Chapter 7

Overall Results and Discussion

This page is intentionally left blank.

Chapter 7

Overall Results and Discussion

In this chapter discussion on various elements of this research are presented including selection of study districts / state and the results obtained through the three sub-studies. Gujarat state was selected as the study state because it is one of the significant contributors in agricultural production of India and is has been in news for long time due to its recent good progress in agriculture sector (Gulati et al., 2009; Shah, 2009; IBTL, 2011; Kashyap, 2012; Swain et al., 2012). The crop sown area in Gujarat is significant at the National Level in Bajara, maize, Tur, onions, cabbages, brinjal and okra, bananas, papaya, sapodilla (Chiku), groundnut, castor, sesame (Til), cotton seed, fennel, cumin, fenugreek, carom seeds (Ajwain), garlic, cotton and tobacco. At the state level, the following crops are dominant in terms of area: rice, wheat, Bajara, maize, potatoes, groundnut, and cotton (Alagh, 2014). Agriculture in Gujarat influences the economic activity in the region (directly) through its dominant share and (indirectly) through its inter-linkages with other sectors of the economy…instability in agriculture, on its own, affects around 40 per cent of economic activity and 70 per cent of the population of the region (Kashyap, 2012).

Mehsana district was selected as a study district in Sub-study-A because it is one of the multiple crops district of Gujarat with considerable share of crop area under wheat, mustard, cumin, and potato along with cotton and castor crops (basically long duration Kharif crops extending into Rabi season months). Presence of these long duration Kharif crops poses challenge of discrimination of Rabi crops from long duration Kharif crops using single data LISS-III data. It made a good case of early estimation of Rabi crop area using single date LISS-III and multi-date AWiFS data.

7.1 Overall Results

The overall result of this research is that the information extracted from multi-source data could be integrated for early estimation of current year's crop sown area at the district as well as the state scales. The integration of multi-date AWiFS data with single date LISS-III data could provide early estimate of rabi crop sown area in Mehsana district for 2011-12 season. The classified images based on the integrated data resulted in early estimation of the crop sown area of 58.9 thousand hectare which is close to the expected value of 62 thousand hectare. Error matrix based accuracy assessment of the classified images was carried out. While overall accuracy of single date classification using MLC was 83.4percent, it was 79.3 percent in case of single date ISODATA classification. The overall accuracy of 96.5 percent with Kappa coefficient of 0.954 was obtained in case of integrated dataset classification.

Similarly, based on the decision fusion of the information derived from the multi-date AWiFS and MODIS data added with crop history information (derived from multi-year RS data) at state level could provide early crop sown area of Gujarat for 2011-12 season. The estimated area based on 9-dates MODIS data was found to be 2.370 million hectares, which is higher than the reference value. The integrated dataset (AWiFS + MODIS + crop history) yielded Rabi crop sown area of 1.125 million hectares. Rabi crop area at the end of crop season was reported to be 2.045 million hectares (DES, 2016). The Rabi crop sown area as on mid-December 2011 was reported to be 1.238 million hectares (DES, 2016). The government's reference data showed that almost 60% of full season Rabi crop was sown as on December 16, 2011. The full season crop sown area determined from the reference classified image was 1.972 million hectares. Hence, the proposed methodology based on decision fusion of multi-sensor information could estimate the crop sown area. The mapping accuracy of Rabi sown area was found to be 88.6% with respect to the reference full season AWiFS image based Rabi sown area.

7.2 Discussion on the Sub-studies A, B and C

Sub-study-A demonstrated that a reliable early estimation of Rabi crop area at district level can be made by synergistic use of the multi-sensor data. If single date high spatial resolution cloud-free data is available in the early phase of Rabi season it can be integrated with multi-date AWiFS data for discriminating Rabi crops from long duration Kharif crops and plantations. When single date LISS-III data of November 22, 2011 was classified using MLC it yielded a Rabi crop sown area of 72.8 thousand hectares. However, analysis of full season multi-date data shows that the Rabi crop area of not more than 62 thousand hectare will be detectible by satellite sensor by November 22. It meant that the LISS-III based area was over estimated. Further investigation of the classified image showed that some area of long duration Kharif crops and some area of plantation got mixed with Rabi crop area to yield higher estimate. To solve the problem of this mixing the multi-date AWiFS data over the study area was integrated with the LISS-III data. In temporal domain the mixed classes became separable due to different temporal patterns of the Kharif crops / plantations and the Rabi crops. While Rabi crops clearly showed increasing NDVI patterns, Kharif crops and plantations showed decreasing NDVI patterns. Rajak and Jain (2016a) reported the primary results of this sub-study and demonstrated the integrated use of multi-source remote sensing data for early Rabi crop acreage estimation at district scale by using Resourcesat-2 LISS-III and AWiFS data.

Sub-study-B has produced another important result in two ways. One is that it has demonstrated the usefulness of crop history derived from previous multi-years data in estimating current year's crop area. Secondly, it has shown that the multi-source data can be effectively used for early estimation of Rabi crop sown area at state level. Two dates AWiFS data could not provide the distinct temporal patterns for identifying Rabi crops against the other classes but the addition of multi-date MODIS derived NDVI provided well separable temporal patterns. Even at this stage there were some clusters (belonging to crops) which were not showing considerable increasing NDVI trends. Such clusters were classified to crop if the crop history showed that they belonged to crop during previous all the 5 years and this year they are not showing decreasing NDVI trend. The decision fusion by employing decision tree approach was carried out to integrate the information derived from two dates AWiFS, multi-date MODIS and crop history data. The classification accuracies of the classified images were assessed with respect to the in-season field data collected during 2011-12. The estimated early crop sown area was compared with the government season-end estimates and the proportion of this crop area expected to be identified by satellite by mid-December (using full-season multi-date AWiFS data derived classified image). The results of the sub-study were published (Rajak et al., 2016b) to demonstrate the feasibility of integrated AWiFS and MODIS data for early estimation of Rabi crop sown area over a large area.

Sub-study-C demonstrated that multi-date coarse resolution data like MODIS can be used for multi-year crop assessment over a large area after suitable normalization of the data. It also demonstrated that HANTS can be successfully used for multi-date normalization of MODIS NDVI time series data which has been classified using an unsupervised classification. The results showed that there was overall increase in the Rabi crop sown area over Gujarat from 2002-03 to 2011-12. The annual rate of change of crop area has also been calculated. Crop specific annual crop area monitoring for two major Rabi crops (wheat and musterd) also has been presented.

The results of the three sub-studies were published (Rajak and Jain, 2016a; Rajak et al., 2016b; Sharma et al., 2014) and the feasibility of early estimation of Rabi crop area at district and state level using multi-source RS data was demonstrated.

7.3 Discussion to Address the Research Questions

This section is meant for discussing the answers obtained through this research with respect to the research questions (RQ are listed in Chapter 2, Section 2.5) and the subsequent results obtained to answer those RQs.

Discussion with respect to RQ1: The first stated research question was "what are the potential issues that affect integration of multi-source data for crop studies and how these issues can be resolved?" and a number of issues were identified and described in Chapter 3. In general, we can say that only comparable datasets can cohesively be combined to extract any integrated result in any multi-source data study. The datasets need to be compared with respect to their basic characteristics before integration e.g. what is the type of data (spatial or non-spatial, spatial raster or spatial vector etc.), in which format the spatial data is available (CEOS, GeoTIFF, HDF-EOS etc.), what is the spatial resolution of the data, which physical quantity is stored in the data (reflectance, radiance, emittance, digital number etc.), what is the radiometry of the data (8-bit, 16-bit signed/unsigned, float etc.) etc. Herewith, I list the major issues affecting multi-source data integration:

(i) **The multi-source data may or may not have geo-referencing information or the geo-referencing schemes of two or more datasets may be different**. Most of the remote sensing data have geo-referencing information either stored with the data in the same data file or in an associated file. If any data doesn't have proper geo-referencing information, it needs to be geo-referenced before its integration. Most of the image processing software packages have image-to-image and image-to-vector registration modules for this task. In the studies presented here, all the datasets were either already geo-referenced or geo-referencing metadata was available with the raw data; however with different geo-referencing schemes. All the datasets were accurately co-registered and re-projected to a common geo-referencing system. The details of the common referencing framework adopted in all the three sub-studies are given in Chapter 3, Section 3.5.

(ii) **The multi-source RS data are acquired by different sensors with different spatial resolutions**. In such a case, the spatial details of the objects contained in the datasets are also of different scales with different pixel sizes. For a good integration of datasets all the spatial details available should be preserved at integration level. It is a complex issue. If finer resolution data is resampled to coarser resolution data, the finer details are lost. If coarser resolution data is resampled to finer resolution, over sampling increases the data volume and hence computer storage and analysis time. Radiometry of the data also suffers when it is resampled. An optimum way of common spatial resolution needs to be adopted for the best use of information available at two different resolutions. In the studies presented here, the classified images obtained at coarser resolution were compared with the classified images obtained at finer resolution. In Sub-study-A, the classified image obtained through ISODATA clustering of multi-date AWiFS data was compared with the classified image obtained from single date LISS-III data. For analysing the temporal spectral profiles of the clusters, AWiFS data was accurately registered and resampled with LISS-III data. A procedure adopted for accurate registration of two images with different spatial resolutions is presented in Chapter 3, Section 3.6. In Sub-study-B, ISODATA clustering has been used to classify multi-date MODIS data and two-date AWiFS data separately. The Decision Rules based integration of these images and an a-priori image obtained through multi-year multi-date data has been done.

(iii) **The data available from different sources are acquired under different viewing and atmospheric conditions.** Different viewing conditions and varied atmosphere change the actual response of the object differently. In such cases the absolute data responses cannot be compared and direct integration of such data cannot be done. The data need to be normalised for different viewing and atmospheric conditions before integration of such data. If source-object-sensor geometry and composition of atmospheric constituents at the time of RS data acquisitions are known, the dataset can be normalised using Radiative Transfer (RT) models. Getting source-object-sensor geometry is a relatively easier task but obtaining composition of the atmospheric constituents at the time of data acquisitions is difficult. In case of historical data, you may not get any information on atmospheric composition. The other way to normalise such multi-source data is to empirically normalise the datasets based on the features available in the images. It is called image based normalisation. Image based Haze Optimisation Transformation (HOT) and Harmonic ANalysis of Time Series (HANTS) are described in Chapter 3, Section 3.7.

(iv) **The multi-source data comprises of different physical quantities of the target i.e. spectral radiance, reflectance, emittance, backscattering coefficient etc.** Direct comparison of such datasets cannot be done hence direct integration of such data also cannot be made. In case of the Sub-study-B, the multi-date and multi-year MODIS data gave the surface reflectance value of the object while AWiFS data gave spectral radiance values (converted from the Digital Numbers). These two physical quantities cannot be directly compared, hence NDVI (Normalised Difference Vegetation Index) derived from the two quantities were used in the study. The NDVI is a normalised ratio of the two related variables (radiance and reflectance), it may be considered comparable especially in case of clear atmospheric conditions.

(v) **The raster data are stored in different formats** e.g. CEOS (Committee on Earth Observing Sensors), TIFF (Tagged Image File Format), GeoTIFF (Geographic TIFF), IMG (Image format by ERDAS), HDF 4/5 (Hierarchical Data Format), HDF-EOS (HDF Earth Observing System), MrSID (Multi-Resolution Seamless Image Database), NetCDF (Network Common Data Form) etc. Similarly vector data is also available in different formats. Data available in different formats have to be brought in a common format. ERDAS Imagine image (.img) was taken as the common data format because most of the image analysis was carried out using ERDAS Imagine software.

(vi) **The types of spatial data are different.** For example, crop area estimation using current year's RS data and land use land cover vector layer of the study area available from any other source is a difficult task to do. The two datasets are spatial in nature but the types are different. Either the vector layer needs to be rasterised or RS data need to be vectorised before using them synergistically. If a software package accepts two different types of data as input, it converts internally one data into other data type and does the analysis accordingly. In such cases, the raster and vector data characteristics need to be taken into account before integration.

(vii) **The types of available data or information are different**. For example, while analysing RS data for crop area estimation over a village or taluk the information about the excess/deficit of rainfall cannot be directly used because the two things are not of similar types. The RS data being spatial in nature and rainfall excess / deficit information being non-spatial in nature, their combination is not a state forward process.

Discussion with respect to RQ2: The 2nd research question was whether a methodology can be developed to use multi-source data for early crop sown area estimation at district level and state level. Estimating early crop sown area at any scale has remained a challenging task for long time and has not been attempted much. The availability of high frequency RS data from GISAT over India will offer an opportunity for early estimation of crop sown area. In the present case, a study was taken up for an early estimation of Rabi crop sown area for Mehsana district in Gujarat (India) for 2011-12. The data from Indian satellite sensors Resourcesat-2 AWiFS (multi-date) and Resourcesat-2 LISS-III (single date) were used in the study.

A major problem encountered while classifying LISS-III data of November 22, 2011 was mixing of the two vegetation classes i.e. crops and plantations/long duration Kharif crop. The spectral values from the two classes, as measured by LISS-III were found to be closure to each other in all the four bands (Red, Green, NIR & SWIR). The NDVI values for the two classes were also in the same range (Absolute NDVI values between 0.34 and 0.39). This correspond to the period when Rabi crops are in growing stage and most of the plantations and long duration Kharif crops (cotton and castor) are in declining stage of vegetation vigour. Resourcesat-2 AWiFS data is available every 5-day over an object in India. It was found that AWiFS data over the study area was available for three dates before November 22, 2011 (the acquisition date of LISS-III) i.e. October 10, October 24, and November 11, 2011. The NDVI values calculated from single date LISS-III and three dates AWiFS data provided the temporal spectral profiles at the resolution of LISS-III data (AWiFS data was re-sampled at LISS-III spatial resolution). The temporal spectral profiles of the four classes which were not discriminable at 4-bands LISS-III spectral plot became separable when 4-dates NDVI values were plotted.

The ISODATA based classified images obtained from single date LISS-III and 4-dates integrated data were compared for classification accuracies. The problem of mixing of crop and plantation/Kharif crops classes observed in LISS-III data was resolved with use of integrated dataset. The estimated Rabi crop sown area using LISS-III and integrated data were found to be 72.8 ‘000 ha and 58.9 ‘000 ha, respectively. Crop calendars of two major Rabi crops of the district i.e. wheat and mustard show that most of the mustard crop was sown and should be visible to satellite sensor by November 22 but a large fraction of wheat crop was either not sown or not visible to the sensor by that date. To assess the accuracy of these estimates, a multi-date LISS-III and multi-date AWiFS data (full season data) based classification was performed, separability analysis was carried out, and classified image was obtained. This full season classified image was considered as the reference image for the accuracy assessment of the classified images obtained from limited period datasets for early estimation. The full season data analysis showed that not more than 62 ‘000 ha of combined crop sown area of wheat and mustard crops should be identifiable by the satellite by November end. In the light of this result, the early crop sown area (58.9 ‘000 ha) assessed by integrated dataset upto November 22, 2011 is considered close to the expected value (62 ‘000 ha). The results of this study answered the research question posed by RQ2.

While first part of RQ2 dealt with district scale early crop sown area, the second part posed the question whether a methodology can be developed for early crop sown area at state level. At state level, one cannot expect even a single coverage of cloud-free LISS-III data at an early stage of crop growth. Only wide swath sensor can cover the large study areas frequently. As stated earlier, Gujarat state was selected to develop the methodology of early Rabi crop estimation using multi-date 2011-12 AWiFS and MODIS data. Two-dates AWiFS data and multi-date MODIS data for 2011-12 along with the crop history information derived from previous 5-years MODIS data were used to early estimate the crop area. In such cases, a large number of feature classes due to spatial and temporal variation over large area usually prohibit the use of supervised classifiers. Training of supervised classifiers would require a large number of training data sets that will not be available for such a large area especially in an early stage of crop growth. Hence, ISODATA clustering, an un-supervised approach of RS data classification was used to classify integrated dataset prepared from AWiFS and MODIS data. The ISODATA clusters identical to reference crops temporal patterns were classified to Rabi crops but there were some clusters that did not show the distinct increasing growth patterns. The integration of crop history information to classify such clusters into crops and non-crops classes through decision tree approach was found very useful.

Discussion with respect to RQ3 and RQ4: The 3rd research question posed was whether a methodology can be developed to use the information derived from previous multi-years RS data for current year‘s crop sown area estimation. To get an answer of this question, firstly the crop history information was extracted using multi-date 5-years MODIS data (2006-07 to 2010-11). Then an approach was developed to integrate this information with the information obtained from in-season RS data (2011-12). This integration helped in identifying the crop clusters which were not distinctly showing increasing NDVI patterns in the in-season temporal profiles. On comparing such clusters with the full season multi-date AWiFS based classified image, it was found that the crop history information was useful in identifying crop clusters for early estimation of crop sown area. Further details are already given in Chapter 5, Sections 5.3 and 5.4.

To answer the fourth and the last research question whether coarse resolution RS data like MODIS can be used for assessing multi-year Rabi crop area changes at state level, a sub-study was carried over Gujarat state. The radiometrically normalized multi-date NDVI data could classify the Rabi crops. The crops estimates show that the moderate spatial resolution data with high temporal frequency could be used for monitoring year-after-year changes in rabi crop area. The results and discussion on the sub-study are presented in Chapter 6, Section 6.3.

This page is intentionally left blank.