. The software is platform-independent. SARCROPS is written in EASI Plus (EASI/PACE Macro Language) for software linking and EASI Modeling Language for new modules. ANSI C has been used in the existing modules of the sample segment approach (see Section 3). The package, at present, uses RADARSAT ScanSAR data. However, it can be easily modified for any other SAR data. This paper describes the methodology used and results obtained.

### 2. Objectives

The objectives are:

(i) To standardise the procedures for district/state level rice crop inventory using multi-date SAR data; and
(ii) To develop a software package for use of SAR data in remote-sensing-based crop inventory applications in an image processing software environment.

3. Data used

RADARSAT SNB data have been used. Narrow beam data were selected because of its large swath (300 km) and the B beam position because of its shallow incidence angle range (31°–46°). SNB data are generated by combining the beam positions W2, S5 and S6 (RSL, 2000a). Four states covered by 10 ScanSAR scenes (one pass) and 11 states covered by 27 ScanSAR scenes (one pass) were used during the 1998–1999 and 1999–2000 periods, respectively. Descending mode data with acquisition time of 0500–0630 h (local time) were used in this study. The descending mode was selected in order to match the orbit inclination with optical RS data. The sample segment approach (SAC, 1995a) has been designed and tuned for the inclination of the satellite track in the descending mode. In addition, the morning pass reduces the probability of rains and also the effect of temperature or moisture stress on rice plants is less. Three-date to four-date data acquired with 24-day repeat cycle from July to October were used. The data acquisition plan was made keeping in mind the rice-planting calendar of each state. First-date data were acquired coinciding with the puddling to transplanting stage. Subsequent data were acquired with a 24-day repeat cycle. The total numbers of ScanSAR scenes used during 1998–1999 and 1999–2000 were 45 and 89, respectively.

4. Methodology

Multi-date ScanSAR data were used as dataset to derive information on rice area. The first two-date data acquired within 24–30 days of planting were used to obtain an early estimation of crop prospects. The estimates were updated using third and fourth date data, well in advance of crop maturation. To carry out the analysis, SARCROPS was used. SARCROPS has basically three distinct processing steps. These are the following.

(i) Data preprocessing and dataset preparation. This is the most time-consuming part of the entire analysis chain of the SARCROPS package. For the implementation of this step, software has been written for linking various components and generating new ones that were missing. The following EASI/PACE programs have been used in the processing steps, viz. CDSAR for downloading the image and header data on computer hard disk, FELEE for enhanced Lee filtering, FTF and FTI for the forward and backward fast Fourier transforms and GCPWorks, the stand-alone image geo-referencing software. New software functionality includes incidence angle calculation, calculation of backscattering coefficient (the SARSIGM function of EASI/PACE has been rewritten to reduce disk space, time of processing and recoding back to an 8-bit image channel) and the descalloping submodule.

(ii) Sample segment extraction and joining, segment-wise crop proportion calculation and district/state level crop acreage estimation. Software modules for these steps were already developed under the CAPEMAN/CAPEWORKS software package (SAC, 1995b). In SARCROPS, links to the required modules have been established.

(iii) Classification of image data for the given crop. This is carried out for each of the sample segments. This is the input to the segment-wise crop proportion mentioned in step (ii) above. In SARCROPS, this is implemented using the EASI Plus modelling language implemented in the stand-alone and interactive ImageWorks package of EASI/PACE. Optionally, one may use the maximum likelihood classifier (MXL function) available in the Classworks submodule of ImageWorks.

Fig. 1a and b show the overall flow diagram of the SARCROPS package. The following subsections describe the steps of the procedure.

4.1. Data preprocessing and dataset preparation

This comprises two submodules. The first module includes the steps of data downloading, data inspection and, if necessary, descalloping of the data (see Appendix A), speckle suppression using a predefined adaptive low-pass filter, conversion of pixel digital numbers (DNs) to backscatter values and coding back into 8-bit DN using a predefined scaling. The
Fig. 1. (a) Flow diagram of SARCROPS software package. The connected blocks of the flow diagram shown on the left-hand side of the figure describe the data preparation steps in the use of multi-date SAR data (here RADARSAT ScanSAR) for crop inventory. The blocks of the flow diagram shown on the right-hand side describe the sample segment approach software modules. (b) Flow diagram of SARCROPS software package (continued). The connected blocks of the flow diagram shown on the left-hand side of the figure describe the descalloping procedure for ScanSAR data. The blocks of the flow diagram shown on the right-hand side describe the classification steps.
second module includes the steps of image georeferencing (i.e. develop image-to-map transformation model) and multi-date image dataset preparation. In this step, georeferencing data, provided in the satel-
lite header, are used. In addition, the user has to include ground control points obtained from GPS or 1:50,000 scale topographic map sheets. Multi-date datasets are generated using image-to-image registration. The steps of each module are given below.

4.1.1. Module 1

4.1.1.1. Data downloading. Image data from the CD-ROM (CEOS format) are downloaded on the disk using the CDSAR program of EASI/PACE. For ScanSAR data, it creates the necessary 8-bit image channel of required image size. For ScanSAR data of descending mode, it carries out the necessary east–west flipping of the data. Transfer of all required geometric and radiometric header information into the same image file is also carried out.

4.1.1.2. Data inspection. The data are checked for its coverage and quality. If scalloping image banding is observed, then a frequency-domain-based filter is used to remove this. The details are given in Appendix A and also shown in the left column of blocks in Fig. 1b. Image banding is observed in some ScanSAR Narrow Beam data appearing as a periodic variation of the grey values of the image, oriented nearly east–west slanted to the south.

4.1.1.3. Speckle suppression. At present, the package uses an enhanced Lee filter (FELEE function of EASI/PACE) with $5 \times 5$ window size. This is based on the experience of the last 5 years using ERS and RADARSAT data for the Indian agricultural conditions. The operation of this filter is carried out in-place in the same image channel as the original.

4.1.1.4. Data calibration. Using the orbit information stored in the header (transferred to the image data as a segment), the incidence angle of each pixel $i$ in a scanline is computed as follows:

$$I_i = \cos^{-1}\left[\left(\frac{H + H - R_{s_j} \times R_s}{2 \times R_s + R}\right)/\left(2 \times R_{s_j} + R\right)\right],$$

where, $I_i$ = incidence angle; $R_{s_j}$ = slant range; $H$ = orbit altitude, and $R$ = radius of the Earth.

Here, the area covered by a pixel is implicitly assumed to be horizontal in orientation. This is not true in undulating terrain; however, for agricultural areas, this assumption is quite appropriate.

Using this computed incidence angle, the backscattering coefficient $\sigma_0$ is computed as follows:

$$\sigma_0 = 10 \times \log_{10}\left[\left(DN^2_j + A_o\right)/A_j\right] + 10 \times \log_{10}\left[\sin(I_j)\right] \text{(dB)},$$

where $DN_j$ is the DN, $A_j$ is the scaling gain value, $I_j$ is the incidence angle at the $j$th range pixel and $A_0$ is the fixed gain offset, usually zero.

The backscattering coefficient $\sigma_0$ thus computed is then linearly scaled back to the range 0–255 and written in the same image channel. In this process, the range of backscatter values that can be stored is $-26.05$ to $-0.46$ dB with quantisation error of $\pm 0.05$ dB. This is acceptable as the whole useful range of $\sigma_0$ can be accommodated (observed for different land cover from very calm water or smooth surfaces to dense forest) and the quantisation error is much less than the relative and absolute calibration accuracy of $\sigma_0$. Only dense urban areas show saturation. This approach is necessary to reduce the disk space requirement and data handling and enables the use of most image processing modules, which operate on 8-bit data only.

4.1.1.5. Quadrant image formation. ScanSAR full scenes are subdivided in four overlapping quadrants. This step is carried out for the following reasons: (a) for easy data handling and faster response; (b) improvement of image-to-map registration accuracy; (c) restriction of the number of sample segments to fit within the display screen, when a sampling approach is used. These quadrants are extracted automatically using software called QUADEXT, which uses the image size information from the header.

4.1.2. Module 2

4.1.2.1. Image-to-map registration. The first-date quadrant data are treated as base or master image. Image-to-map registration is carried out in this dataset using the procedure described below.

From the orbit information contained in the header file of the full ScanSAR data, latitude/longitude values of first, middle and last pixels of each scan
line of data are generated. At least six appropriate bounding latitude/longitude co-ordinates are transferred to each quadrant image as header GCP (ground control point) segment using an add-on program (SARFIT). A few additional GCPs (obtained using GPS or from 1:50,000 scale maps) are used to validate and improve the accuracy of the header GCPs. The GCPs are edited, if required, and image-to-map and map-to-image fitting equations are developed using second-order polynomial models. This image-to-map GCP segment and the fitting equations are used to georeference the image or to extract sample segments. In the sample segment approach, a georeferenced output image is not required as the image chips corresponding to the sample segments are extracted from the original image using the map-to-image transformation.

4.1.2.2. Multi-date dataset preparation. The subsequent date data are registered to the first-date quadrant data using an image-to-image registration procedure. A first-order affine transformation is used to create multi-date registered datasets using nearest neighbour resampling.

The above two steps are carried out using the GCPWorks module of EASI/PACE. The left column blocks of Fig. 1a show the flow diagram of these steps. This completes the dataset preparation for multi-date SAR data. One can then use either the complete enumeration approach, using district/state boundaries, or a sampling approach (as used in the SARCROPS package) to estimate the crop area.

4.2. Sampling approach

The sample segment approach has been used in the estimation of crop acreage and production in order to cover a large area within reasonable time and cost. A stratified random sampling method (Cochran, 1977) has been developed over time for various crops in different states of India, and is being implemented routinely using optical remote sensing data (SAC, 1995a). A state forms the sampling frame, which is then divided into $5 \times 5$ km$^2$ units, aligned to the satellite pass (descending mode). All grids with more than 95% agricultural area are treated as the population. The population is then stratified based on various other criteria, often depending on crop proportion estimated from remote sensing data. Finally, 5–10% of the samples is randomly selected from each stratum. A database is prepared in a specific format, giving latitude and longitude information of each segment and its identification. Images of these selected sample areas are extracted from the original satellite image.

Ground truth, generally, is acquired in these sample areas and used to classify the images. For convenience, sample images falling in each scene are sorted out as per the study districts and joined together. The results are aggregated following the standard statistical approach of stratified random sampling (Cochran, 1977) to obtain district and state level crop estimation. The details of the procedure are given in the manual of the CAPE procedure (SAC, 1995a).

Extraction of sample segments, joining of segments, training statistics, classification, histogram of crop classes to statistical aggregation for acreage are carried out using existing submodules of CAPEMAN (SAC, 1995b). Since this package is not within the scope of this paper, it is not discussed further in detail.

The core functional blocks of the software for the sample segment approach are shown on the right-hand side of Fig. 1a. These consist of locating the sample segments in the image, extracting them from the full image, and joining or tiling them for convenience, and rejection of bad segments. After the sample segment images are classified, calculation of crop proportion in each segment and finally district or state-wise crop acreage estimation are carried out.

4.3. Classification of multi-date SAR data

Per-pixel classifiers have been selected for use in this study. The classifiers used are:

(i) Maximum likelihood classifier, and
(ii) Decision-rule-based classifier.

4.3.1. Maximum likelihood classifier

This is the most robust and well-understood per-pixel classifier. However, it is not well suited for multi-date datasets (even more so for SAR data) used in the present study. The reason for this is the
large variability in the multi-date SAR dataset. This leaves a large proportion of unclassified pixels even with an exhaustive ground truth training dataset.

4.3.2. Decision-rule-based classifier

An alternate approach of a decision-rule-based classifier has been developed for this study (Chakraborty and Panigrahy, 1997). In this approach, logical and mathematical conditions are used to classify a pixel as crop pixel or not. In addition, qualitative tags like early, normal, late, good, poor, etc., can be attached to the crop pixel (based on the outcome of the logical/mathematical conditions). Unlike the maximum likelihood and other supervised classifiers, class identification of other land use/land cover classes is not needed. The following steps are used in this approach to classify rice crop pixels:

(i) The ground truth sites are first scrutinised to categorise the rice areas into early, late and other cultural practice based distinct categories;
(ii) The backscatter values from identified puddled fields and water bodies are used to develop thresholds (lower and upper) for puddled fields;
(iii) The growth characteristics of different categories of rice areas in temporal datasets are used to develop decision rules for identification of rice classes; and
(iv) The broad decision rules are further fine-tuned for each area to take care of the specific growing conditions of rice in different regions.

This classifier is well suited to handle the very high variability in the multi-date SAR data and gives meaningful categorisation of rice classes based on their growth and management practices. Qualitative crop condition assessment is also feasible using this classifier. The block diagrams on the right-hand side of Fig. 1b show the classification steps.

5. Discussion

Automated software packages are one of the essential requirements for operational crop survey methods not only to meet timeliness targets but also to maintain uniformity of the approach when a number of interdisciplinary users are involved. This is more so when SAR data are concerned, as a user has to carry out a number of data processing steps before it can be used for classification. SARCROPS is an attempt towards this and was found to be very effective in increasing the efficiency of user participation in SAR-based crop inventory.

Data calibration and speckle suppression are two steps essential for SAR-data-based studies. SAR, being an active sensor, has the advantage of retrieving calibrated multi-temporal target properties. This can be used for knowledge-based classification of the crops, and modelling of the crop growth parameters and crop condition. The necessary software was developed to generate backscatter images.

An enhanced Lee filter with $5 \times 5$ window size has been found to be optimum for the Indian agricultural conditions. However, it can be replaced in the future by a more advanced adaptive filter.

Accurate image-to-map and map-to-image transformations are one of the critical requirements of such studies. Unlike optical sensors, it is extremely difficult to identify and locate good quality GCPs in SAR data due to the presence of speckle and lack of well-formed geometric shapes. Use of satellite header information was extremely useful in quick location of user-supplied map/GPS GCPs and minimising their requirement. Four to six bounding GCPs obtained from the data header were found to be essential to keep the overall error within bounds throughout the image. Identification of map GCPs was the most time-consuming step in the study. At present, a user is required to supply a few map GCPs (6–10). Since the sample segments were of size $5 \times 5$ km$^2$, an average map location error of 125 m ($2.5\%$ of the segment size) of the sample segments was considered acceptable. Hence, the acceptance criteria of the accuracy of the ScanSAR Narrow Beam GCPs (for image-to-map registration) were:

(i) The residual root mean square error (RMSE) should be less than five pixels (125 m); and
(ii) The maximum residual error must not exceed 10 pixels (250 m).

For ScanSAR data, an average RMSE error of five pixels or less could be achieved. It is essential that the desired accuracy using the satellite ephemeris be obtained, so that this step can be fully automated.
Multi-date SAR data were used as one combined dataset similar to multispectral optical data. Thus, it requires very accurate image-to-image registration. Lack of well-defined GCPs in SAR data coupled with changes in geometrical shapes in different dates was the major constraint to achieve this. It may be mentioned that one cannot achieve the desired error criterion of less than 0.25 pixels, as in the case of band-to-band registration of multispectral optical sensors. For ScanSAR data, an RMSE of 0.5 pixel could be achieved using the GCPWorks module of EASI/PACE. Hence, the acceptance criteria of the accuracy of the GCPs were:

(i) Residual RMS error should be less than 0.5 pixel; and
(ii) Maximum residual error must not exceed one pixel.

This can be improved using self-similarity detection algorithms (and other automated image-to-image fitting software) which are, at present, not found in most of the commercial image processing softwares.

6. Conclusion

Rice crop was separable from most of the other land cover classes in two and three date ScanSAR Narrow Beam data and was found to be cost-effective for large area rice crop monitoring. The cost of data was US$4.15 per thousand hectares of rice crop. The total geographical area covered during the 1999–2000 season was approximately 210 million ha (of which about 35 million ha are occupied by rice crop). The analysis required approximately 2350 man h to complete (ca. 17,000 ha/h). The total cost (in Indian terms) was approximately US$7.10 per thousand hectares of rice crop. This is of particular interest in the Indian context, where rice is grown in about 40 million ha, often in large contiguous areas, as the most dominant crop during the rainy season. SARCROPS was used during the 1998–1999 season for four states and again validated for most of the rice-growing states during 1999–2000. The classification accuracy for rice crop ranged from 91% to 95% over the 11 states under study. Around 40 scientists of various departments have used this package for state and national rice crop estimation during the 1999–2000 season. The analysts were spared the nitty gritty of the SAR data preprocessing. This package was found to be very effective in achieving the timeliness target and standardisation of the analysis procedure. The final estimates for rice acreage could be released more than 30 days prior to harvesting in each of the states. Attempts are being made to incorporate a knowledge-based classification algorithm in this package.

Appendix A. Image descalloping

Scalloping (image banding) is observed in some of the ScanSAR data of 1998 and 1999. This effect is caused by difficulty in calculating the azimuth beam pattern in the ScanSAR mode. It is observed as near-horizontal data banding tilted slightly from southwest to northeast at a spacing of approximately 57 lines (for SNB data). In the frequency domain, scalloping is observed as discrete frequencies, aligned slightly left of the vertical axis. These are equally spaced, indicating a periodic behaviour which is observed to be constant over different images. This allows one to construct a fixed band-reject filter in the frequency domain, which removes the spikes corresponding to the scalloping.

The approach is to remove only these unwanted frequencies using a band-reject mask in the two-dimensional frequency domain. Descalloping has to be carried out prior to any other preprocessing step. The procedure used for descalloping is as follows:

(i) Select the area to be descalloped. The size of the area must be a multiple of 1024 pixels in both pixel and line directions;
(ii) In case of a large area to be processed, the area is divided into 1024 × 1024 pixel subareas and processed one by one;
(iii) Fast Fourier transform of each subarea is carried out to obtain the frequency domain;
(iv) Using the band-reject mask prepared earlier, mask out the unwanted frequencies associated with scalloping;
(v) Inverse fast Fourier transform of the image to obtain the descalloped area. After processing of all subareas, the resulting image is used for further processing.
References


