Food and Agriculture Organization
of the United Nations

Mid report on the project

IMPROVE AGRICULTURE MONITORING SYSTEMS THROUGH SATELLITE IMAGERY FOR THE ISLAMIC REPUBLIC OF **IRAN**

Mid report on the project

IMPROVE AGRICULTURE MONITORING SYSTEMS THROUGH SATELLITE IMAGERY FOR THE ISLAMIC REPUBLIC OF **IRAN**

Food and Agriculture Organization of the United Nations
Rome, 2019

Required citation:
FAO. 2019. *Improve agriculture monitoring systems through satellite imagery for the Islamic Republic of Iran.* Rome.

ISBN 978-92-5-131965-9

Cover photograph:
©FAO/Shay Mazloom

CONTENTS

PART 1: Project objective and methodology

PART 2: Current data collection practices by Ministry of Agriculture Jihad - Approaches, gaps and limitations of current procedures

PART 3: User expectations of a new data collection method - Identification of areas for improvements in methodology adopted at contry level and in the area frame

PART 4: Application of remote sensing in agricultural data collection

PART 5: Previously implemented pilot project in Hamedan Province - Description of the area frame test to date. Analysis of its gaps and limitations

Figures

ACRONYMS

Acronym	Description
AEZ	Agro-ecological zone
AMS	Agriculture monitoring system
AOI	Area of interest
ASF	Area sampling frame
ASIP	Agricultural survey improvement programme
BOA	Bottom of the atmosphere
CG	Center of gravity
CGMS	Crop growth monitoring system
CRS	Coordinate reference system
CV	Coefficient of variation
DEM	Digital elevation model
EC	Euclidean distance
ESA	European space agency
FAO	Food and Agriculture Organization of the United Nations
GDB	Geo-database
GIS	Global information system
GSD	Ground sampling distance
Ha	Hectare
HR	High resolution
IC	Image classification
ITC	International institute for aerospace survey and earth science
LC	Land cover
LCCS	Land cover classification system
LF	List frame
LU	Land use
MAJ	Ministry of Agriculture Jahad
MLC	Maximum likelihood classification
MVC	Maximum value composite
NDVI	Normalized difference vegetation index
NIR	Near infrared
PJAO	Provincial Jahad Agriculture Organization
PSU	Primary sampling unit
QGIS	Quantum geographic information system
RD	Rooting depth
SAM	Spectral angle mapper
SCI	Statistical Center of Iran
SNAP	Sentinels application platform
SPG	Soil physical group
SSU	Secondary sampling unit
SUs	Simulation units
TCP	Technical cooperation project
TOA	Top of atmosphere
UTF	Unilateral trust fund
UTM	Universal transverse mercator


© FAO

ACKNOWLEDGEMENTS

This report pertains to the "Improved agriculture monitoring systems through satellite imagery project for the Islamic Republic of Iran". This is a pilot project conducted in 2017 for three provinces/regions of Iran: Mazandaran and Zanjan provinces and South of Kerman region.

We acknowledge with gratitude the guidance, technical inputs and support extended by Elisabetta Carfagna and Gianluca Franceschini and their colleagues to formalize the methodological aspects of using geospatial technology to produce reliable and timely agricultural statistics based on area frame and satellite data analysis.

It is hoped that this report will help to strengthen and improve the agriculture monitoring system based on integral use of advanced geospatial technologies to support the development of the techniques, policy and investment conditions to achieve sustainable agricultural development under climate change.

EXECUTIVE SUMMARY

To implement the project "Improved agriculture monitoring systems through satellite imagery for Iran", three parts of the country were selected as pilot areas: Zanjan and Mazandaran provinces, and south Kerman region. For each area, some basic information was obtained, including crop calendars, statistics on the status of major agricultural products, political boundaries, land use and land cover maps and climatic agro-ecological zone (AEZ) maps.

In the next stage, Sentinel-2 level -1C satellite images of pilot areas were downloaded for the period from 1 January 2017 to 31 December 2017. Pre-processing (i.e. atmospheric correction) was performed on downloaded images using Sentinels application platform (SNAP) software.

The normalized difference vegetation index (NDVI) was then calculated for each image for 12 months of the year. Following that, the average monthly NDVI was assigned to each segment that had been provided by an international consultant from the FAO Geospatial Unit (CBDS). An up-to-date land use/land cover map was needed for designing sample units; therefore, using the NDVI trend review over one year and setting the threshold limits for NDVI values, these land use/land cover maps were generated. The maps include orchards, irrigated lands, water bodies, rangelands and bare-area classes for the three pilot sites.

After visually checking maps based on expert knowledge and with comparisons to Google Earth images, incorrectly classified segments were refined. The land cover layer was re-classified, focusing on agricultural classes. All non-agricultural classes were classified as 'no agriculture', and the agricultural layer was generated. Areas with where no agriculture were removed from the agricultural layer. Roads were included since these are mainly linear features.

These maps, along with climatic AEZs were used for stratification and sample allocation and selection. Regular grids (fishnets) at increasing sizes (20 ha, 30 ha and 40 ha) were generated to provide an idea of the average number of parcels per segment to be surveyed on the ground. This was because no previous area frame had been implemented in the pilot areas.

To account for different conditions affecting crop growth, a layer of climatic AEZs previously obtained was combined with the agricultural land cover layer to generate the stratification of the three pilot areas by intersect function in GIS software. Then, the intersection of the fishnet of the agricultural land cover layer with the layer of the climatic AEZ was performed. The intersection was imported into Excel software, unnecessary attributes were removed and the area for the intersected polygons was calculated. The grid with cells (segments) of 20 ha size was chosen for collecting ground data in the current test for Kerman and Mazandaran areas and 30 ha was chosen for Zanjan province. That size proved to be the most suitable to balance the average size of the parcels with the estimated effort for data collection (to be tested through the ground survey).

To avoid biased estimates, each segment with even a small portion (at least 5 percent) of agricultural area was included in the population. Several segments presented parts of their area covered by more than one stratum given by the intersection of land cover and agro-ecological zones. The percentage of the different strata in each segment was computed and the segment was attributed to the stratum with the highest percentage. Within each stratum, the segments were grouped on the basis of the percentage of the area of the different strata in the segment.

Consequently, the final stratification was based on land cover, agro-ecological zones and percentage of agriculture. Since no previous information on the variability inside each stratum was available, the proportional allocation was adopted; that is, the total number of segments to be selected was subdivided among the final strata according to the stratum size, provided that at least two segments were selected in each stratum for computing the variance of the acreage of crops in each stratum.

The sample selection was made using the permanent random numbers. In each final stratum, a random number was assigned to each segment in the population of segments. The segments were then ordered according to those numbers. This procedure was conducted in each final stratum, independently; and corresponds to the sample selection.

Then, the first subset of segments corresponding to the number of segments allocated to each stratum became part of the group of segments to be surveyed on the ground. This procedure allowed for performance of the ground survey on additional segments in a stratum, if needed; following the order of the segments generated by the ordered random numbers and treating all surveyed segments as one-step simple random samples.

For each segment, a form was pre-filled with universal transverse mercator (UTM) coordinates, final stratum ID and name, province, county and a screenshot from Google Earth to support parcel identification. Pilot samples of 90,72 and 68 segments were selected for ground truth collection from Zanjan, Mazandaran and South Kerman, respectively. Training workshops on how to collect segment information was conducted in three pilot areas for the enumerators from MAJ. Then, the field data collection was done by them. Segments' ground information was then imported to Excel and ArcGIS environments for statistical analysis and satellite image processing.

Once the first ground survey has been conducted, collected data can be used for computing the variability inside each final stratum. That allows estimates of the sample size needed for obtaining estimates with prefixed accuracy and performing the optimal allocation for the next surveys.

Due to the adoption of the permanent random numbers, there is no need to select ex novo all the segments in each stratum. The additional segments can be selected following the order of the permanent random numbers. Thus, the segments selected for the first survey and the additional ones can be treated as just one simple random sample inside each stratum. Consequently, the parcel identification already performed on the screenshots from Google Earth for the first survey can be used for the next surveys.

In this project, satellite image analysis was also performed (including classification and NDVI) for the review and assessment of the existing agriculture monitoring procedures, Moreover, geospatial technology is exploited at the design level (for efficient ground data collection) as well as at the estimation and monitoring levels for developing an improved, innovative, up-to-date and complementary agriculture monitoring systems through the integration of satellite remotely sensed data.

PART 1

Project objective and methodology

© FAO/H. Null

INTRODUCTION

An agriculture monitoring system (AMS) has become a priority of the Ministry of Agriculture Jahad (MAJ) of the Islamic Republic of Iran. This is due to a number of environmental trends which suggest an urgent need for a comprehensive, systematic and accurate agricultural monitoring system. More frequent extreme climate events, such as floods, drought and frosts, are adversely affecting agricultural production in the country. Changes in precipitation amounts, seasonality, intensity and distribution, are impacting rainfed agriculture, and warming temperatures are changing growing seasons.

In the face of this environmental complexity, MAJ asked the Food and Agriculture Organization of the United Nations (FAO) to provide assistance in setting up an improved agriculture monitoring system, based on integral use of advanced geospatial technologies to support development of the techniques, policy and investment conditions to achieve sustainable agricultural development under conditions of climate change.

It was agreed that in the first stage, FAO would provide technical assistance for the technical cooperation project (TCP) under the title, "Improved agriculture monitoring systems through satellite imagery for Iran".

This TCP is seen as a precursor to a larger unilateral trust fund (UTF) project that will focus on applying the methodological approaches developed under the current project for country crop monitoring, area estimation and yield forecasting.

Likewise, the TCP has focused on identification of state-of-the-art methods and strategy for acreage and yield estimation, based on an assessment of the existing monitoring methodology and optimized through the use of remote sensing. It will identify sustainable methodology for field data collection and approaches to produce improved, harmonized and consolidated statistics, utilizing methodology that leverages data collected by provincial agriculture offices. In addition, the project will benefit from the availability of multi-temporal satellite images for testing and monitoring of a range of crops for the selected areas as a demonstrator for the larger UTF.

Estimation through sampling in the field will be carried out to identify sites for crop area and yield estimates. The employment of multi-temporal high-resolution satellite images will enable MAJ to collect near-real time crop field information and use it operationally for the monitoring.

1. PILOT PROJECT AREA

The pilot areas should demonstrate the geographical and natural characteristics of different conditions of entire country. The selection should be based on factors such as:

- availability of already prepared land use/land cover maps
- new, large-scale aerial photos
- crop variety
- average size of crop fields
- appropriate cooperation of provincial Jahad-agriculture organizations
- diversity of climatic conditions, etc.

The project area includes Zanjan and Mazandaran provinces and South Kerman region shown in Figure 1 with their locations and false color images.

FIGURE 1 - Iran project study area

2. OBJECTIVES

The primarily objective of the project is to establish a prototype demonstrator operational agriculture monitoring system in a limited number (three) of priority provinces. This prototype will also improve quality of agriculture information and reporting, based on geospatial technology, with the expected impact of adoption of improved strategies for increasing and diversifying production potentials by the government. As well, it will focus on identifying state-of-the-art methods and strategies for acreage and yield estimation, based on assessment of the existing monitoring methodology and optimized through the use of remote sensing.

The following executive objectives were considered:

- review existing survey methods for agricultural statistics to identify gaps and limitations;
- transfer relevant methods, good practices and learning materials;
- distance learning for land cover mapping, agriculture monitoring and crop area and yield estimation;
- improved provincial capacity to develop reports/ bulletins based on geospatial technologies;
- exchange of knowledge, technology and experiences.

3. METHODOLOGY FOR CROP AREA ESTIMATION

3.1. Land use/land cover map production

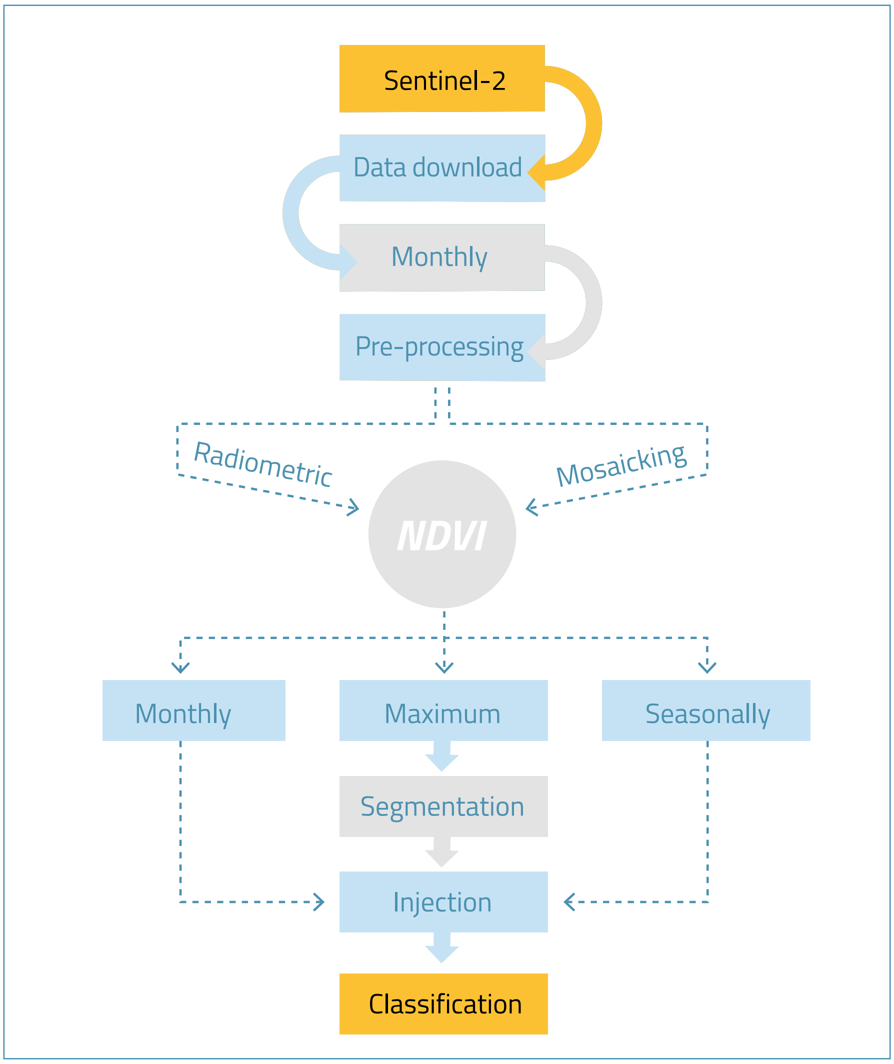
The main goal of the FAO classification system approach is to discriminate among different kinds of land covers and agricultural areas with acceptable accuracy. This is based on monthly NDVI change thresholds during a year, because NDVI values show the vegetation density in farmlands and forests. In other words, higher densities lead to higher NDVI and the reverse. In fact, it enables specialists to detect the pheonological cycle of crops, orchards and other land covers through the satellite imagery. According to this, monthly NDVI reflects the growth cycle and helps to discriminate among different kinds of vegetation.

Remote sensing land cover mapping methodology based on the FAO standard approach includes several steps:

- download Sentinel-2 monthly data;
- pre-process each image radiometrically;
- compute NDVI for each month;
- make a segmentation layer based on maximum vegetation in study area;
- insert monthly NDVI values from images to the segmentation layer;
- define NDVI thresholds for various classes.

3.2 Sentinel-2

FIGURE 2 - Land cover classification workflow



Several remote sensing satellites are useful for land use and agricultural mapping. However, Sentinel-2 is more relevant, because it provides many multi-spectral bands in red and near-infrared portions which are very suitable for computing NDVI. Sentinel-2 also has a 10 meter spatial resolution, which provides more spatial details for land-use mapping. Higher spatial resolution provides better border detection for different classes and it improves the accuracy of estimated area for croplands and orchards.

Sentinel-2 data is distributed as free, remote sensing data at two levels:

1) Level-1C (L1C)
2) Level-2A (L2A)

L1C is composed of 100 km^2 tiles (ortho-images in UTM/WGS84 projection), and L1C product results from using a digital elevation model (DEM) to project the image in cartographic coordinates.

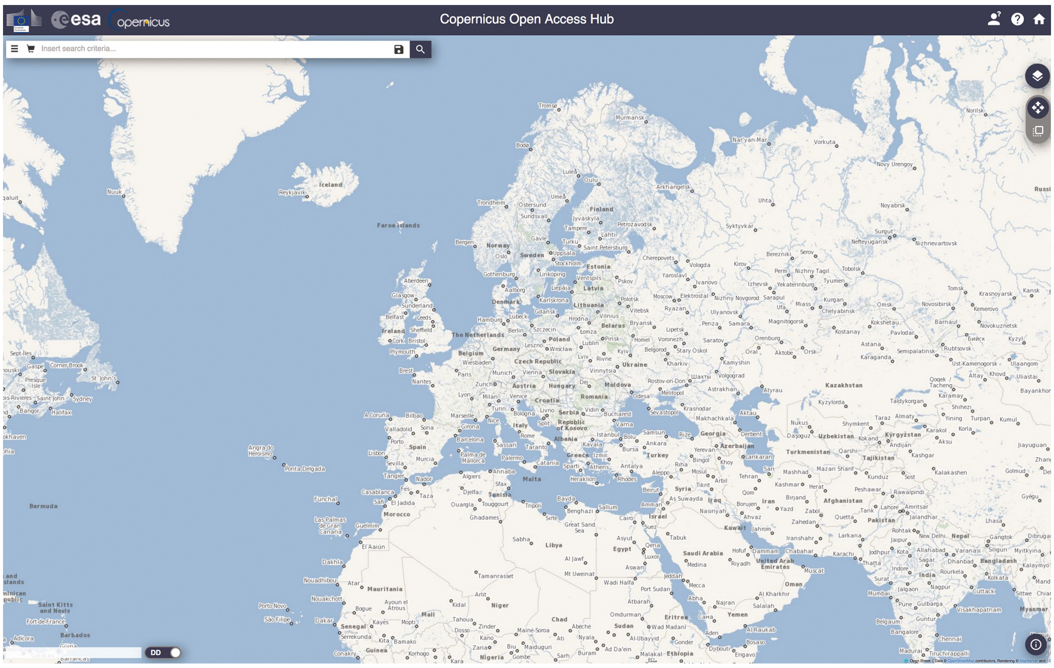
Per-pixel radiometric measurements are provided in top of atmosphere (TOA) reflectance with all parameters to transform them into radiances. L1C products are re-sampled with a constant ground sampling distance (GSD) of 10 m, 20 m and 60 m, depending on the native resolution of the different spectral bands.

Table 1 - Sentinel-2 bands properties

Bands	Name	Wavelength
1	Coastal aerosol	442.7
2	Blue	492.4
3	Green	559.8
4	Red	664.6
5	Vegetation red-edge	704.1
6	Vegetation red-edge	740.5
7	Vegetation red-edge	782.8
8	NIR	832.8
9	Narrow NIR	864.7
10	Water vapor	945.1
11	SWIR-cirrus	1 373.5
12	SWIR	1 613.7
13	SWIR	2 202.4

L2A product also provides bottom of atmosphere (BOA) reflectance images derived from the associated L1C products. Therefore, each L2A product is also composed of 100 km² tiles in cartographic geometry (UTM/WGS84 projection). In this project, Sentinel-2 LIC was used for land use and land cover mapping since it covers the entire area. Sentinel-2 data supplied by European Space Agency (ESA) through Copernicus sentinel data hub registration[1]. Monthly Sentinel-2 data (L1C format) was downloaded for each province. After that, a pre-processing step started for correcting data atmospherically and mosaicking geometrically with SNAP software.

FIGURE 3 - Sentinel-2 data hub (Copernicus website)



1 https://scihub.copernicus.eu/dhus/#/home

FIGURE 4a - Monthly false color composite of Zanjan province

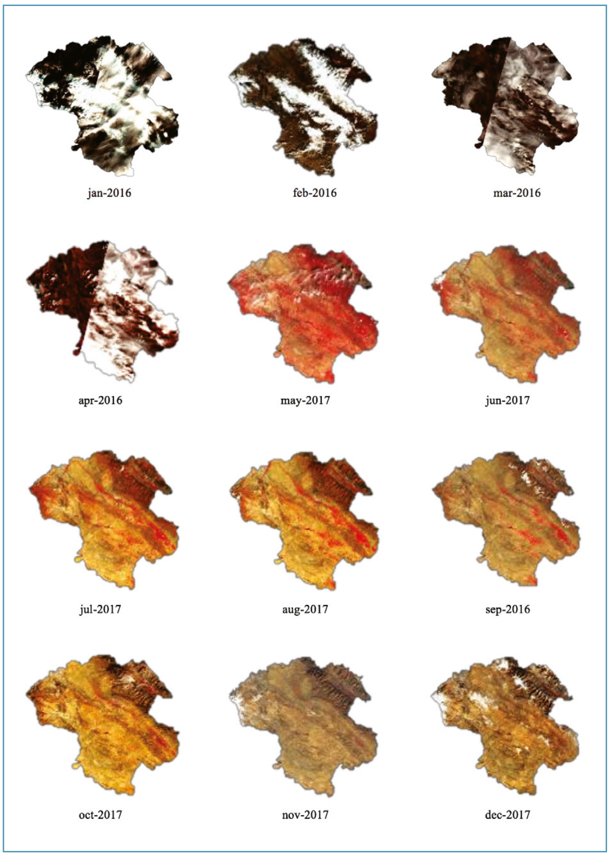


FIGURE 4b - Monthly false color composite of South Kerman

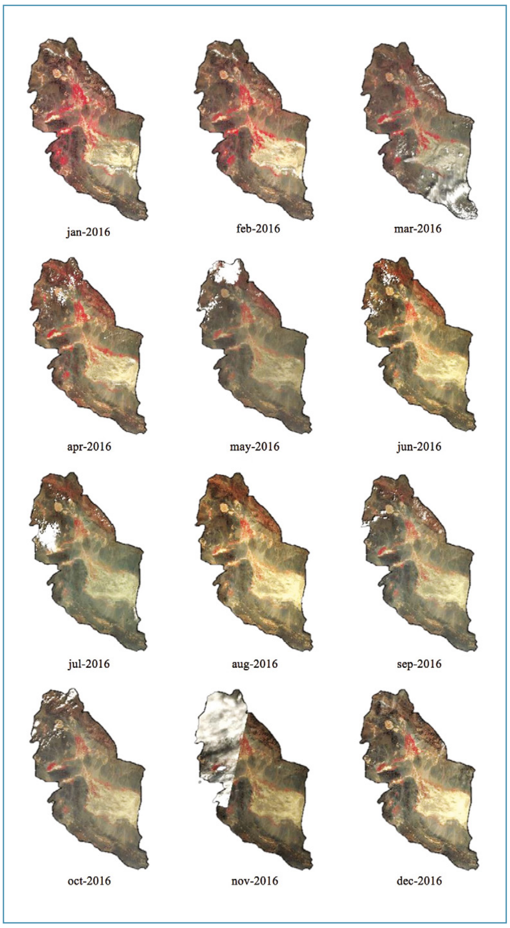
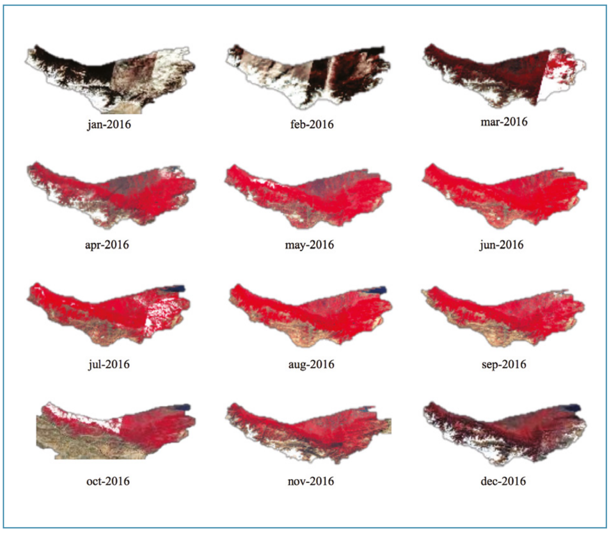


FIGURE 4c - Monthly false color composite of Mazandaran province

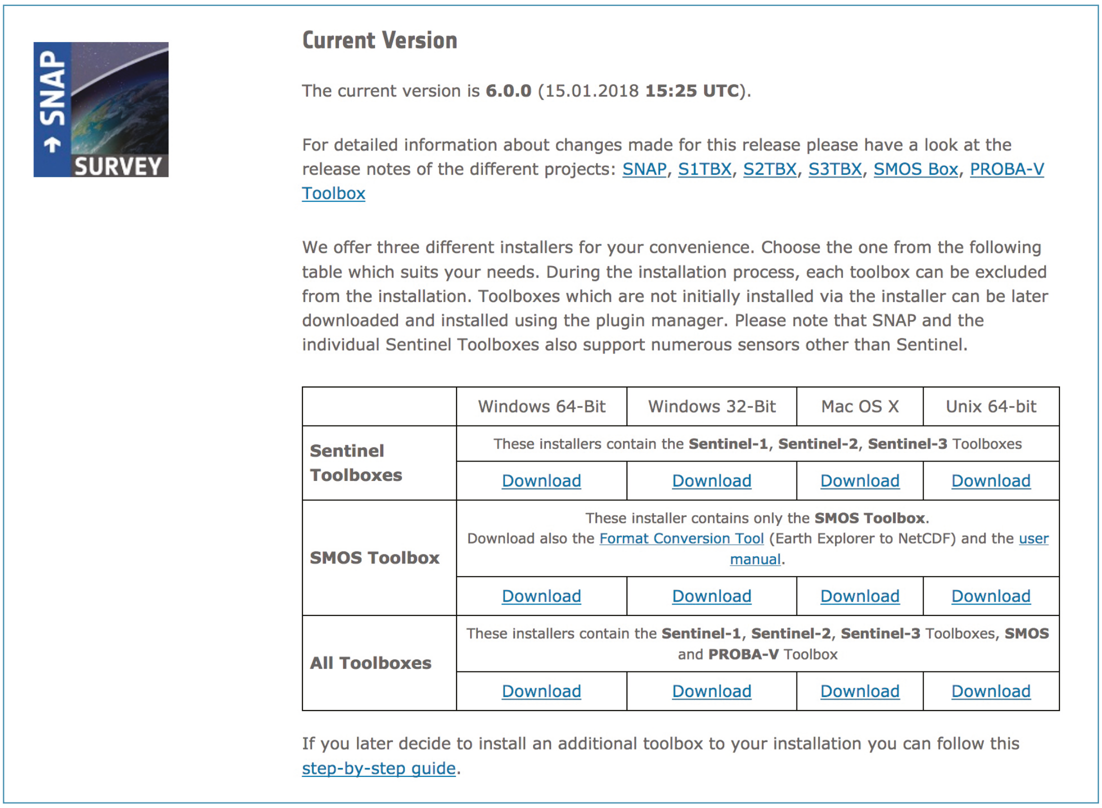


3.3 SNAP software

Sentinels application platform or SNAP is a special software for the Sentinels series and RADAR data. This is a free software for all platforms provided by ESA[2]. Sentinel-2 data correction and mosaicking for the three provinces was done in SNAP. After that, the NDVI was computed for each month using this software.

Note: Sentinel-2 data download, corrections, mosaicking and index calculations can be done using Google Earth Engine (another open source hub for remote sensing data processing). It works very quickly with a cost/benefit method as an alternative methodology.

FIGURE 5 - SNAP software



Current Version

The current version is **6.0.0** (15.01.2018 **15:25 UTC**).

For detailed information about changes made for this release please have a look at the release notes of the different projects: SNAP, S1TBX, S2TBX, S3TBX, SMOS Box, PROBA-V Toolbox

We offer three different installers for your convenience. Choose the one from the following table which suits your needs. During the installation process, each toolbox can be excluded from the installation. Toolboxes which are not initially installed via the installer can be later downloaded and installed using the plugin manager. Please note that SNAP and the individual Sentinel Toolboxes also support numerous sensors other than Sentinel.

	Windows 64-Bit	Windows 32-Bit	Mac OS X	Unix 64-bit
Sentinel Toolboxes	These installers contain the **Sentinel-1**, **Sentinel-2**, **Sentinel-3** Toolboxes			
	Download	Download	Download	Download
SMOS Toolbox	These installer contains only the **SMOS Toolbox**. Download also the Format Conversion Tool (Earth Explorer to NetCDF) and the user manual.			
	Download	Download	Download	Download
All Toolboxes	These installers contain the **Sentinel-1**, **Sentinel-2**, **Sentinel-3** Toolboxes, **SMOS** and **PROBA-V** Toolbox			
	Download	Download	Download	Download

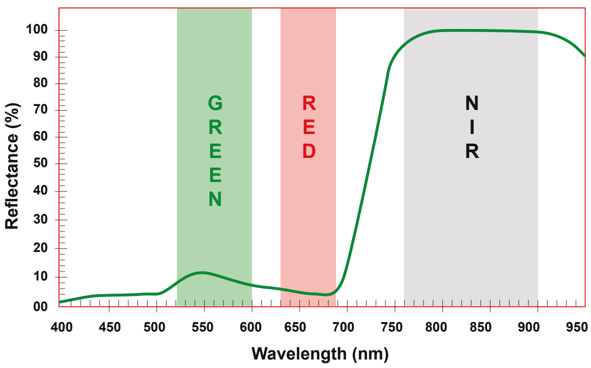
If you later decide to install an additional toolbox to your installation you can follow this step-by-step guide.

3.4 Methodology: normalized difference vegetation index (NDVI)

Normalized difference vegetation index (NDVI), a remote sensing global index for vegetation mapping through satellite imagery, is calculated by a normalization equation between red and near-infrared bands that is affected by chlorophyll content in vegetation. Higher chlorophyll leads to higher reflectance in near infrared (NIR) and absorption in red bands.

Healthy vegetation spectral behaviour is shown in Figure 6.

FIGURE 6 - Healthy vegetation spectral behaviour



[2] http://step.esa.int/main/download/.

As shown in Figures 7a, 7b and 7c, there is clearly a different reflectance between NIR and red portions, which is caused by chlorophyll. Higher reflectance objects appear lighter, and lower reflectance appears darker. According to this NDVI equation, vegetation appears as black in the red band; and as white in NIR bands, which makes vegetation detection easier.

Normalized difference vegetation index for Sentinel-2 data computed with this equation:

$$NDVI\ (Sentinel\text{-}2) = \frac{B8 - B4}{B8 + B4}$$

where
$B8$ indicates RED Band and $B4$ indicates NIR Band.
The NDVI values range is distributed from -1 to +1.

Negative values show water and moisture, zero shows non-vegetated lands and higher values up to 1 show vegetation with different densities. For instance, 0.5 and 0.9 values show vegetation cover, but some areas with 0.9 are denser than other areas with 0.5 values. As shown in Figure 7b and Figure 7c, differences between farmlands in winter and summer are distinguishable by NDVI changes; for example, winter NDVI shows decreases and in summer, increases.

FIGURE 7a - NDVI in Mazandaran province

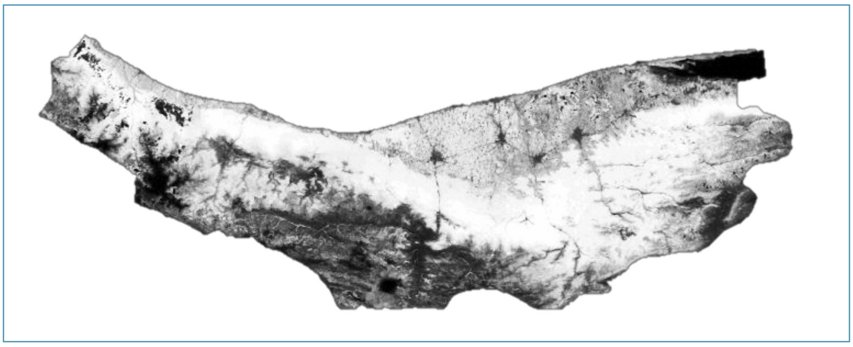
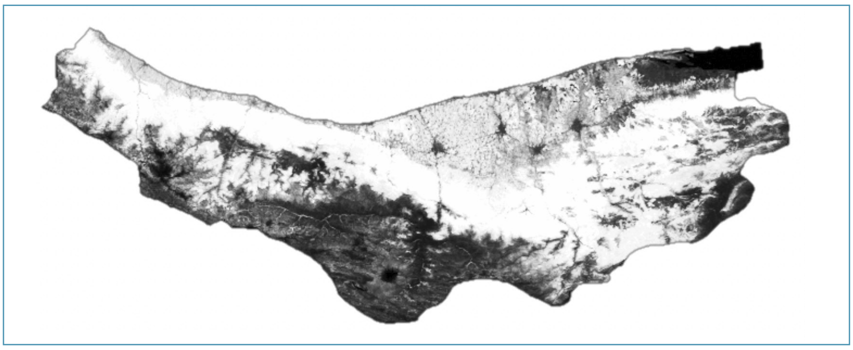


FIGURE 7b - Farmlands in winter (Mazandaran province)

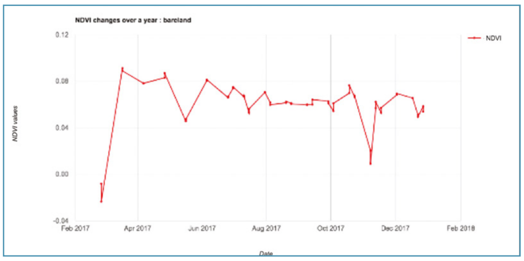
FIGURE 7c - Farmland in summer (Mazandaran province)



3.5 Methodology: normalized difference vegetation index thresholds

As mentioned previously, NDVI changes over a year show different phenomena. For example, in bare land areas the NDVI threshold changes as shown in Figure 8:

FIGURE 8 - NDVI changes over a year: bare land



The NDVI range changes between 0 and 0.2. However, orchards appear differently. For example, orchard NDVI values – shown in Figure 9 and its range – between 0.3 and 0.8 because the constant presence of trees has led to higher values for this class across all seasons.

FIGURE 9 - NDVI changes over a year: orchard

FIGURE 10 - NDVI changes over a year: rangeland

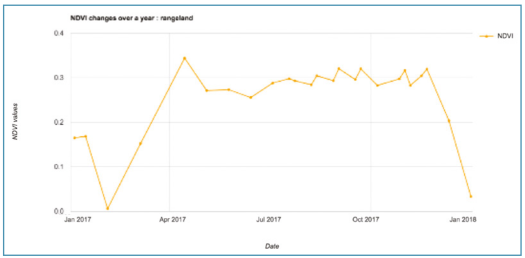
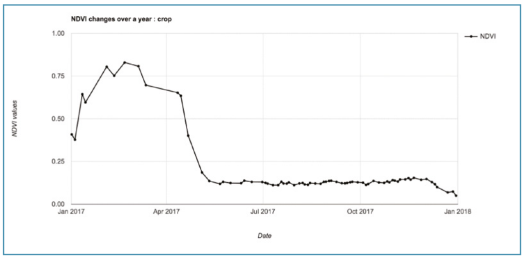


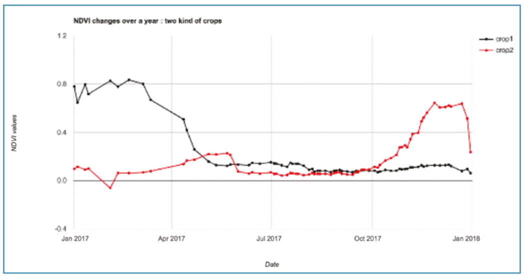
FIGURE 11 - NDVI changes over a year: crop



Compared with other classes, NDVI behaviour in crop lands fluctuates significantly because vegetation is present at one time but absent at another time. So, in the presence of vegetation, NDVI increases and in its absence, decreases. This assumption of different phenological cycles helps us in separating different kind of crops.

To detect different crop types, a crop calendar is essential. For example, in Figure 12, two kinds of crops were compared with NDVI behaviour over a year; however, the crop calendar tells us which crop will grow in winter and which in autumn.

FIGURE 12 - NDVI changes over a year: two kind of crops



After the NDVI trend is extracted, it is matched with a crop calendar as ground reference data. Crop calendars provide useful data for crop monitoring, particularly for rice, which is a strategic product in Iran's irrigated lands. Crop calendars tell us which product grows at what time in what province.

See the crop calendar for each of the 3 pilot areas in the next pages.

CROP CALENDARS

Table 2 - South Kerman crop calendar

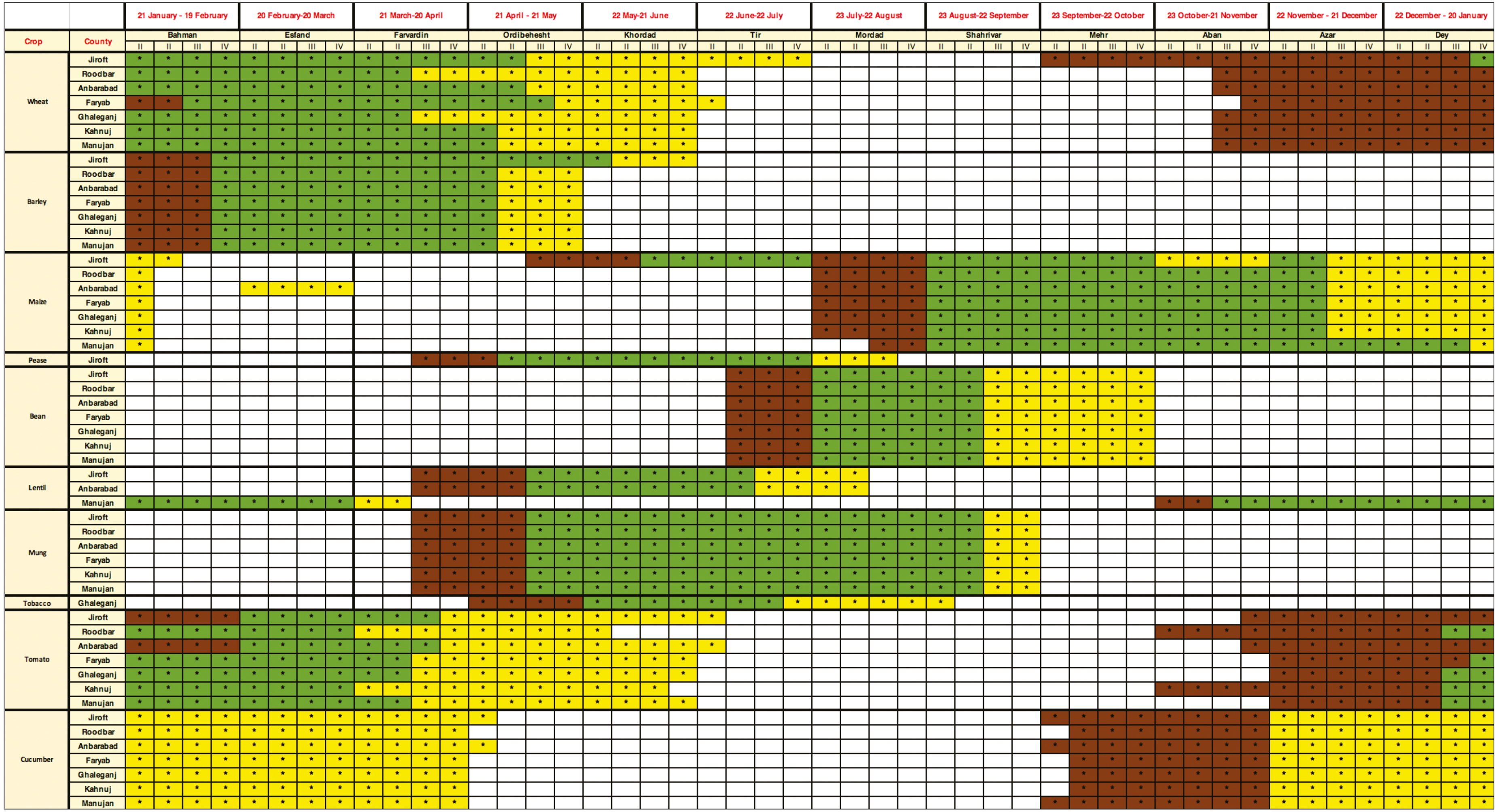
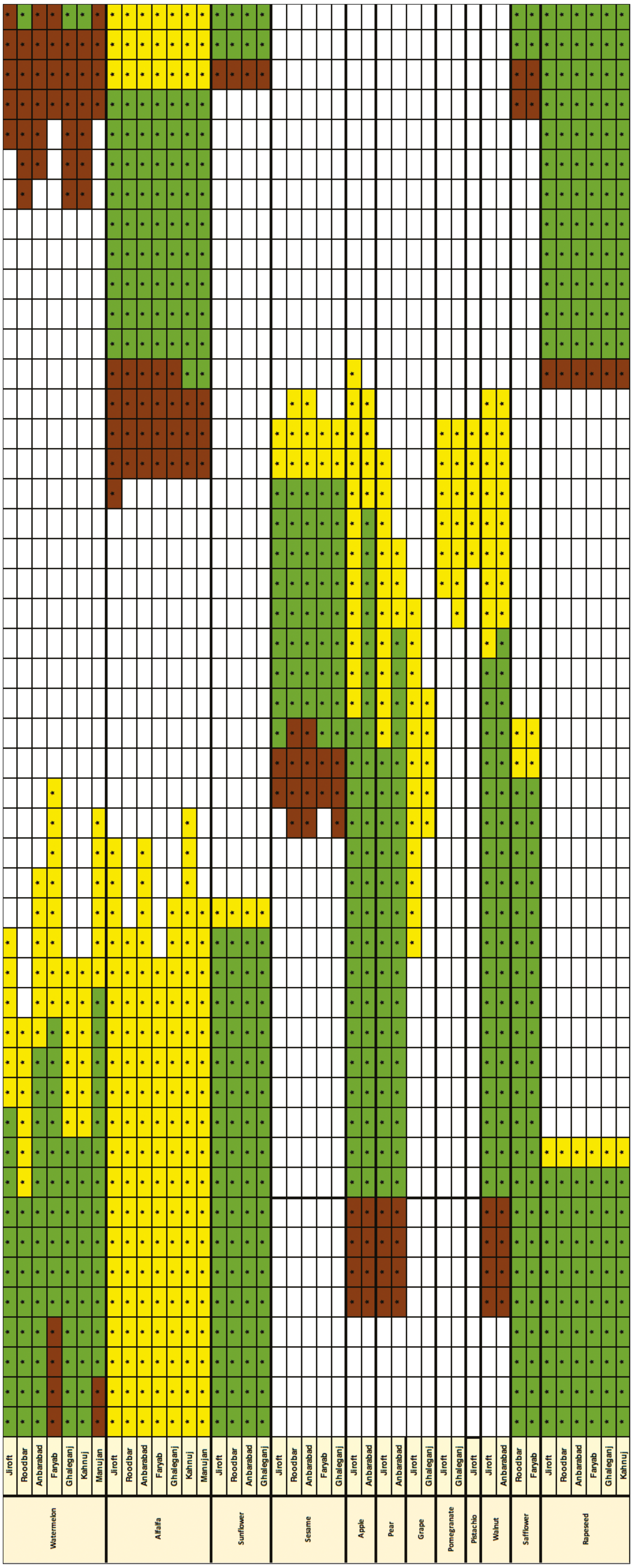


Table 2 - (continued)



LEGEND

- Sowing
- Growth
- Harvesting

CROP CALENDARS

Table 3 - Zanjan crop calendar

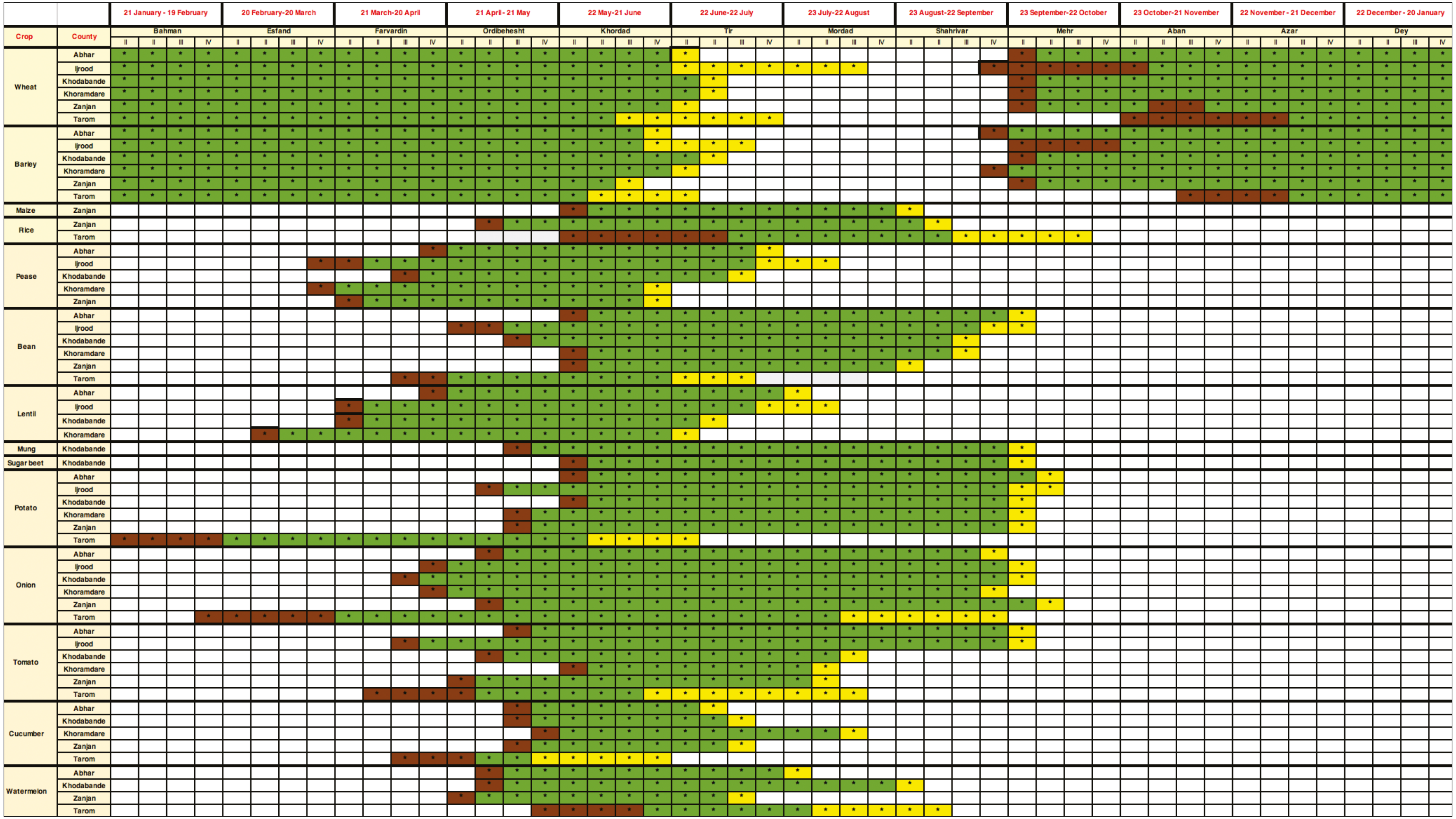
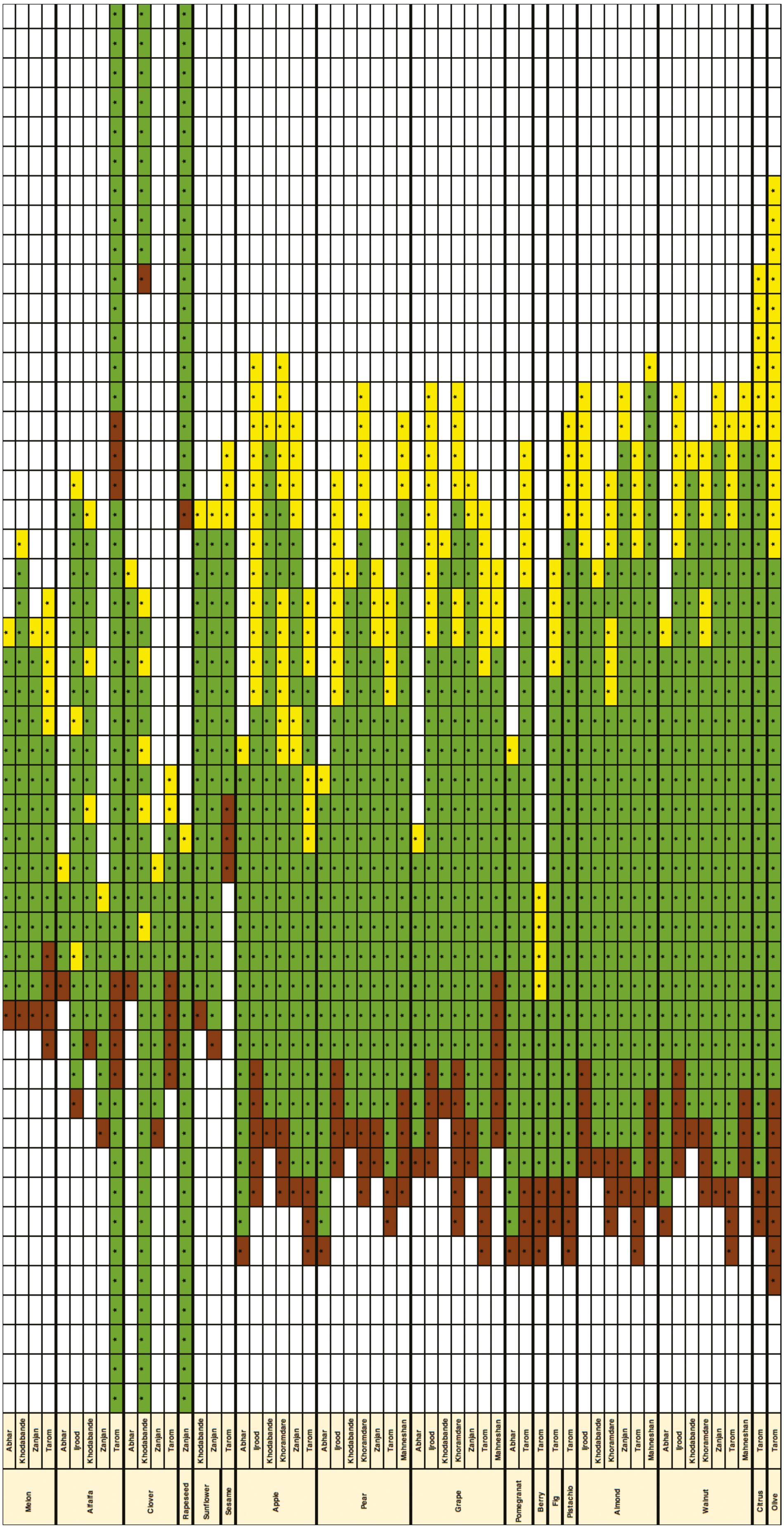


Table 3 - (continued)

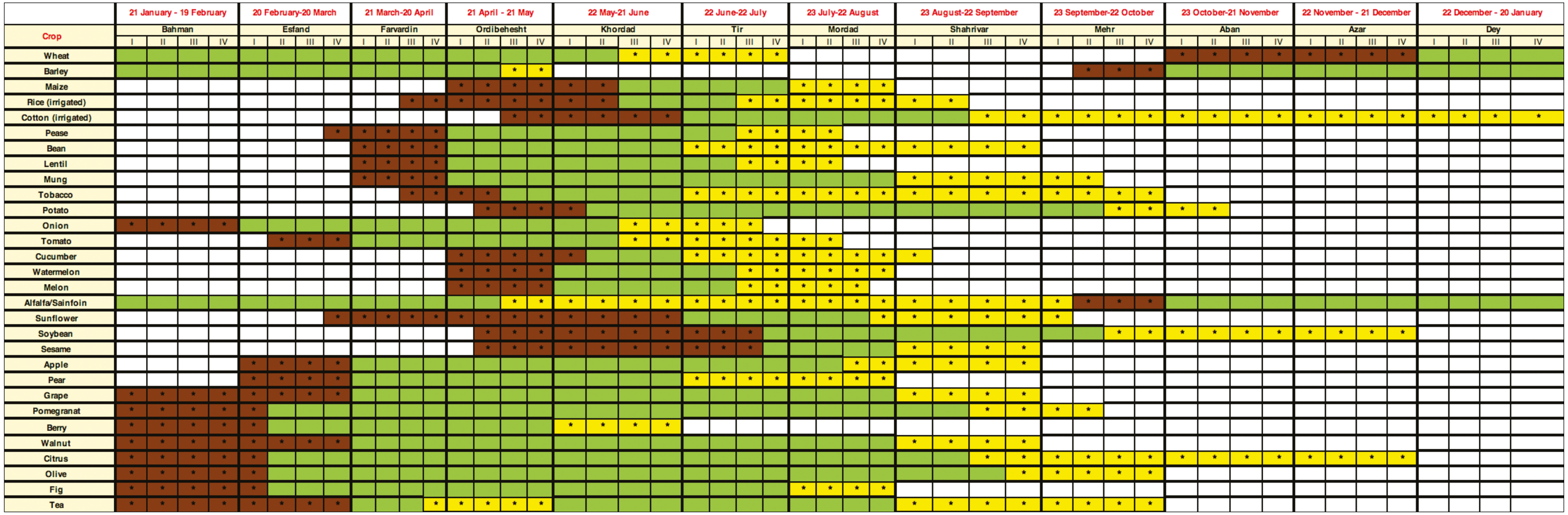


LEGEND

- Sowing
- Growth
- Harvesting

CROP CALENDARS

Table 4 - Mazandaran crop calendar



LEGEND

- Sowing
- Growth
- Harvesting

3.6 Crop type visualization with normalized difference vegetation index series

To better detect and programme agricultural classes, it's possible to make a median NDVI for each season. Then, compositing different seasonal NDVIs allows for a good differentiation between different classes.

For example, one year (12 months) is represented by three, four-month NDVI images which are computed from maximum value composite (MVC) and can be used for detection of many classes[3]. Figures below are composited from three NDVIs in RGB color system:

- NDVI-1-MVC: January to April 2016: red
- NDVI-2-MVC: May to August 2016: green
- NDVI-3-MVC: September to December 2016: blue.

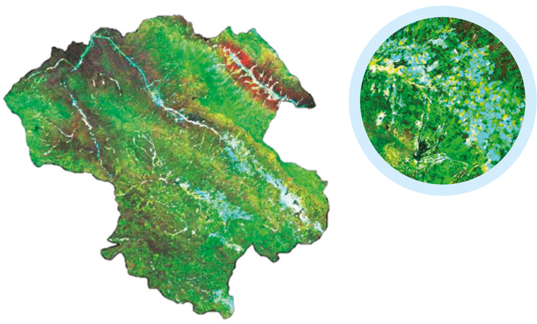
FIGURE 13 - NDVI-MVC in Mazandaran province

MAZANDARAN

Colors seen in Figure 13 represent:

- White: areas with high NDVI values over a year: forest and orchard. Green: areas with high NDVI values only in May to August
- Yellow: areas with high NDVI values only in January to August: crops
- Black: areas with low NDVI values over a year: bare land, urban areas, water
- Cyan: areas with high NDVI values only in May to December: crops
- Violet: areas with high NDVI values only in January to April and September to December: crops.

FIGURE 14 - NDVI-MVC in Zanjan province

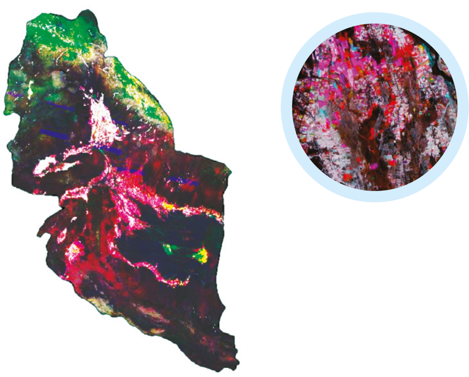


ZANJAN

According to different colors that can be seen in Figure 14:

- Green: shows area with high NDVI values only in May-August: crops
- Yellow: shows areas with high NDVI values only in January-August: crops
- Cyan: shows areas with high NDVI values only in May-December: crops
- White: shows areas with high NDVI values over a year: orchard.

FIGURE 15 - NDVI-MVC in South Kerman region



SOUTH KERMAN

Colors seen in Figure 15 represent:

- Red: area with high NDVI values only in January to April: crops
- Pink: areas with high NDVI values only in January to April and September to December: crops
- White: areas with high NDVI values over a year: orchard
- Cyan: areas with high NDVI values only in May to December: crops
- Grey: area with low NDVI values (0.1-0.3) over a year.

[3] Note: high NDVI in this report refers to values greater than 0.3; low NDVI is less than 0.3.

3.7 Segmentation layer

Before image classification, phenomena and farmland borders must be defined by a segmentation layer which is created by maximum NDVI layer in SAGA-GIS software. Maximum NDVI is computed through the Google Earth Engine. Maximum NDVI for each province is as in the images below.

FIGURE 16 - Maximum NDVI for the three pilot project areas

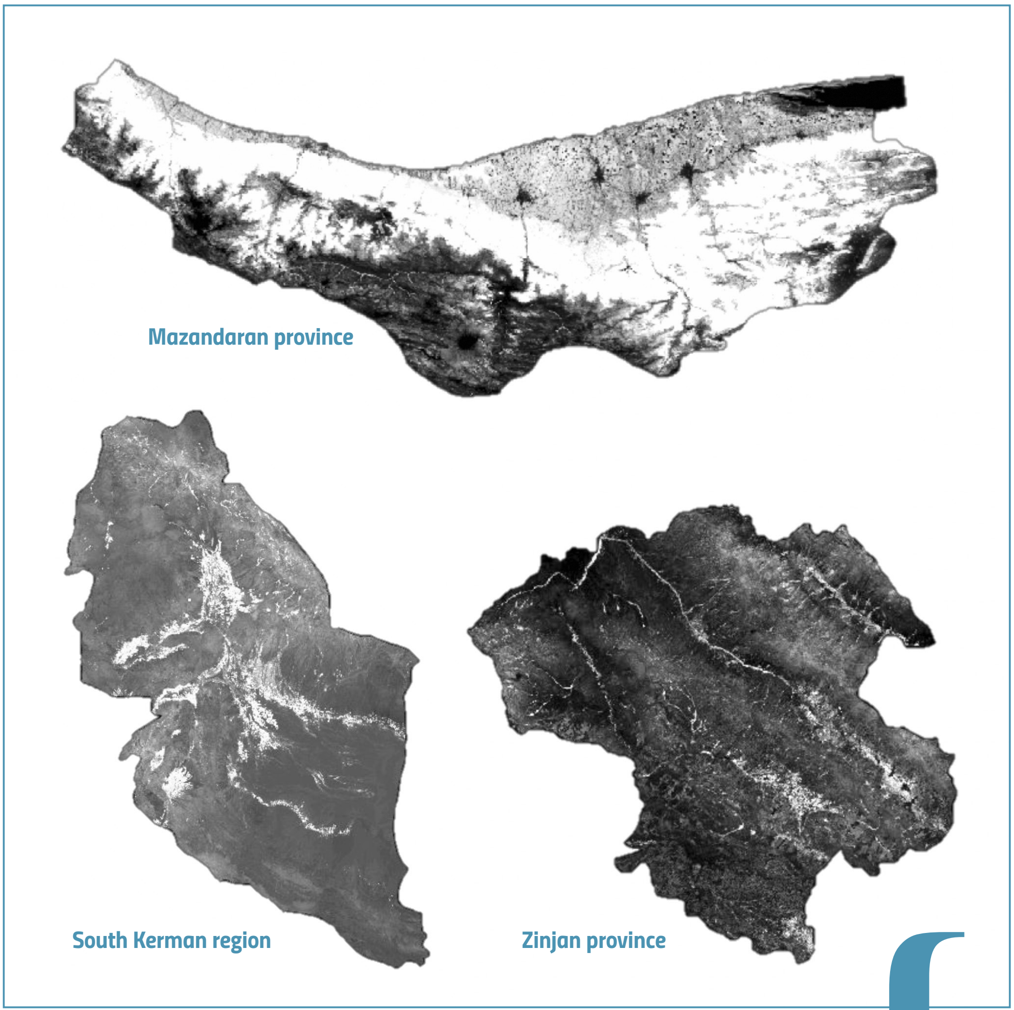
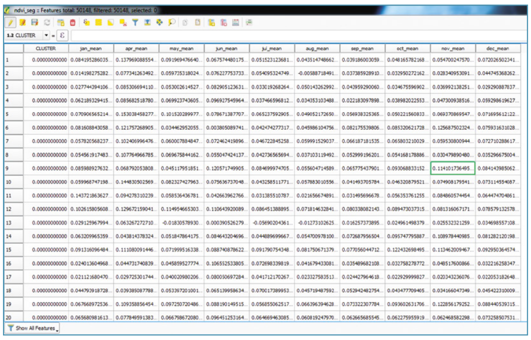


FIGURE 17 - Segmentation layer

3.8 Normalized difference vegetation index inserted

After segmentation, monthly NDVI values were inserted into the segmentation layer (in Attribute table) via the zonal statistics tool in Quantum geographic information system (QGIS) software. Monthly NDVI changes help to define the proper class for each segment through query reporting with SQL server programming language in QGIS. Note: in addition to monthly values, seasonal values for NDVI have led to easier classification. According to this, mean NDVI for each season is computed and inserted as in Figure 18.

FIGURE 18 - NDVI values in QGIS table



ndvi_seg :: Features total: 50148, filtered: 50148, selected: 0

	CLUSTER	jan_mean	apr_mean	may_mean	jun_mean	jul_mean	aug_mean	sep_mean	oct_mean	nov_mean	dec_mean
1	0.0000000000	0.084195286035…	0.137969088554…	0.091969476640…	0.067574480175…	0.051523123681…	0.043514748662…	0.039186003059…	0.048165782168…	0.054700247570…	0.072026502341…
2	0.0000000000	0.014198275282…	0.077341263492…	0.059735318024…	0.076227753733…	0.054095324749…	-0.00588718491…	0.037385928910…	0.032950272162…	0.028340953091…	0.044745368262…
3	0.0000000000	0.027744394106…	0.085306694110…	0.053002614527…	0.082905123631…	0.033019268264…	0.050143262992…	0.043959290060…	0.034675596902…	0.036992138251…	0.029290887837…
4	0.0000000000	0.062189329415…	0.085682518780…	0.069923743605…	0.096927545964…	0.037466596812…	0.034353103488…	0.022183097898…	0.038982022553…	0.047300938516…	0.059298619627…
5	0.0000000000	0.070906565214…	0.153038458277…	0.101520289977…	0.078671387707…	0.065237592905…	0.049052172650…	0.056938325365…	0.050221560833…	0.069370869547…	0.071695612122…
6	0.0000000000	0.081608843058…	0.121757268905…	0.034462952055…	0.003805089741…	0.042474277317…	0.045986104756…	0.082175539806…	0.085320621728…	0.125687502324…	0.075931631028…
7	0.0000000000	0.057820568237…	0.102406996476…	0.060007884847…	0.072462419896…	0.046722845258…	0.059991529037…	0.066187181535…	0.065803210029…	0.059530800944…	0.072710288617…
8	0.0000000000	0.054561917483…	0.107764966785…	0.069675844162…	0.055047424137…	0.042736565694…	0.037103119492…	0.052999196201…	0.054168178886…	0.030479890480…	0.035296675004…
9	0.0000000000	0.085988927632…	0.068792053808…	0.045117951851…	0.120571749905…	0.084699974705…	0.055604714589…	0.065775437901…	0.093068833152…	0.114101736495…	0.084143985062…
10	0.0000000000	0.059968747198…	0.144830502569…	0.082327427963…	0.075636737048…	0.043258511771…	0.057883010556…	0.041493705784…	0.046320879521…	0.074908179541…	0.073114554087…
11	0.0000000000	0.143721863627…	0.094278310239…	0.058536436781…	0.042663962766…	0.031385510787…	0.021656674891…	0.031495696678…	0.056353761255…	0.084860574454…	0.064474704861…
12	0.0000000000	0.102615805608…	0.129672159041…	0.114954665303…	0.110643920089…	0.086451388895…	0.071814632841…	0.080338082143…	0.089473037315…	0.081316067171…	0.078579132578…
13	0.0000000000	0.029125967994…	0.063267272710…	-0.01830578930…	0.000390526279…	-0.05690204361…	-0.01273102625…	0.016257373895…	0.024961498379…	0.025532321259…	0.034698557108…
14	0.0000000000	0.063209965359…	0.043814378324…	0.051847864175…	0.084643204696…	0.044889699667…	0.054700978100…	0.072687956504…	0.095747795887…	0.108978440985…	0.081282120198…
15	0.0000000000	0.091316096484…	0.111080091446…	0.071999516338…	0.088740878622…	0.091790754348…	0.081750671379…	0.077056044712…	0.122432698495…	0.113462009467…	0.092950364574…
16	0.0000000000	0.024013604968…	0.044731740839…	0.045859527774…	0.106552533805…	0.072698339819…	0.041679433081…	0.035489682108…	0.032758278772…	0.048517600866…	0.032216258347…
17	0.0000000000	0.021121680470…	0.029725301744…	0.040020980206…	0.080050697284…	0.041712170267…	0.023327583513…	0.024427964618…	0.022929999827…	0.020343236076…	0.022053182648…
18	0.0000000000	0.044793918728…	0.039385087788…	0.053397201001…	0.065139958634…	0.070017389953…	0.045719487592…	0.052942482754…	0.043477709405…	0.034166047349…	0.045422310009…
19	0.0000000000	0.067668972536…	0.109358856454…	0.097250720486…	0.088190149515…	0.056855062517…	0.066396394628…	0.073322307784…	0.093602631706…	0.122856179252…	0.088440539315…
20	0.0000000000	0.065680981613…	0.077849591383…	0.066798672080…	0.096451253164…	0.064669463085…	0.060819247970…	0.062665685545…	0.062275955919…	0.062468582298…	0.073258507531…

Show All Features

3.9 SQL codes

In order to classify segmentation layers, NDVI changes over a year – measured either monthly or seasonally – must be used. This project used seasonally method. Some SQL codes which used in this project for identifying different classes are presented in Figure 19.

FIGURE 19 - SQL codes

```
Orchard: (winter >= 0.3) and (spring >= 0.3) and (summer >= 0.3) and (fall >=
0.3)

Crop: ((winter >= 0.3) and (spring =< 0.3) and (summer =< 0.3) and (fall =<
0.3)) or ((winter >= 0.3) and (spring >= 0.3) and (summer =< 0.3) and (fall
=< 0.3)) or …

Water: (winter < 0) and (spring < 0) and (summer < 0) and (fall < 0)

Rangeland: (winter > 0.1 and winter < 0.3) and (spring > 0 and spring < 0.3)
and (summer > 0 and summer < 0.3) and (fall > 0 and fall < 0.3)

Bare land: (winter >= 0 and winter =< 0.1) and (spring >= 0 and spring =<
0.1) and (summer >= 0 and summer =< 0.1) and (fall >= 0 and fall =< 0.1)
```

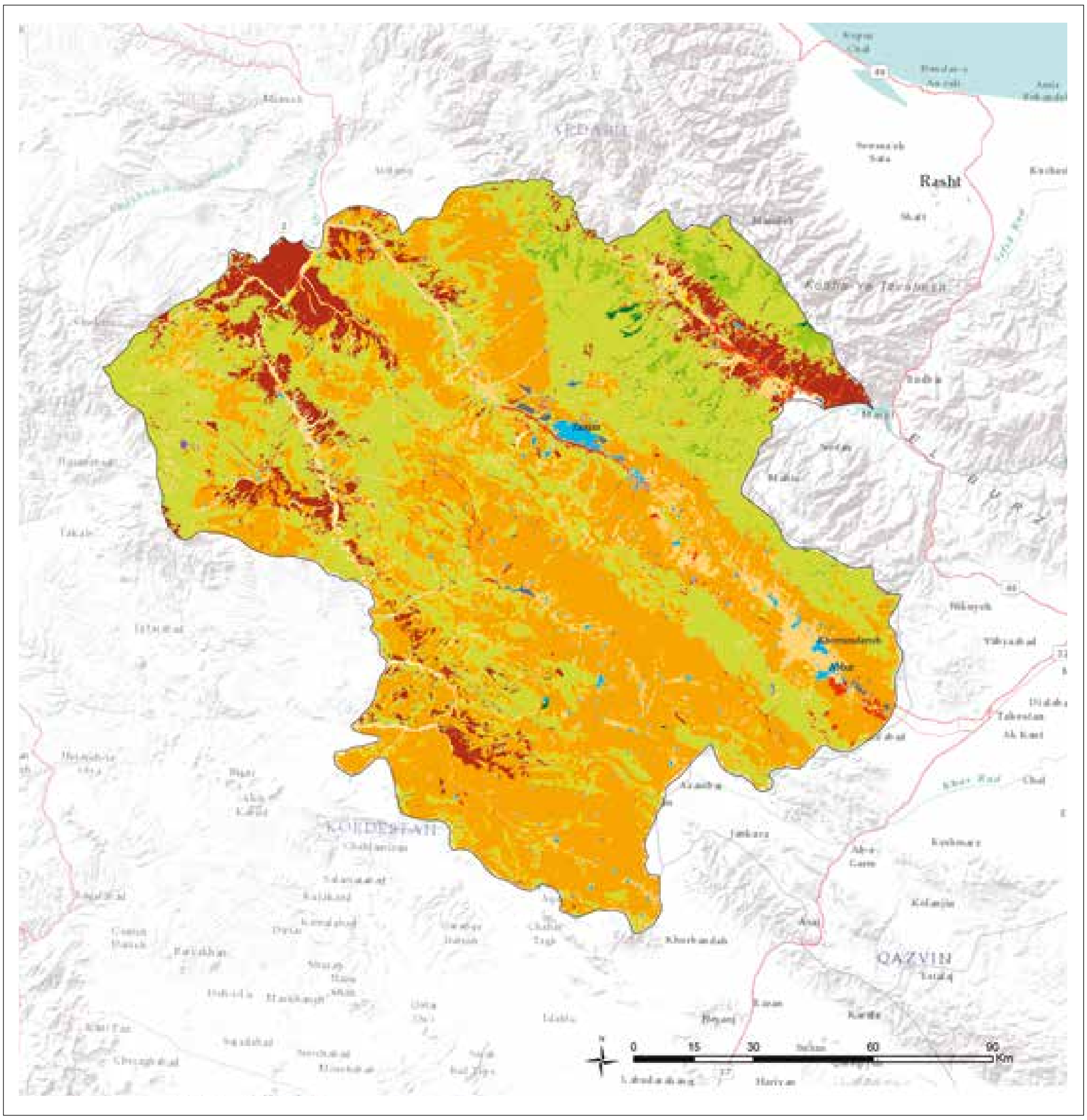
3.10 Classified maps

The FAO classification system defines 16 classes for land use maps in Iran according to the LCC standard. In fact, all phenomena in Iran's environment must be translated to these classes.

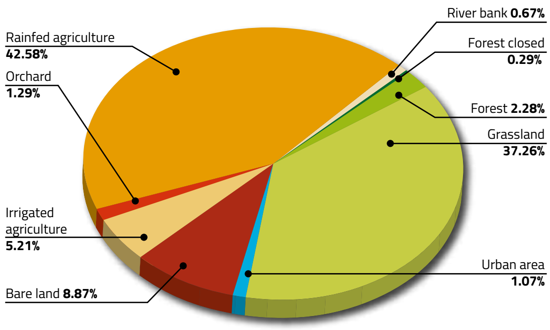
Table 5 - Land cover classes

LCCS code		Name	Description
	BA	Bare land	Rock or soil sometimes with very sparse natural vegetation (0-15%)
	FC	Forest closed	Woodland with closed (60-100%) trees and/or shrubs
	FO	Forest open	Woodland with open (20-60%) trees and/or shrubs and herbaceous natural vegetation
		Grassland	Relatively dense grassland natural vegetation, with very sparse shrubs and/or trees (0-15%)
	UA	Urban area	Urban and/or rural settlement
	UAM	Urban area mixed	Urban build-up areas + small cultivated herbaceous crops + orchards and other plantations
	HCI	Irrigated agriculture	Irrigated herbaceous crops
	TCP	Orchard and other plantations	Orchard crops or other plantation
	HCR	Rainfed agriculture	Rainfed herbaceous crops
	IA	Industrial area	Non-urban build-up areas and other constructions (e.g. industrial)
	RD	Roads	Roads
	MN	Mines	Extractions sites
	RB	River bank	River bank (bare soil)
	RC	Rice crop	Irrigated rice crop
	WB	Water body	Lake or river - perennial fresh water

FIGURE 20 - Land cover map of Zinjan province



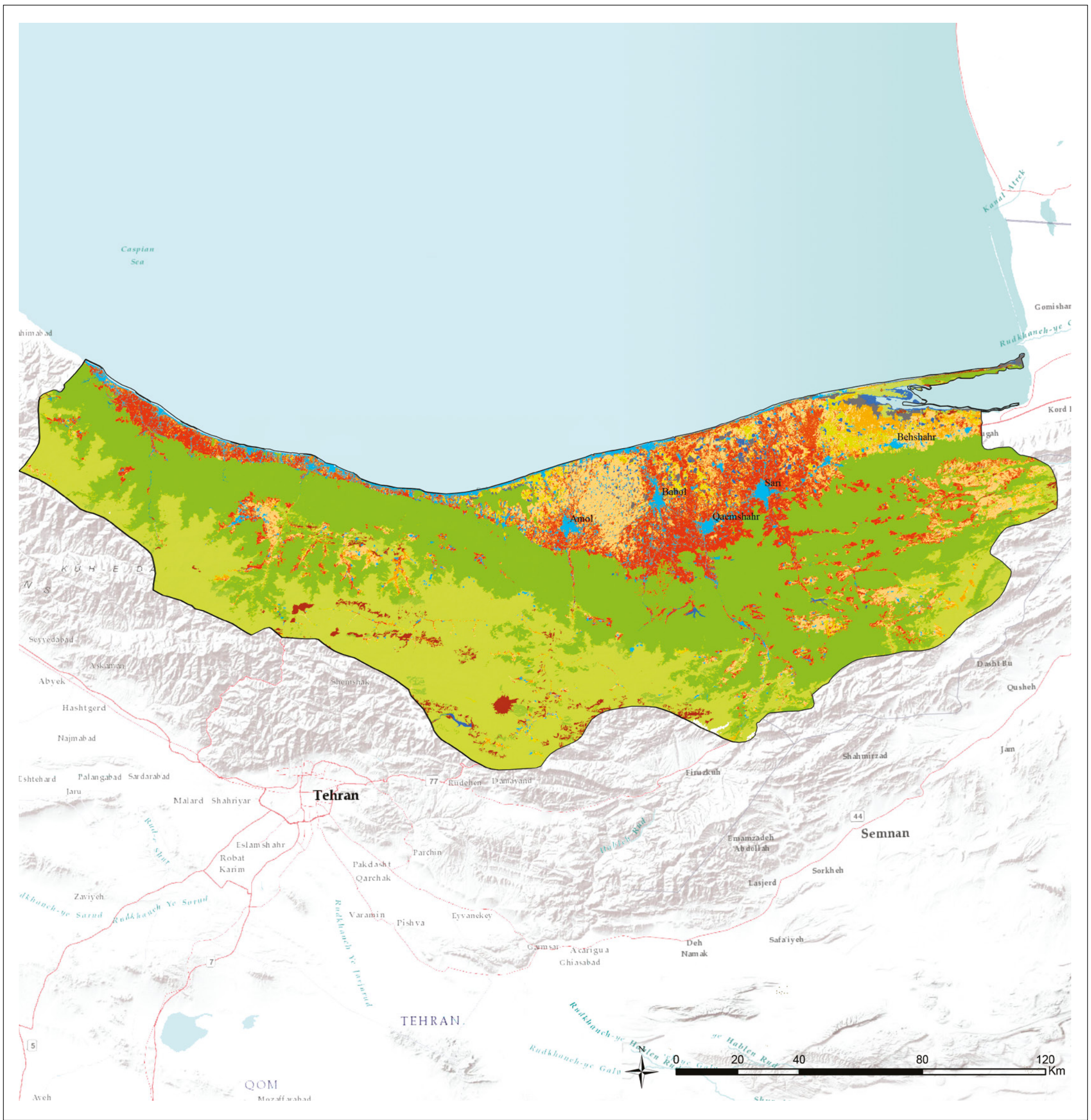
Database and map source: Geospatial Unit, FAO



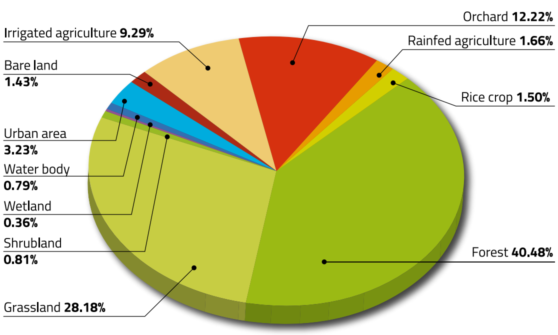
Legend		km²	%
	Bare land	1 921.00	8.87
	Forest closed	65.10	0.29
	Forest	492.98	2.28
	Grassland	8 068.71	37.26
	Urban area	232.29	1.07
	Urban area mixed	8.94	0.04
	Irrigated agriculture	1 128.38	5.21
	Orchard	278.84	1.29
	Rainfed agriculture	9 219.00	42.58
	Industrial area	32.92	0.15
	Roads	20.81	0.10
	Mines	6.34	0.03
	River bank	146.15	0.67
	Rice crop	19.91	0.09
	Water body	15.33	0.07
	Total	**21 653.32**	**100**

The boundaries and names shown and the designations used on the map above do not imply the expression of any opinion whatsoever on the part of FAO concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers and boundaries. Dashed lines on maps represent approximate border lines for which there may not yet be full agreement.

FIGURE 21 - Land cover map of Mazandaran province



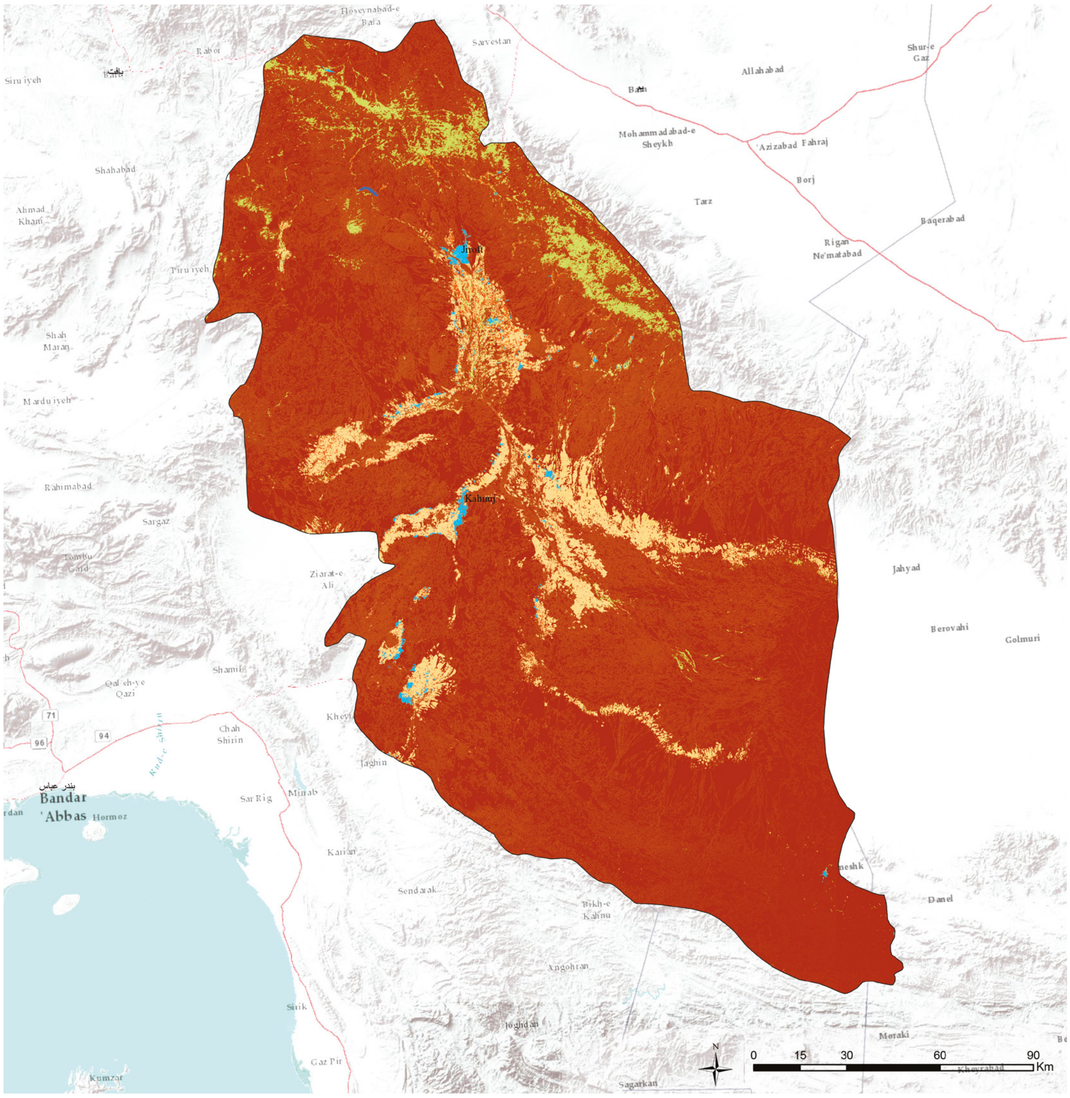
Database and map source: Geospatial Unit, FAO



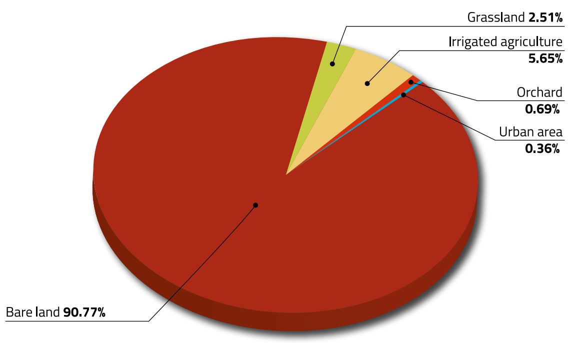
Legend	km²	%
Bare land	341.24	1.43
Snow and ice	10.02	0.04
Forest	9 646.08	40.48
Grassland	6 715.77	28.18
Shrubland	192.44	0.81
Urban area	770.53	3.23
Irrigated agriculture	2 213.93	9.29
Orchard	2 913.38	12.22
Rainfed agriculture	395.47	1.66
Wetland vegetation	85.67	0.36
Rice crop	355.61	1.50
Water body	188.54	0.79
Total	**23 828.68**	**100**

The boundaries and names shown and the designations used on the map above do not imply the expression of any opinion whatsoever on the part of FAO concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers and boundaries. Dashed lines on maps represent approximate border lines for which there may not yet be full agreement.

FIGURE 22 - Land cover map of South Kerman region



Database and map source: Geospatial Unit, FAO



Legend		km²	%
	Bare land	34 231.95	90.77
	Grassland	946.51	2.51
	Urban area	137.63	0.36
	Irrigated agriculture	2 131.34	5.65
	Orchard	259.60	0.69
	Water body	5.72	0.01
	Total	**37 712.74**	**100**

The boundaries and names shown and the designations used on the map above do not imply the expression of any opinion whatsoever on the part of FAO concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers and boundaries. Dashed lines on maps represent approximate border lines for which there may not yet be full agreement.

4. AREA FRAME METHODOLOGY

4.1. Overview

The purpose of any statistical survey is to acquire knowledge of a target population, the aggregate of elementary units about which information is needed. Information can be collected through a complete enumeration census or a sample survey.

A complete agricultural, enumeration census is a survey examining all the elementary units forming a population. Historically, the primary objective of such a census was to provide a detailed classification of the agricultural structure of the inspected area and of agricultural holdings and data for small administrative units.

Due to the high cost of a complete enumeration census, some countries instead try to obtain a similar kind of information by conducting sample censuses, instead of complete enumeration censuses.

In an agricultural sample survey, the values of survey variables for the target area are inferred from the values of the variables in a sample of elementary units. Sample surveys can be further classified as probability and non-probability (or purposive). In the probability sample survey, each elementary unit of the population has a given and a non-null probability to be included in the sample. In contrast, the non-probability sample survey chooses the sampling units with no randomness criterion.

The conventional area frame sample survey is a type of agricultural probability sample survey activity introduced as a vehicle for conducting surveys on crop acreage, cost of production, farm expenditures, yield and production, livestock inventories and other agricultural items. It is a costly and time-consuming method that requires a basic data set (aerial photographs/images) and topographic and thematic (land cover/land use) maps or databases. Meticulous and laborious work is necessary to delineate strata, primary sampling units (PSUs) and then segments; and qualified staff with a strong background in statistics must be employed.

Area frame sampling consists in "dividing the total area to be surveyed into *N* small blocks (segments) without any overlap or omissions (and) furthermore, select a random sample of *N* small blocks and get the desired data for reporting units of the population that is in the sample blocks" (Madana, 2002). When stratifying the survey area into a number of strata (which are then divided into PSUs), it is important to ensure homogeneity within a stratum: this will make it possible to reduce the sampling variance and obtain more accurate estimates. Stratification and the correct definition of strata are, therefore, crucial steps. An accurate land cover database can facilitate and improve the process and represents the baseline for the subsequent stratification. The specifications for building an area frame consist of strata definitions and target sizes for both PSUs and segments within each stratum.

To prepare an area frame, the first requirement is up-to-date cartographic materials. Their resolution must be sufficient to enable land stratification and the definition of PSUs. PSUs that are usually constructed from photographs or satellite images must be demarcated by recognizable permanent physical boundaries (roads, rivers, etc) and associated with a particular land use strata.

For each land use stratum, a number of PSUs are selected; these are, in turn, divided into secondary sampling units. In the last stage, the final sampling units of an area frame are defined. These land areas are called segments. They should not overlap and should cover the entire survey area. To gather agricultural information, segments should be visited by an interviewer. These enumerators collect field data, completing a questionnaire via personal interviews with people working in agriculture or other respondents who can provide information on the tract included in each sample segment selected. In addition to these questionnaires, data collection involves the identification and measurement of cultivated areas.

For each sample segment, the enumerator uses an enlarged aerial photo (or a map) that shows the boundaries of the segment. This is called the segment photo. For each tract within a given sample segment, an enumerator delineates the boundaries of the tract and of all the fields included therein. The enumerator verifies the crops planted and other land

uses for each field and also obtained this information from the agriculture holders. During the interview, the enumerator may also superimpose a transparent grid over the segment photo for an approximate verification of the fields' reported area.

The agricultural areas identified in each sample segment can be measured. Area frames are critical to the production of quality estimates, because they provide complete coverage and land areas are represented in a probability survey with a known chance of selection (Cotter and Tomczak, 1994). This frame does not suffer rapid obsolescence unless the population extends into areas that are not covered by the frame (FAO, 1996).

4.2 Area frame construction and stratification

The area sampling frame was not developed identifying the PSU and then splitting the selected PSUs into segments. In fact, an international statistics consultant reported that currently, with the support of geospatial information, ground data collection does not require segments with permanent physical boundaries. The consultant suggested developing an area sampling frame with units made of segments generated by a regular grid.

The land use/land cover maps, which had been prepared for the three pilot provinces, were reclassified, focusing on agricultural classes according to Table 6. All non-agricultural classes were designated as "no agriculture" and the agricultural layer was generated.

Then, "no agriculture" classes were removed from the land cover layer and a map of the agriculture mask was produced. Figure 23 shows the agriculture mask for all three provinces (Zanjan, Kerman and Mazandran); where the use of white represents areas with agriculture activities and black areas with no agriculture activities. To account for the different conditions affecting crop growth, an existing layer of climatic agro-ecological zones was combined with the mask of agriculture map layer to generate the stratification of the three provinces via the intersect function in GIS software. A section of this intersection is shown in Figures 27. An intersection was created between the agricultural land cover layer and the climatic agro-ecological zones layer. A section of this intersection is shown in Figure 28. Regular

Table 6 - Re-classification of land cover layer for generating agricultural land cover layer

LCCS code	LC description	LC re-classified
BA	Bare land	No agriculture
FC	Forest closed	No agriculture
FO	Forest open	No agriculture
GRWS	Grassland	No agriculture
HCI	Irrigated agriculture	HCI
HCR	Rainfed agriculture	HCR
IA	Industrial area	No agriculture
MN	Mines	No agriculture
RB	River bank	No agriculture
RC	Rice crop	RC
RD	Roads	RD
TCP	Orchard	TCP
UA	Urban area	No agriculture
UAM	Urban area mixed	UAM
WB	Water body	No agriculture

grids (fishnets) at increasing size (20 ha, 30 ha and 40 ha) were generated for provide an estimate of the average number of parcels per segment to be surveyed on the ground, since no previous area frame was implemented in the selected provinces.

- a fishnet of 20 ha (200,000 m²) has a side 447 m
- a fishnet of 30 ha (300,000 m²) has a side 548 m
- a fishnet of 40 ha (400,000 m²) has a side 632 m.

Since several segments presented parts of their area covered by more than one stratum, when computing percentage and area for all strata within segments the percentage of the different strata in each segment was calculated using the tabulate intersection tool in ArcGIS software.

FIGURE 23 - Agriculture masks in the 3 project pilot areas

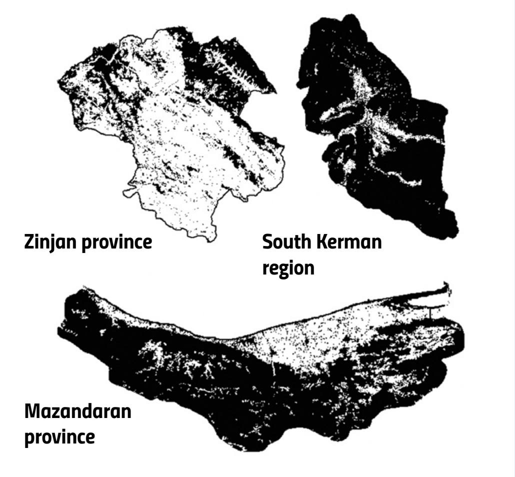


FIGURE 24 - South Kerman climate map

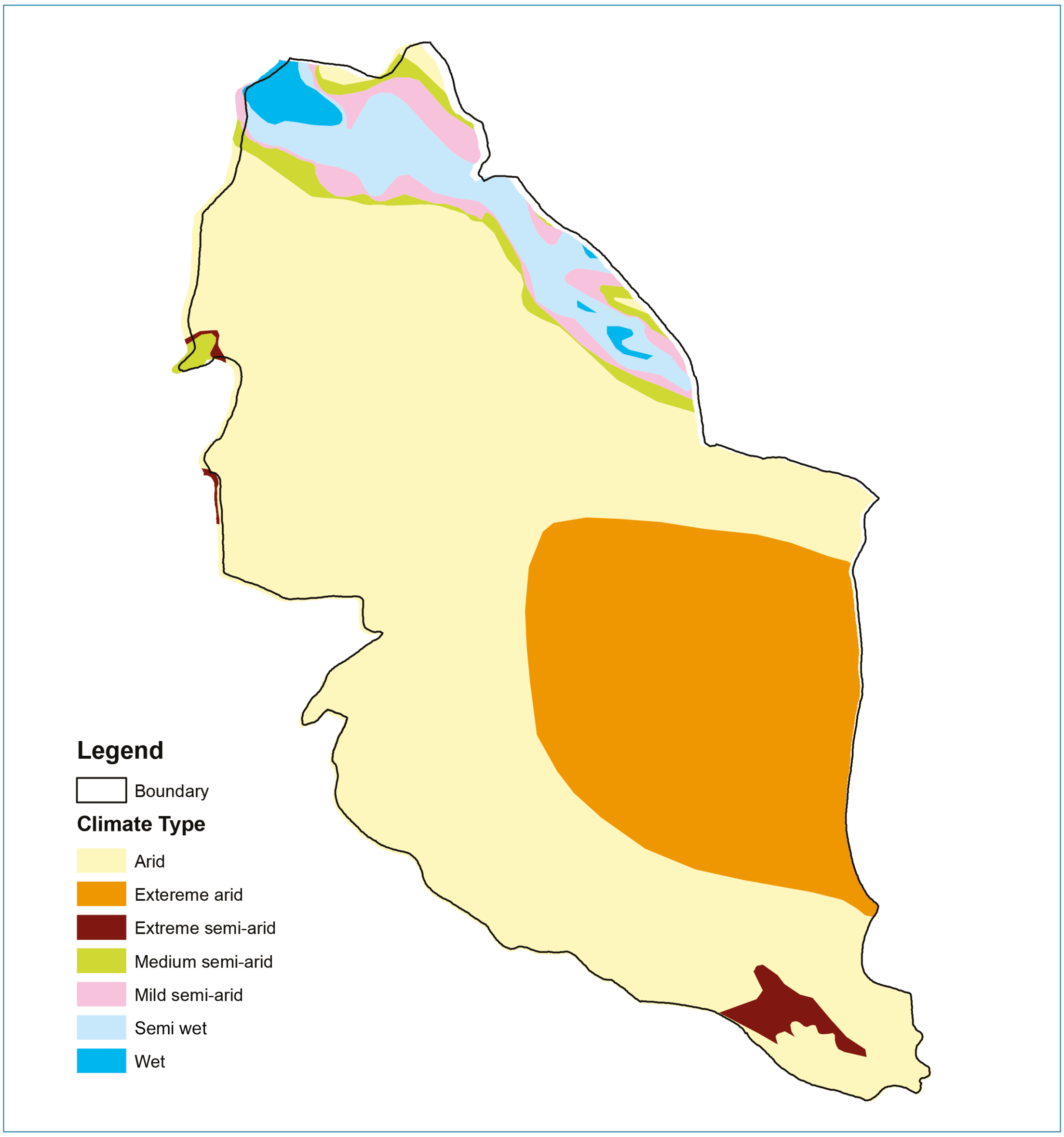


FIGURE 25 - Zanjan climate map

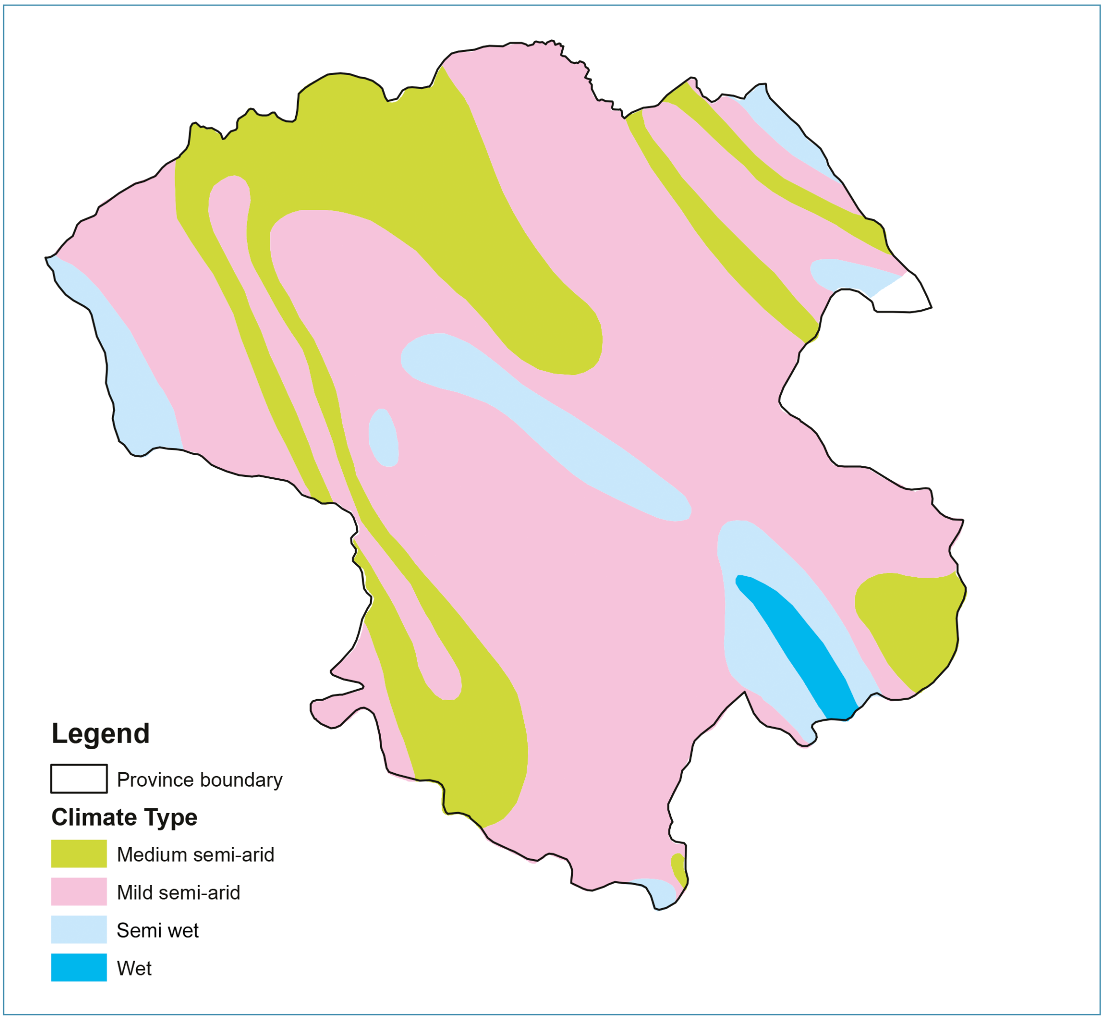
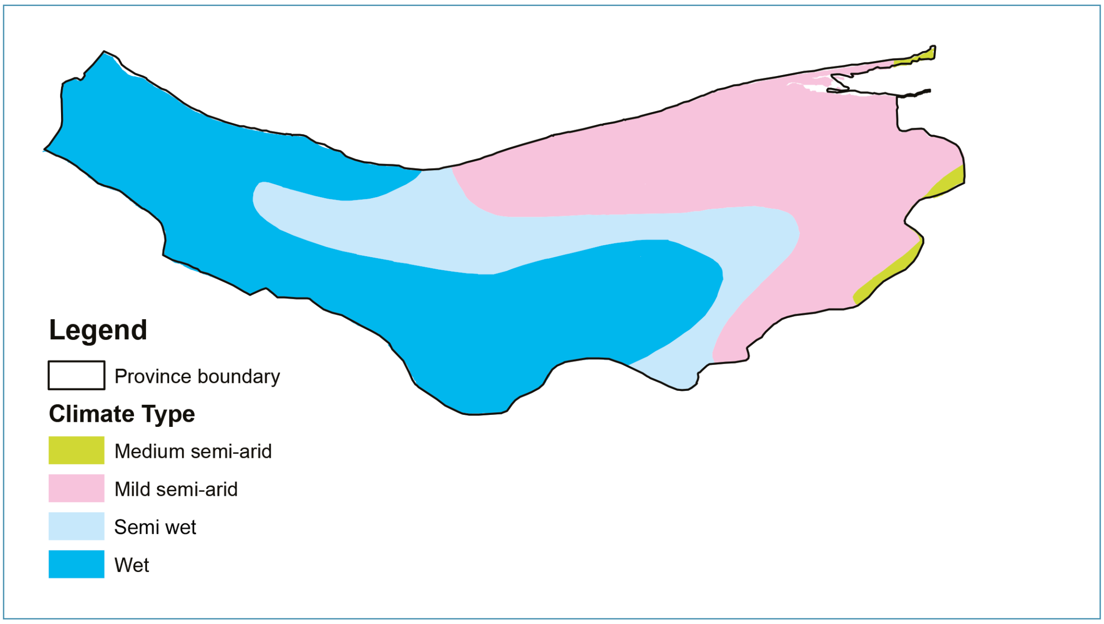


FIGURE 26 - Mazandaran climate map



The output table contained records for each stratum in one segment, where *n* is the number of strata within one segment. For example, if a segment contains four strata, the output table will have four records for that segment ID; accordingly, we have similar segment IDs in records.[4]

To fix this redundancy in segment IDs in records and also discover the exact number of remained segments, the pivot table tool in ArcGIS software was used to transform the output table into a table that contains one record for each segment ID with strata attributes as separate attribute fields.[5] This table shows the exact number of agriculture segments without redundancy, and those that contain strata with different percentages of agricultural types. For each segment, the percentage of different strata was computed and then the segment was attributed to the stratum with the highest percentage, using ArcGIS software. In an Excel file, this is specified with a field named MAX percentage_Strata. Segments were then grouped, based on the percentage of agricultural area in the segment. Consequently, the final stratification was based on land cover, agro-ecological zones and percentage of agriculture.

The intersection was imported into Excel software, unnecessary attributes removed, and the area for the intersected polygons calculated. The grid with 20 ha cells (segments) was chosen for collecting ground data in the current test for Kerman and Mazandaran provinces and 30 ha for Zanjan province. This size proved to be the most suitable to balance the average size of the parcels with the estimated effort necessary for data collection (to be tested through the ground survey). To avoid biased estimates, all the segments with even a small percentage (at least five percent) of agricultural area were included in the population. Several segments presented parts of their area covered by more than one stratum. The percentage of the different strata in each segment was computed and the segment was attributed to the stratum with the highest percentage. The segments were grouped based on the percentage of area of the different strata in the segment.

Consequently, the final stratification was based on land cover, agro-ecological zones and percentage of agriculture. Since no previous information on the variability inside each stratum was available, the

FIGURE 27 - Intersection of agricultural land cover and climatic agro-ecological zone

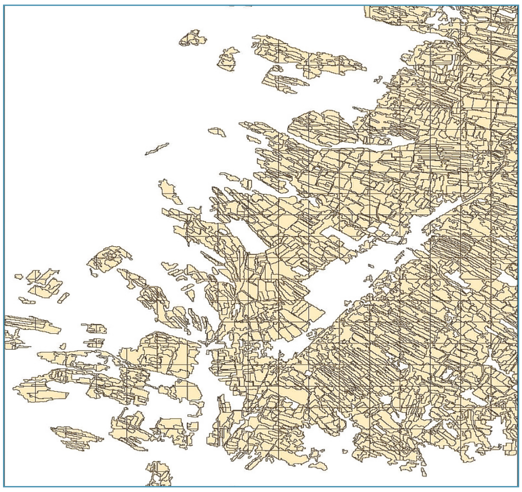
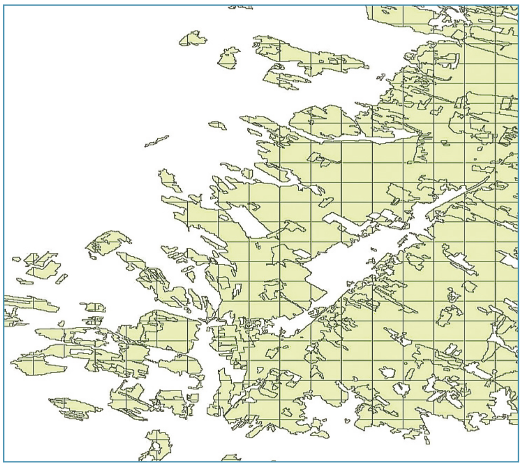


FIGURE 28 - Dissolved of intersection layers agricultural land cover and climatic zone

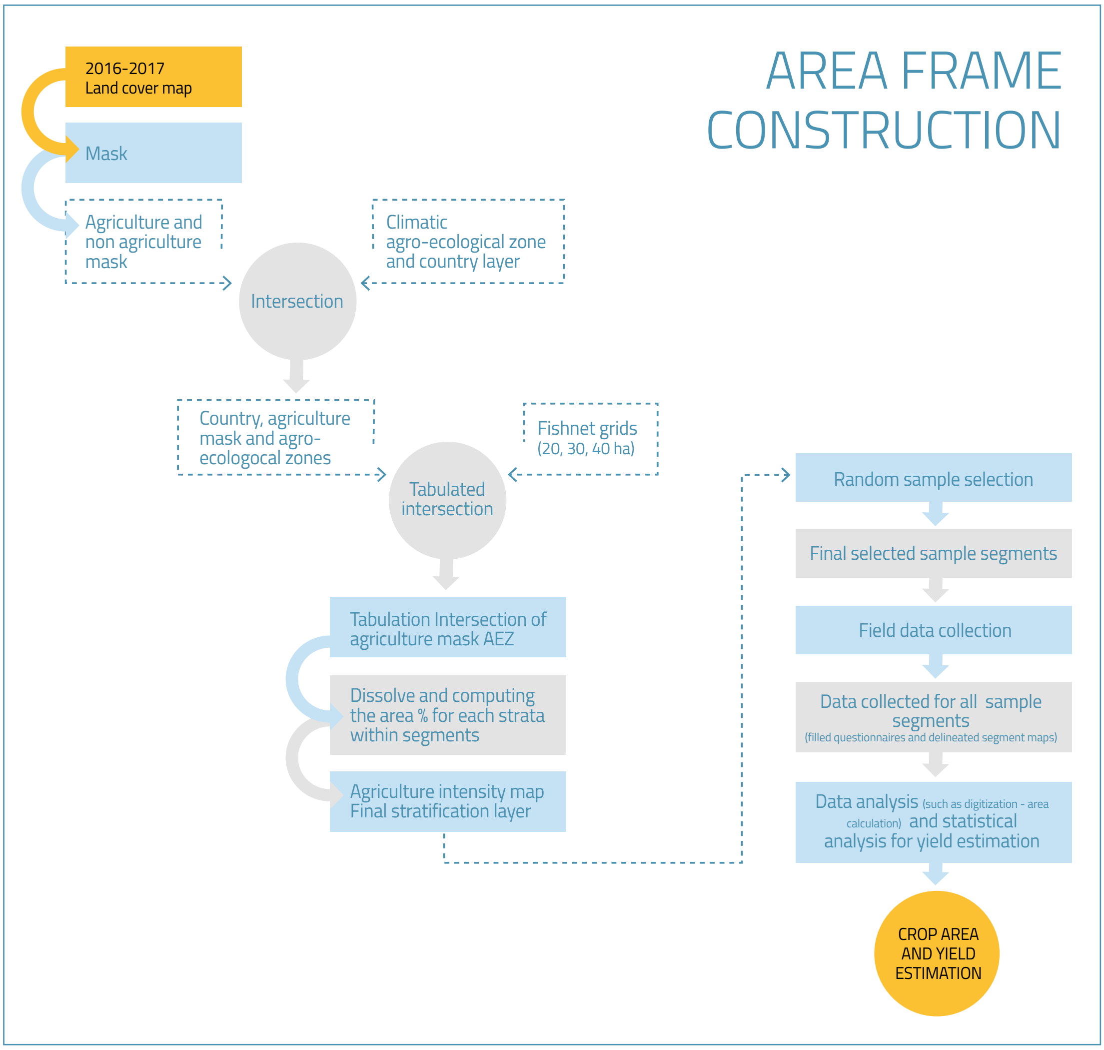


proportional allocation was adopted; that is, the total number of segments to be selected was subdivided among the final strata according to the stratum size, provided that at least two segments were selected in each stratum, for computing the variance of the acreage of crops in each stratum. The sample selection was made using the permanent random numbers. In each final stratum, a random number was assigned to each segment in the population. The segments were then ordered according to their random numbers. This procedure was followed in each final stratum independently and corresponds to the sample selection. The first subset of segments

4 The output table was saved in geodatabase format and the Excel file was saved with title "tabulate intersection table".

5 The output table and Excel file was pivot tabulate intersection.

FIGURE 29 - Area frame construction flow-chart



corresponding to the number of segments allocated to each stratum became part of the group of segments to be surveyed on the ground. With this procedure, the ground survey could be performed on additional segments in a stratum, if needed, following the order of the segments generated by the ordered random numbers and treating all surveyed segments as one-step simple random sample. For each segment, a form was pre-filled with UTM coordinates, final stratum ID and name, province, county and screenshot from Google Earth to support parcel identification. Pilot samples of 90, 72 and 68 segments were selected for ground truth collection respectively in Zanjan, Mazandaran and South of Kerman. Training workshops on how to collect segment information was conducted in three pilot areas for the enumerators from MAJ. Then, the field data collection was done by them (the process of field work and data collection is presented in the next session). . Segments ground information was then imported to Excel and ARC/GIS environment for statistical analysis and satellite image processing. Once the first ground survey is conducted, collected data can be used for computing the variability inside each final stratum that allows estimating the sample size needed for obtaining estimates with prefixed accuracy and performing the optimal allocation for the next surveys. Due to the adoption of the permanent random numbers, there is no need to select ex novo all the segments in each stratum. The additional segments can be selected following the order of the permanent random numbers. Thus, the segments selected for the first survey and the additional ones can be treated as just one simple random sample inside each stratum. Consequently, the parcel identification already performed on the screenshots from Google Earth for the first survey can be used for the next surveys.

4.3 Field work and data collection

One of the main objectives of the project was to provide practical training and capacity building towards development of efficient methodologies for area frame sampling. Therefore, before starting the main activity for data collection, several workshops were planned to meet these aims for each of three pilot provinces. These workshops included two days of training. Day one looked at theory of image processing, concepts and techniques of area frame sampling, GPS application including GPS for positioning and segment finding. In addition, the questionnaire forms and the segment photos were introduced and the enumerators were trained how to complete them.. Day two focused on field-based techniques for area frame sampling in a test segment. The segment included a wide variety of land cover/use types, particularly irrigated, orchards, bare lands and fallow-lands.

For more efficient, practical training and report generation, enumerators were divided in several groups, each assigned a national consultant. Field practice included how to find segments by GPS, how to define and draw boundaries of parcels on the Google image plot and how to complete the questionnaire After training enumerators in collecting the main data from the selected segments of each pilot province, printed sample segments were distributed among the enumerators for a data collection exercise. The coordinates of the four corners of each segment – which had been marked on the plot of Google Earth images of the segment, and also the number of corresponding segment – were imported to GPS before going to each segment on the ground, Then, the location of, and access to, each segment was planned using Google Earth images.

After arriving at a segment on the ground, information including the enumerator's identification data, date of enumeration, travel time and time of arrival were entered in the questionnaire. Then, the position of a specific corner of a segment was identified on the ground. In addition, with respect to the position of the enumerator and the image, the other corners were identified as well. During the enumeration procedure, the boundaries of all parcels within the segment were delineated on the plotted image. A unique number was assigned to each parcel, and the numbers was also marked on the plot. Also during enumeration, basic information was entered in the questionnaire, such as crop name, cultivation type (irrigated versus rainfed), irrigation systems, and other information. Only for cereals close to harvest was information collected such as average number and size of ears, crop status (parasites, color, vigour) and yield predictions were entered in the questionnaire.

After completing the survey with required information, all segments were digitalized with relevant information from questionnaires to calculate the area of parcels and perform other required analyses.

FIGURE 30 - Segments photo map

PART 2

Current data collection practices by Ministry of Agriculture Jihad

5. SAMPLING AND SURVEY METHODOLOGY ADOPTED

Four types of agricultural surveys are annually completed, including:

- wheat and barley (surveyed in July/August)
- rice (surveyed in November/December)
- other crops (surveyed in November/December)
- survey of production costs (November/ December).

The survey of production costs is part of the crop-type survey. At present, due to budget limitations and sampling difficulties, surveys of wheat and barley as well as other crops, have been integrated and completed under the Agricultural crops survey project.

Sampling and survey methodology is mainly based on the list frames and the last agricultural census data. Stratified, two-stage sampling composed of selecting the sample villages in the first stage and selecting the sample holders in the second stage is used. Agricultural census data with temporal resolution of 10 years are provided by the Statistical Center of Iran (SCI). An agricultural sampling and survey process as described in the technical plan provided by the Statistical Office of the MAJ (Appendix 1) includes the following seven stages:

Stage 1: Definition of the survey targets and variables of interest. e.g. wheat and barley, rice, other crops.

Stage 2: Definition of the sampling spaces and target population. Townships are used as survey and sampling spaces; therefore, existing maps of township borders are used to define the relevant spaces and boundaries. Using agricultural census data, villages in each township are divided in two classes of A: "with agricultural activities"; and B: "without agricultural activities". Villages in class A are then selected and used as the target population for sampling in each township. Also, a number of class B villages may be randomly selected and included in the enumeration.

Stage 3: Stratification. The Dalenius stratification approach (Dalenius and Hodges, 1959) as described in Cochran (1977) and the technical plan of the Statistical Office of MAJ, are used to stratify the selected villages in each township. Total specified agricultural areas of villages are extracted from the agricultural census data, and used as the stratification variable. The objective is to stratify villages such that the standard deviation within each strata is minimized. The number of strata is different for townships and usually varies between two and nine. With this method, village data are sorted into ascending order, based on their area under cultivation. Then, classification is performed using the following statistics:

Table 7 - Statistics used for the classification

Class limits	Frequencies	Sq. root of frequencies	Cumulative sum of sq. roots
$a+u$	f_1	$\sqrt{f_1}$	$F_1 = \sqrt{f_1}$
$a+2u$	f_2	$\sqrt{f_2}$	$F_2 = \sqrt{f_1} + \sqrt{f_2}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$a+iu$	f_i	$\sqrt{f_i}$	$F_i = \sqrt{f_1} + \sqrt{f_2} + ... + \sqrt{f_i}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$a+Ku$	f_K	$\sqrt{f_K}$	$F = F_K = \sum_{i=1}^{K} \sqrt{f_i}$

where:

α is the crop area of the village with the least area under cultivation

u is the class width

K is the initial number of classes
f_i is the frequency of ith class (no. of villages that their area under cultivation is between $\alpha+(i-1)u$ and $\alpha + iu$).

After the initial classification, a secondary classification (main) is performed. First, D is calculated as $D= F/L$

where:

L is the maximum number of classes in a township and F_i are used to calculate the limits of classes as follows:

- D and F_1 are compared, if $D < F_1$, L is reduced by one and new value for D is calculated. This continues until D becomes greater than F_1.
- by comparing D with F_i s (last column of the Table), two absolute frequencies of F_i and $F_i +1$ that D lies between them are identified. Absolute differences of D with F_i and $F_i +1$ are calculated. Either of the F_i or $F_i +1$ showing the least absolute difference is selected and used as the lower limit of class 1. Using the $2D$, $3D$,...and $L–1D$ and their comparison with F_i, s, the same process is repeated to define limits of other classes.

Stage 4: Definition of the sample size. Sample size expressed as the number of villages (n) for each township is independently defined by considering the total number of villages of the township, expected precision and heterogeneity of the township as follows:

$$n = \frac{(\sum_{h=1}^{L} N_h\, S_h)^2}{V + \sum_{h=1}^{L} N_h\, S_h^2};$$

where:

L is the number of strata in the township
N_h is the number of villages in strata h
S_h is the standard deviation of the variable of interest (agricultural areas) between villages of strata h
V is the maximum expected variation (errors) for estimation of the agricultural areas and computed as: $v = (\alpha \,.\, \gamma)^2$.

where:

α is the coefficient of variation (usually 5%)
γ is the total agricultural areas in each township.

Stage 5: Allocation of sample size to strata. Sample size for each strata is defined using the optimum Neyman allocation approach, which tries to maximize the survey precision for a given fixed sample size. As well, where more samples are allocated for strata with higher number of villages and higher standard deviation between the villages, as follows:

$$n_h = \frac{N_h . S_h}{\sum_{h=1}^{L} N_h . S_h} \times n$$

where:

n_h is the sample size (number of sample villages) for stratum h
n is the sample size for each township as defined in stage 4,
N_h is the number of villages in stratum h
S_h is the standard deviation of specified agricultural areas of villages in stratum h.

Stage 6: Selection of the sample villages. By defining the sample sizes for different townships and strata, villages of each strata are ranked by the size of their specified agricultural areas. A systematic sampling approach with sampling intervals of l_h ($l_h = N_h/n_h$) is then used to select the sample villages in each strata. This ensures that villages with varying rates of the specified agricultural areas are selected.

A special sample is selected containing major crops or groups of crops (one sample for each of wheat and barley, rice and other crops).

Sub-samples of the selected sample villages for area and yield studies are randomly chosen and used for both the area plus the yield and production cost studies. Examples of regular sample villages and holders for different purposes (area and yield of varying crop types and production costs) are outlined in Tables 8 and 9 (Statistical Office, MAJ, personnel communication).

Table 8 - Sample villages and holders for area and yield studies 2014-2015

Crop surveys	Sample villages	Sample holders
Wheat, barley and other crops	6 944	86 519
Rice	893	10 512
Total	**7 837**	**97 031**

Table 9 - Sample villages and holders for production cost studies 2014–2015

Crop surveys	Sample villages	Sample holders
Wheat, barley and other crops	6 944	14 542
Rice	893	1 835
Total	**7 837**	**16 377**

Stage 7: Selection of the sample holders**.** For the selection of sample holders, a list of farm holders in each of the sampled villages was prepared and ordered by the area of their specified agricultural crop holdings. Holders were then divided into either 'large' or 'other' holders. Large holders were defined as the top ten percent in the ordered list of holdings. All the large holders (at least two) and ten percent of others (at least three) were then selected for enumeration.

For rice surveys, sampling rates of six percent were used for each of the 'other' holders. A sub-sample of selected holders with minimum and maximum numbers of two and four were selected and used for surveys of production costs. It is notable that surveys of the agricultural holdings of the sample holders were limited to the selected sample villages. Therefore, holdings located outside of the enumerated sample villages were not considered. It is assumed that all the sampled holdings belong to the relevant sample village.

6. METHODOLOGY ADOPTED FOR YIELD ESTIMATION

Like the area estimation, yield estimation is mainly based on interviews with sample holders.

Production per unit area of different sample holders are used to compute the average production of each crop. Total production in each strata is then computed by multiplying average production per unit area, in total area of each crop in a strata.

By using the area and production rates of the relevant strata, total production per province and country is then computed and reported.

6.1 Administrative data used for producing agricultural statistics

Geographical features and boundaries of regional administration units such as roads, watersheds and villages, townships and provinces are effectively used in planning and implementation stages of the survey process.

As mentioned, townships are used as the primary sampling spaces for design and implementation of the survey.

In addition, some registered data[6] such as the records of agricultural imports and exports, extension of agricultural crops and bank loans for development of new farms and orchards as well as data of government subsidies (e.g. due to agricultural disasters such as drought, pests and disease) are used for production and updating of the relevant annual information in province and national levels.

[6] Including registered data

6.2 Methodology adopted for producing agricultural statistics

The available administrative and registered data such as the agricultural tax data, land ownership records, government subsidies, import/export data, administrative farm registers and other registration or licensing systems, farmers and private businesses data, provide considerable information and high potentials for surveying. But at present, there is no formal and nationally adopted procedure for effective use of administrative data for producing agricultural statistics.

Lack of a national, dependable and standard infrastructure for registered data is an important barrier for this use. However, some registered data are used for estimation of the inter-annual variability of crop areas, improvement and updating of the yields and crop areas in province and national levels. These applications are usually based on the expert guess work and subjective decisions.

6.3 Quality analysis of current statistics focused both on sampling and non-sampling errors

Quality and reliability of current statistics are not formally tested and reported. However, quality of the current statistics from the sampling and non-sampling errors points of view are briefly discussed below.

6.4 Sampling errors

Sampling errors arise when selected samples do not perfectly represent the population of interest. Here, agricultural holders are used as the sampling units and agricultural areas, costs and yields of targeted crop types (e.g. wheat, barley, rice and other crops) are the variables of interest. These errors may be due to use of outdated agricultural census data for stratification and sample size allocation; inadequate sample size; problems in allocation of sample sizes to different strata; problems of stratification; faulty choice of sample villages and holders; errors in choice of statistics or substitution of sampling units by the enumerators. Inadequacy of the total sample size, use of the old agricultural census data (as the base information), arbitrarily substitution of sampling units by the enumerators, and difficulties of access to the target sampling units. These problems deserve particular attention.

6.5 Non-sampling errors

Non-sampling errors mainly arise from human errors, such as mistakes in choice of method or procedure. These are usually divided into two types of response and non-response errors. Response errors arise due to inaccurate answers by respondents (respondent errors), or misinterpretation of their answers (errors by interviewer).

Because of the shortage of expert enumerators, budget inadequacies, misallocation of the target sampling areas and low levels of interest among agricultural holders in interviews, non-sampling errors can have a greater influence on the quality of the outputs than the sampling errors. Errors caused by unwilling respondents, mistakes in selecting the most useful respondents, misinterpretation of results and possible cheating by enumerators can be significant. (Getting data from the agriculture holders can be difficult since they don't have adequate training to provide such reliable information).

However, adverse effects of human and sampling errors are more important in sampling projects than in agricultural census projects. That's because errors in the latter projects are mainly local and these are not used for extrapolation to larger areas. However, due to generalization and multiplication of sampling data, any error in sampling is exposed to considerable amplifications. Therefore, the required training, experience and skills in agricultural sampling projects is considerably higher than that of the agricultural census.

7. DIFFICULTIES FACED IN THE IMPLEMENTATION OF THE CURRENT METHODOLOGIES

The main problems faced in the implementation of the currently adopted methodologies include:

- difficulties with interviews, agriculture holders lack interest and confidence in interviews;
- financial and budget limitations on surveys;
- issues with accessibility of sample villages and sample agriculture holders, due to physical conditions (roads, terrain), climate and transportation conditions;
- shortages in trained and expert surveyors and inefficiencies in training programmes because of high rates of change of jobs and responsibilities in the corresponding responsible organizations.

8. CURRENT METHODOLOGIES' GAPS AND LIMITS

8.1 Lack of spatial-based procedures and outputs

Existing methods do not provide dependable spatial data about locations and distributions of crops and their yields. Therefore, the use of outputs is limited to decision making and import/export management problems at the national and, to some degree, provincial levels. Results would suggest serious deficiencies for applications such as large-scale agricultural monitoring, early-warning systems and/or studies of spatial patterns of agricultural productivity, efficiency and costs.

8.2 Current survey' incompleteness

Regardless of their importance, the current survey procedures do not include livestock and orchards. Information about these are mainly provided by province-level expert knowledge, local surveys and subjective judgements. In addition, there is no regular procedure for crop forecasting.

8.3 Lack of timely and cost effective procedures

Cost effectiveness can be defined a producing good results with minimum cost. Good results are timely and acceptable, or have the required qualities that satisfy user needs. The current approaches are capable of controlling costs, but use time consuming procedures meaning these do not lead to on-time products. For example, survey results are usually published six months after collection. Yet, regardless of the need for more frequent agricultural statistics, these data are provided once per year and usually with delays of more than six months. Only past data is available and required data of present and future must be inferred from data of the past.

8.4 Relying on subjective judgments and agricultural holders' interviews

Existing agricultural data (crop areas and yields) provided by the current methods are mainly based on the interviews and statements from farmers and subjective expert judgments. Due to the influence of other factors, such as difficulties in locating agriculture holders, a lack of trust and confidence in the enumerator, and inadequate knowledge about the subjects and their importance means that results are exposed to considerable biases and errors.

8.5 Relying on the old census data and sampling techniques

The dynamics and rapid change in agricultural data in Iran, and reliance on old sampling techniques and census data with update frequencies of 10 years means high uncertainties for selection of the sample villages and agricultural holders. In fact, reliability of the enumeration data and the resulting list frames show a decreasing trend as a function of the age of census data and rate of the change of rural population and agricultural holders.

8.6 Lack of efficient and standard procedures for validation and quality checks

Quality control, quality assurance and quality assessment of survey outputs play an important role in defining the usefulness of outputs. Subjectively defined data, and data from interviews means there are many obstacles and difficulties in validation and quality assurance. Further, there is no procedure for quality control and quality assurance of these data.

8.7 Inconsistency and high variability of spatial errors and uncertainties

Iran is composed of large geographic areas, with different cultural and socio–economical structures. Consequently, knowledge, concerns, interests and trust levels among interviewed agriculture holders can vary significantly. Therefore, results may represent inconsistent, low rates of repeatability and complex structures with high spatial variabilities. Although these results are used for presentation and decision-making at provincial and national levels, they are cannot be used similarly at local levels. In addition, because of the heterogeneity of the environment and related socioeconomic structures, risks of systematic errors are much higher than those of the random errors.

8.8 Lack of a GIS-based and integrated software system and inefficient use of the available Information and web technologies

Agricultural surveys are an important part of the agricultural management system and include different tasks, such as the data collection, storage, access, dissemination, use and analysis. The existing software is mainly based on the specific and separate tasks and not on the efficient use of available Information and web technologies. Development of a GIS-based integrated software for various tasks, from data collection to management and analysis, would yield considerable advantages and unified quality control, by avoiding costly data conversion, and exchange of data between different software. Development of an information system to formalize and support activities of the agricultural survey would significantly reduce time and cost requirements of operations. Many available administrative and particularly registered data are capable of providing valuable information with a minimum time and cost requirements. However, these data are not being effectively used.

8.9 Limited capabilities and inefficient use of outputs

Outputs are suitable for limited uses, rather than multiple uses as expected. Applications of the outputs are very limited and mainly include:

- production of reports, particularly the annual bulletin of agricultural crops and products (in province and national levels) at high management levels
- decision-making in import/export problems
- studies on agricultural economics.

Obstacles to wider use of outputs include its aggregate form, lack of spatial details and lack of information on quality – which hampers knowledge of its fitness for use.

Output may be improperly used, for example, in data for applications that require higher or lower levels of qualities. That can lead to unreliable results and decisions, or inefficient uses of data.

8.10 Limited use of the geospatial technologies

Geospatial technologies – particularly remote sensing, GIS and GPS – provide valuable capabilities for effective sampling, data collection, management and use. Advantages of integrated use of these three technologies in agricultural statistics include:

- analysis of the spatial structures and spatial autocorrelations
- optimization of the sampling design
- more effective planning and facilitation of the fieldworks
- quality management and assurance
- production of crop types and yield maps
- better crop estimates and prediction
- spatial analysis and visualization of patterns of crop types and yields
- time, costs, sampling errors and uncertainties reduced
- detection and monitoring of changes and trends.

8.11 Lack of specific organization, budget, sustainable plans for training and capacity building in agricultural surveys

Stable organizational structure and dedicated staff are essential elements for success in sustainable information systems. Training and capacity building are also important, yet there is no sustainable training programme related to the agricultural survey. Although many training programmes have been held on the subject in the past, instability and frequent turnover in responsibilities has meant that most of these programmes have not been effective in training permanent enumerators in the required qualifications.

PART 3

User expectations of a new data collection method

Identification of areas for improvements in methodology adopted at country level and in the area frame

9. ANALYSIS OF PUBLIC AND PRIVATE CURRENT AND POTENTIAL USERS OF AGRICULTURAL STATISTICS IN THE COUNTRY

Current users of the agricultural statistics include:

- Ministry of Agriculture Jahad, for agricultural planning and developing, implementing, monitoring and evaluating the agricultural policies.
- Statistical Centre of Iran (socio-economic studies and national computations and indicators, analysis of trends).
- Management and Planning Organization of Iran.
- Ministry of Power for Planning and Management of Water Resources.
- Ministry of Interior for Planning and Evaluation of the rural development programmes
- Universities, researchers and research organizations.
- Ministry of Industry, Mines and Trade for Import and Export Regulations and Planning.
- Ministry of Health and Medical Education (food safety, regulations and other health related problems).
- Iran Meteorological Organization (climate related analysis and studies).
- Environmental Protection Organization (emission control, for monitoring of the current land use, land use changes, biodiversity, sustainability, environmental aspects of agriculture and other conservation objectives).
- Local governments and private companies.
- Public and farmers.
- Land use planners and managers.

10. ANALYSIS OF USERS' NEEDS. SET OF CORE STATISTICS, CORRESPONDING ACCURACY, GEOGRAPHIC DOMAIN AND FREQUENCY

Needs of agricultural data users are very diverse and include a wide range of temporal and spatial resolution, precision and data attributes. In an ideal situation, agricultural statistics are used as the input for a diverse subjective or objective processing techniques and objectives, the outputs which are used for decision making. So, by providing a thorough knowledge of crop acreage, yield and related maps, spatial distribution of agricultural holders, holdings, practices and activities, effective monitoring and regulations would be possible. Main applications of agricultural statistics may be grouped into six classes as follows:[8]

- Food security (FS)
- Quick response for disasters (QR)
- Economical aspects of agricultural products, particularly import, export, market and pricing analysis (EAP)
- Environmental, sustainability and regulation aspects of agricultural resources and activities (ER)
- Monitoring and management of agricultural activities (MM)
- Agro-systems, Integrated analysis and optimization (AIAO).

Table 10 - Agricultural statistics for main applications

Code	Spatial	Temporal	Precision	Attributes	Data type	Time-liness	Eary detection/ Prediction
FS	national and provincial levels	every 6 months	high	total yields of different crop types	tables	very important	very important
QR	high resolution	monthly in the growing season	moderate	maps of affected areas	maps/ tables	very important	very important
EAP	national and provincial levels	annual	moderate	total yields of different crop types	tables	important	important
ER	medium to high resolution	every 3 years	moderate	change of agricultural fields and structure-pesticides, fertilizers-agricultural practices	maps/ tables	important	less important
MM	varying resolution (high to medium)	annual	moderate	total area and yields of different crops	maps/ tables	very important	less important
AIAO	varying resolution (high to medium)	annual	high	total area and yields of different crops; agri-environmental and socio- economical data	maps/ tables	important	less important

[8] Annual statistics may not be necessary for all agricultural variables

Regarding the above applications, the required agricultural statistics can be classified as described in Table 10. Timely information is very important; prediction(s) are very important; development of reliable crop forecasting methodology is very important; crops with high annual variation are important for the study and sustainable production.

11. NEEDS' SATISFACTION LEVEL ANALYSIS THROUGH CURRENT METHODOLOGIES

Levels of satisfaction[9] of different users depends on the gaps between the times, attributes, precision and spatial details of what is required and what is provided. The available statistical data is published as the annual reports in three volumes. Volume 1 shows various agricultural crop areas and production at national and province levels.

Volume 2 consists of reports on various subjects, including agricultural machinery, industries and supporting services; livestock, poultry and aqua-cultural production; natural resources, including forests, ranges and watersheds; and education and research activities in the agricultural sector.

Volume 3 covers orchard products. Reports of all three bulletins provide very detailed information about crop types, but with poor spatial details which are limited to the province level. By considering the report types and user needs as described in section 2, it's clear that the lack of timeliness, spatiality and accurate data are the three most important deficiencies of the current statistics.

Their importance depends on the application and user. Although the required accuracy and reliability is very important in almost all applications, for time and spatiality the importance is very application dependent. For example, time is the most important parameter in food security and pricing activities. Yet, time-consuming processes mean agricultural statistics are usually delivered with six- to twelve-month delays. Therefore, all needs requiring timely, spatial data in more detail than the available province levels would not be satisfied.

By providing detailed information on crop types, fruit types and the other reported attributes applications like food security.

© FAO

9 Including the area-frame test

12. DESCRIPTION OF POSSIBLE AREAS FOR IMPROVEMENT IN CURRENT METHODOLOGY FOR THE DIFFERENT KINDS OF AGRICULTURAL STATISTICS

With regard to the gaps and limitations of the current methodology, possible areas for improvement can be listed as follows:

Development of reliable crop forecasting methodology. Lack of crop forecasting is an important gap of the current methodology. Development of a reliable crop forecasting methodology is vital for many applications particularly for food security, import, export and pricing in varying scales from local to national.

Adaptation of geospatial-based technology and cost effective procedures. The current procedures rely on subjective judgments and interviews with agricultural holders; and on old census data and sampling techniques. More effective use of the registered data is recommended.

Completion of the survey attributes.
At present, livestock and orchards are not surveyed.

Development of efficient and standard procedures for validation and quality checks for meeting different users requirements.

Development of a GIS-based and integrated software system for efficient use of the available Information and web technologies.

Development of sustainable plans for training and capacity building in agricultural survey.

Description of possible areas for improvement in the current area frame.

Estimate costs for the area frame–travel and survey cost components.

Estimate spatial autocorrelation function for the different crops. It may differ for different geographic domains and it's based on the current area frame data (in this case, complemented by aerial photos and remote sensing data). It checks the spatial resolution of data available in Google Engine, as the spatial autocorrelation function is strongly scale dependent). It is also very dependent on time and data. Depending on the time and spectral bands of consideration, results may be different.

Analysis of administrative data and geospatial data for zoning of the country; zoning by soil, terrain and precipitation.

Analysis of zoning, using geo-referenced data collected on the ground. in order to develop the stratification (stability of zones, their capacity to minimize the variability inside the strata and maximize the variability among the strata, the sample size, etc.)

Analysis of the available geospatial information to be used for ground surveys. Geospatial technologies includes remote sensing, photogrammetry, cartography, GIS, GPS and information technology (IT).

These technologies deal with the acquisition, storage, processing, production, presentation and disseminationof geoinformation.

GPS allows surveyors to collect the precise locational information and increase the accuracy in mapping. The use of tablets or handheld pen computer (Arcpad) equipped with flat touch-sensitive screen, allows surveyors to draw and click directly into the screen to update map information while on the field.

PART 4

Application of remote sensing in agricultural data collection

13. SATELLITE IMAGE ANALYSIS INCLUDING CLASSIFICATION AND NORMALIZED DIFFERENCE VEGETATION INDEX (NDVI)

13.1 Satellite data preparation

After obtaining shape files of training sites in Zanjan province, based on crop calendar and cloud cover conditions of the region, Sentinel-2 satellite images were downloaded from the Earth Explorer site (Earthexplorer.usgs.gov). Unfortunately, satellite data showed clouds covering the entire area, and only one scene was selected for further processing as shown in Figure 31. From all available scenes, only 11 were cloud-free, on the dates showed in Figure 32. Satellite data for most suitable dates of the growing season from December 2017 until February 2018 were not available. Data were examined for their probable geometric and radiometric distortion; no geometric distortions were found. Radiometric corrections were made using relative radiometric correction as discussed by R. Jensen (2014). In Figure 33 single-image normalization using histogram adjustment was used for radiometric correction.

FIGURE 31 - FCC image of Zanjan overlaid by country boundary shape file

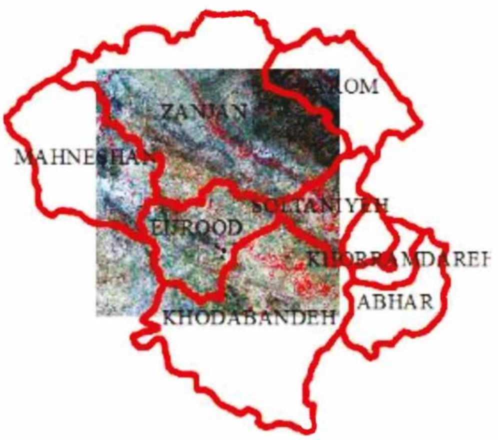


FIGURE 32 - Eleven Sentinel-2 FCC image of different date covering part of Zanjan province

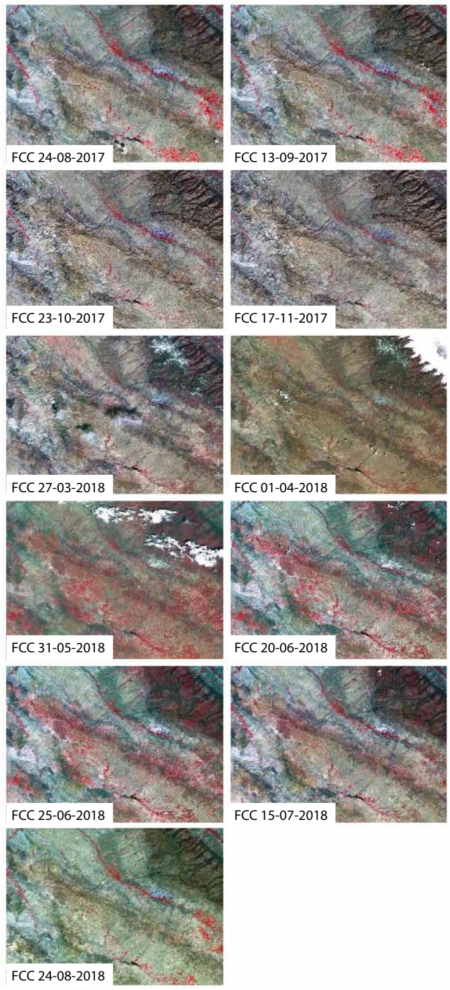
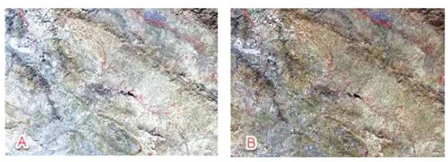


FIGURE 33 - Sentinel-2 FCC image before (A) and after relative radiometric correction (B)



13.2 Image classification

FIGURE 34 - Maximum likelihood classification (MLC) outputs

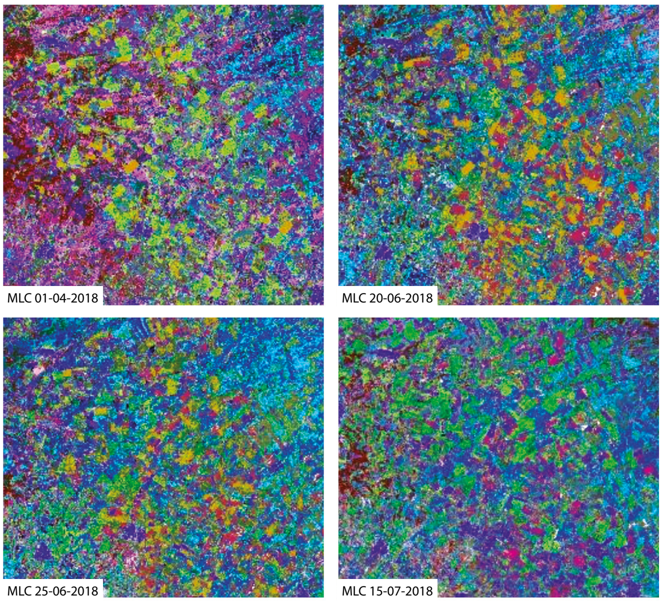
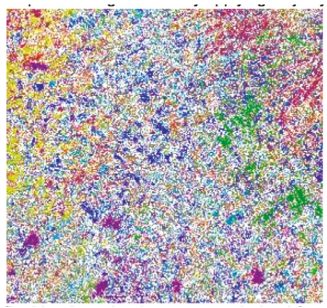


FIGURE 35 - Final land cover use map



Satellite image analysis is the process by which information content of an image is retrieved in the form of maps or statistical tables.

Thematic maps are produced through satellite image analysis either object-based or pixel-based image classification. Image classification is among pattern recognition algorithms that may make use of in situ information to train classification algorithms (supervised classification) or it may not use any training information to classify image (unsupervised classification).

Maximum likelihood classification (MLC) is a well-known and widely used strategy among supervised classification schemes. It benefits from statistical information of land use/land cover training sites to obtain their signatures and uses a probability density function to classify (assigning a class number) every unknown pixel exits in an image.

To implement supervised classification, digitized training sites were divided into two parts and overlaid on FCC images and statistics (including mean, standard deviation, minimum, maximum and range) for different land use/land cover were extracted. Classified output images of different date are shown in Figure 34.

Kappa coefficient of agreement is calculated from training sites and presented in Table 11.

Final land use land cover map could be generated by applying majority voting function (Figure 35).

Table 11 - Kappa coefficient of agreement

Date for MLC output image	Kappa
23-10-2018	0.54
17-11-2018	0.54
27-03-2018	0.60
01-04-2018	0.59
31-05-2018	0.65
20-06-2018	0.64
26-06-2018	0.65
15-07-2018	0.67
24-08-208	0.62
Average	**0.611**
Majority	**0.85**

13.3. Algorithm development based on normalized difference vegetation index trend (NDVI)

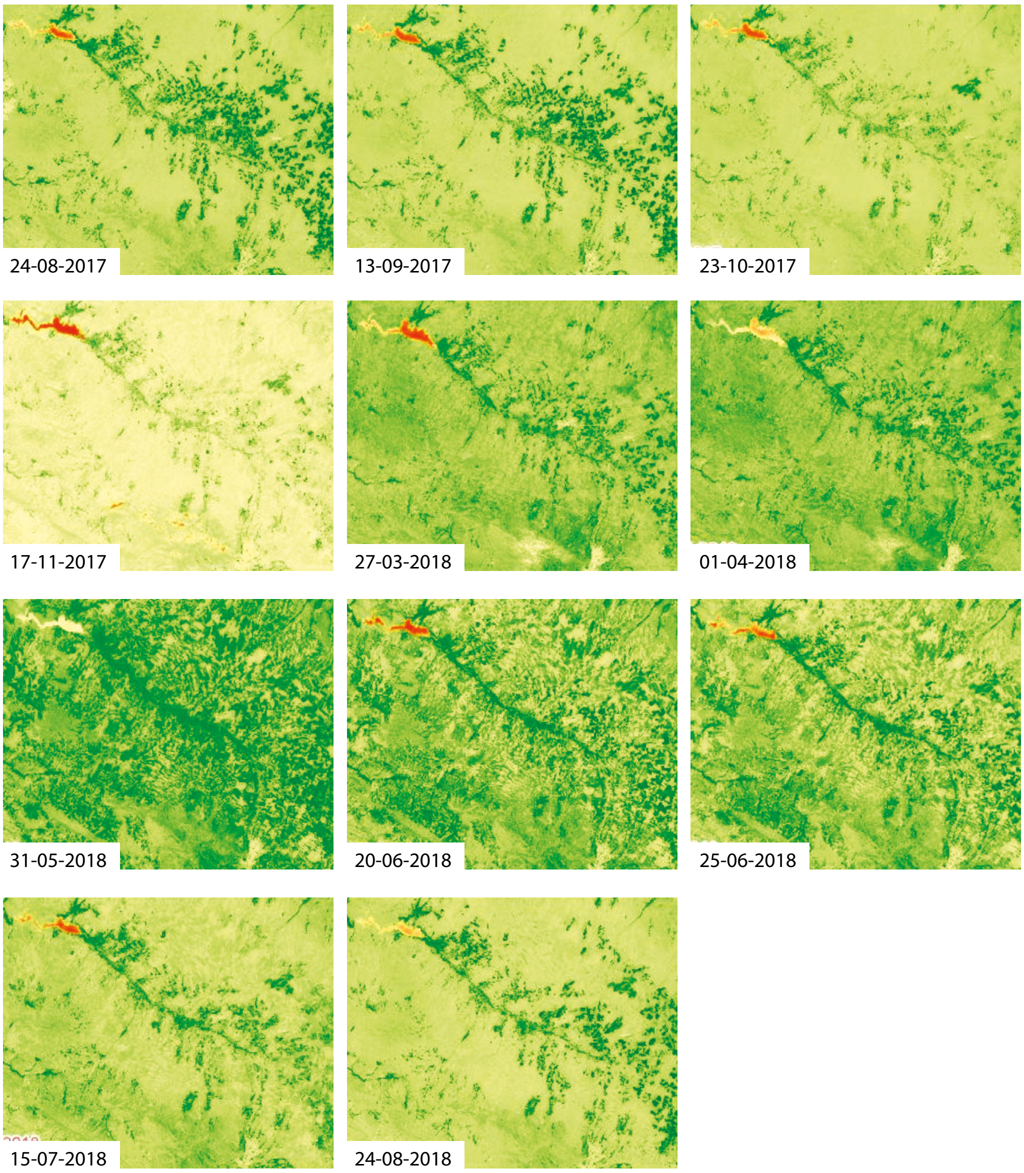
Another way of extracting information from satellite images is to develop and derive Indices indicating the presence of land use/land cover in every pixel of image. One of the widely used vegetation indices is normalized difference vegetation index (NDVI) which uses RED and NIR reflectance bands to generate NDVI image through following formula:

$$NDVI\ (Sentinel\text{-}2) = (\rho NIR - \rho RED)/(NIR + \rho RED)$$

NDVI values range between -1 to +1. Positive values more than 0.2 indicate the existence of vegetation cover. Figure 36 shows a trend of NDVI changes for a part of Zanjan scene.

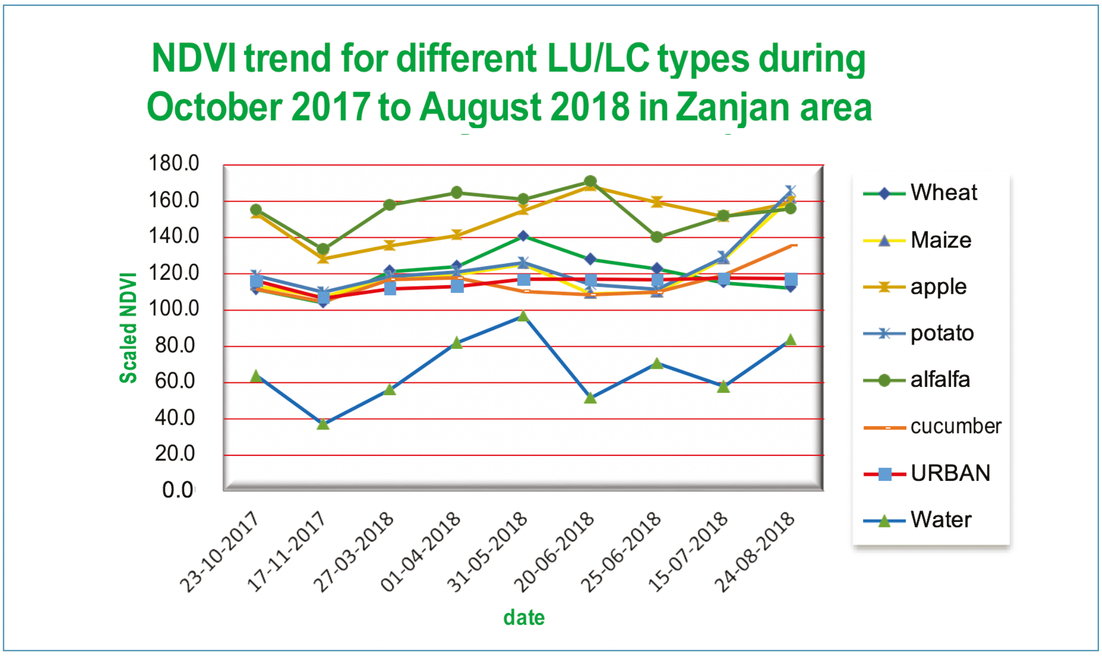
NDVI images of 24–08–2017 and 13–09–2017 belong to end of previous season and will not be consider for further processing.

FIGURE 36 - Trend of NDVI changes for a part of Zanjana scene (water body in red and agriculture lands in deep green)



Digitized training sites were divided into two parts and overlaid on NDVI images and statistics (including NDVI's mean, standard deviation, minimum, maximum and range) for different land use land cover were extracted through zonal statistics tool box in an ArcGIS environment. Figure 37 shows a graph of NDVI mean values for different land use/land cover in different dates.

FIGURE 37 - Graph of NDVI mean values for different land use/land cover in Zanjan area for varius dates



14. MODELLING WITH NORMALIZED DIFFERENCE VEGETATION INDEX TREND

14.1 First approach: modelling with percentage normalized difference vegetation index trend

The trend of mean NDVI values for water at the bottom of graph is clearly distinct from the others. At the top of the graph, alfalfa and apple are distinct not only from each other but also from other LU/LC. Wheat in the middle of graph shows mean NDVI trend not similar to others. Urban areas also show an approximately uniform mean NDVI trend during different dates. In the first approach, one can benefit from one percent of LU/LC or crop NDVI mean values in different dates as indicated in Figure 38 by arrows. Different crop NDVI mean values in Figure 51 are listed in Table 12.

FIGURE 38 - Selection of +/- 1 percent of NDVI values for modelling LU/LC or crops in Zanjan scene

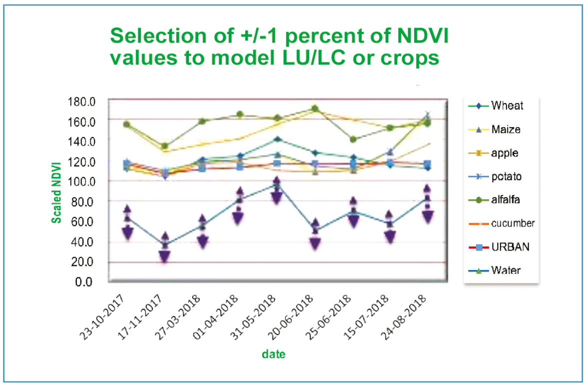


Table 12 - NDVI mean values from different LU/LC crops from 23–10–2017 to 24–08–2018

CLASSNAME	23-10-2017	17-11-2017	27-03-2018	01-04-2018	31-05-2018	20-06-2018	25-06-2018	15-07-2018	24-08-2018
wheat	111.42±1%	104.13±1%	121.19±1%	123.99±1%	140.67±1%	127.83±1%	122.76±1%	114.98±1%	112.25±1%
maize	113.61±1%	107.45±1%	117.94±1%	119.77±1%	125.34±1%	109.25±1%	110.02±1%	128.36±1%	161.59±1%
apple	153.12±1%	128.44±1%	135.64±1%	141.12±1%	155.00±1%	168.24±1%	159.36±1%	151.60±1%	159.36±1%
potato	118.73±1%	109.58±1%	118.34±1%	120.80±1%	126.03±1%	113.84±1%	111.13±1%	129.13±1%	165.73±1%
alfalfa	155.29±1%	133.52±1%	157.88±1%	164.60±1%	160.96±1%	170.68±1%	140.21±1%	151.63±1%	155.88±1%
cucumber	111.54±1%	104.46±1%	116.82±1%	117.82±1%	110.07±1%	108.32±1%	109.89±1%	119.43±1%	135.68±1%
urban	115.97±1%	106.73±1%	111.65±1%	112.93±1%	116.81±1%	116.75±1%	116.46±1%	117.42±1%	117.04±1%
water	**63.57±1%**	**36.73±1%**	**55.83±1%**	**81.79±1%**	**96.51±1%**	**51.30±1%**	**70.44±1%**	**57.72±1%**	**83.42±1%**

To model LU/LC or crop status in each pixel of final crop map, one can use the following models:

- If (NDVI_date1 within ±1% of mean NDVI_water and NDVI_date2 within ±1% of mean NDVI_water and..... NDVI_daten within ±1% of mean NDVI_water) then LU/LC is WATER
- If (NDVI_date1 within ±1% of mean NDVI_wheat and NDVI_date2 within ±1% of mean NDVI_wheat and..... NDVI_daten within ±1% of mean NDVI_wheat) then LU/LC is WHEAT.

14.2 Second approach: similarity measure (distance between known and unknown normalized difference vegetation index trend values)

This approach is based on distance between mean of NDVI values of sample sites (LU/LC or crop) and each unknown pixels of NDVI image. Mean NDVI values are extracted from crops training sites.

To classify each pixel in its proper category, a similarity measure (i.e. euclidean distance measure) based on euclidean distance was used. The following formula calculated the euclidean distances between mean NDVI of crops with pixel NDVI value:

ED_NDVI_CROP_C_ pixel(i,j) = $\sum_{k=1}^{N}(Mean_{NDVI_{CROP_{C_{D_k}}}} - PIXEL_{NDVI})^2$

where:

k is the date number

N is the number of dates that NDVI were calculated

ED_NDVI_CROP_C_pixel(i,j) is the resultant output that shows distance for crop C

$Mean_{NDVI_{CROP_{C_{D_k}}}}$ is mean NDVI for class C in date k.

Graphical representation of this approach is given in Figure 39.

FIGURE 39 - Euclidean distance (ED) between mean NDVI tend and NDVI of unknown pixels

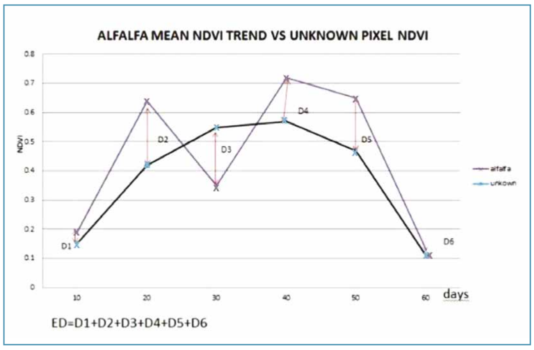
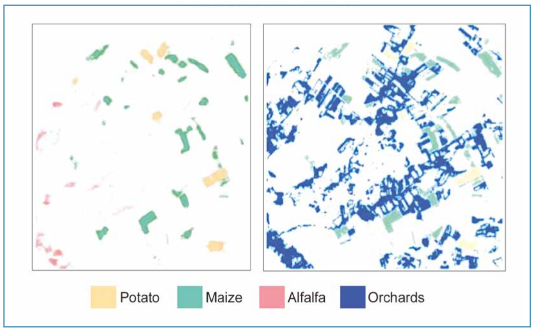


FIGURE 40 - Outputs obtained modelling with euclidean distance approach. In some places, orchards and alfalfa coincide which indicates lands used for simultaneous cropping



14.3 Third approach: spectral angle mapper as a similarity measure

Spectral angle mapper (SAM) is a similarity measure that calculates the angle between known and unknown vectors, as shown in Figure 41. Here, r_i is mean NDVI value for crop at time i and t_i is unknown NDVI at time i.

The known vector here is mean NDVI values of each crop and unknown vector could be any NDVI pixel value in NDVI trend images. Figure 42 shows output images of applying SAM and ED to map maize crop.

FIGURE 41 - Spectral angle mapper (SAM) concept adapted from R. Jensen (2015)

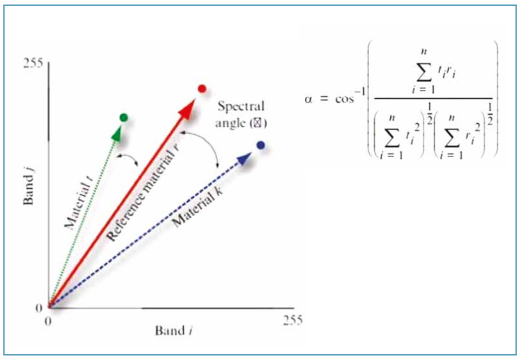
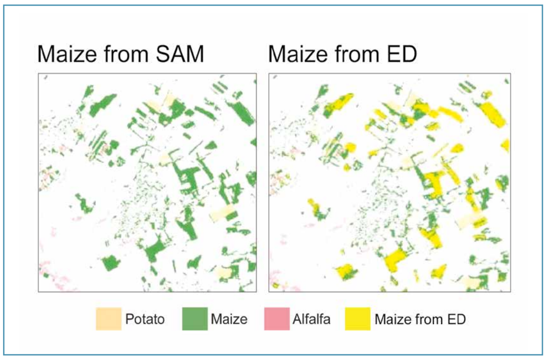


FIGURE 42 - Image with spectral angle mapper (SAM) and euclidean distance (ED) map layer overlaid



14.4 Fourth approach: the centre of gravity and area function

The centre of gravity

This approach introduces the centre of gravity (CG) which is also called centroid. Its position should be fixed in relation to the shape. If a shape is represented by its region function EQ. 1, its centroid ($gx;gy$) is:

This measure was applied to the NDVI images of Zanjan.

$$\begin{cases} gx = \frac{1}{N}\sum_{i=1}^{N} xi \\ gy = \frac{1}{N}\sum_{i=1}^{N} yi \end{cases} \qquad \text{EQ.1}$$

Figure 43 shows its output image. Value of each pixel in this image shows the center of gravity of its NDVI images. This image can be either classified or compare with CG of different crops shown in Table 13 through euclidean distance to give each pixel a proper class.

FIGURE 43 - NDVI different crops and their centre of gravity. This new measure can differentiate crops through their different CG.

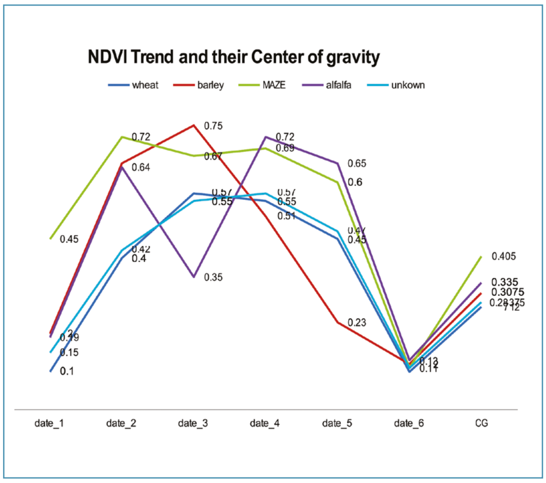


Table 13 - Center of Gravity (CG) of different crops

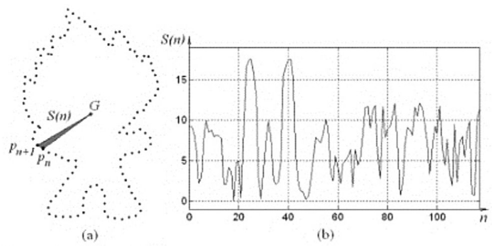
Class name	Mean	Class name	Mean	Class name	Mean
wheat	100 466.647	apple	127 917.832	cocumber	96 737.179
fallow	92 464.444	potato	111 985.641	plow	95 221.324
maize	102 782.451	alfalfa	129 054.794	bare	98 876.482
urban	97 208.333	water	55 740.023	mont	93 157.179

14.5 Area function

When the boundary points change along the shape boundary, the area of the triangle formed by two successive boundary points and the center of gravity also changes. This forms an area function which can be exploited as shape representation. Figure 44 shows an example. Let $S(n)$ be the area between the successive boundary points P_n, P_{n+1} and center of gravity G.

This function is under development for the next phase and will be provided in coming report.

FIGURE 44 - Area function concept



(a) Original contour; (b) the area function of (a).

PART 5

Previously implemented pilot project in Hamedan Province

Description of the area frame test to date. Analysis of its gaps and limitations

Two technical cooperative area frame sampling programmes have been tested in Iran. One is by the International institute for aerospace survey and earth science (ITC) and started in 1995; the other is from FAO, and began in 2004. Both projects have been implemented in Hamadan Province which is the study area.

The first project, "Development of crop inventory and forecasting system for the major agricultural commodities in Hamadan province, Islamic Republic of Iran", has focused on integration of area sampling frame (ASF) and remotely sensed data for crop acreage estimation and crop yield forecasting. The main concern of the second project, "Agricultural survey improvement programme" (ASIP), has been to test ASF and its integration with the list frame (LF) for acreage, yield and production costs estimation. Overall designs of these two projects are outlined in Figure 45 and 46.

15. SAMPLING AND SURVEY METHODOLOGIES FOR AREA ESTIMATION

The tested area frame approaches have been based on the stratified probability sampling and statistical inference. The crop area and production forecasting system as designed in the first project is composed of three main processes: area estimation, yield estimation and production forecasting. These are briefly described in Figure 45.

Process 1. Area estimation (crop inventory): in this process, various techniques and data from different sources, such as remote sensing, field observations and historical data, are used to derive area estimates of major agricultural crops.

Process 2. Yield forecast: crop growth simulation models and detailed data on soil, weather, crop physiology, crop management and historical production data are used to derive periodical estimates of yields of various agricultural crops.

Process 3. Production forecasting: crop area and yield estimates derived from processes 1 and 2, are used to compute the periodical production forecasts for the lowest administration units. Data from these units are then aggregated to produce the relevant and required information in higher administration levels.

The survey design in the second project, as outlined by Bouzaffour and Sharifi (2006), consists of a stratified probability sample of segments with a replicated selection procedure and includes three major components. Those are: the area sampling frame (ASF); the list sampling frame (LSF); and the dual sampling frame, as presented in Figure 46.

FIGURE 45 - Conceptual presentation of the crop area and production forecasting system (after Sharifi, 2000)

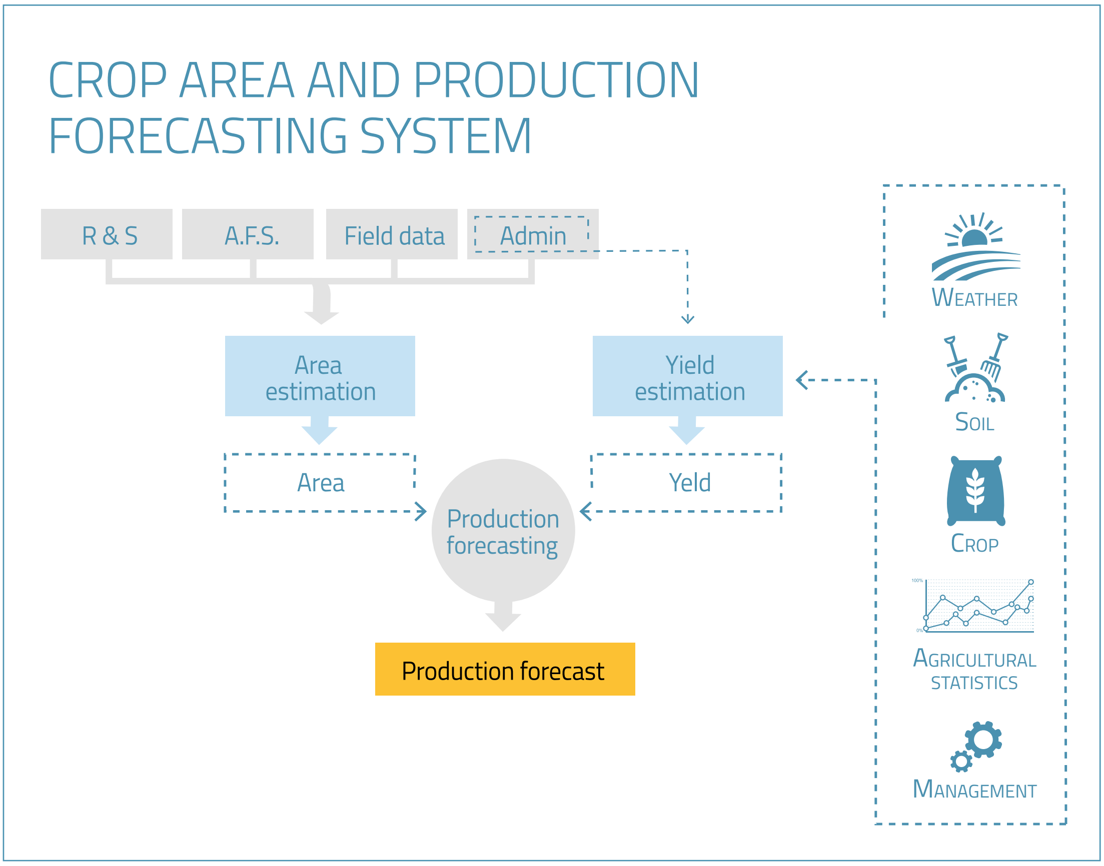
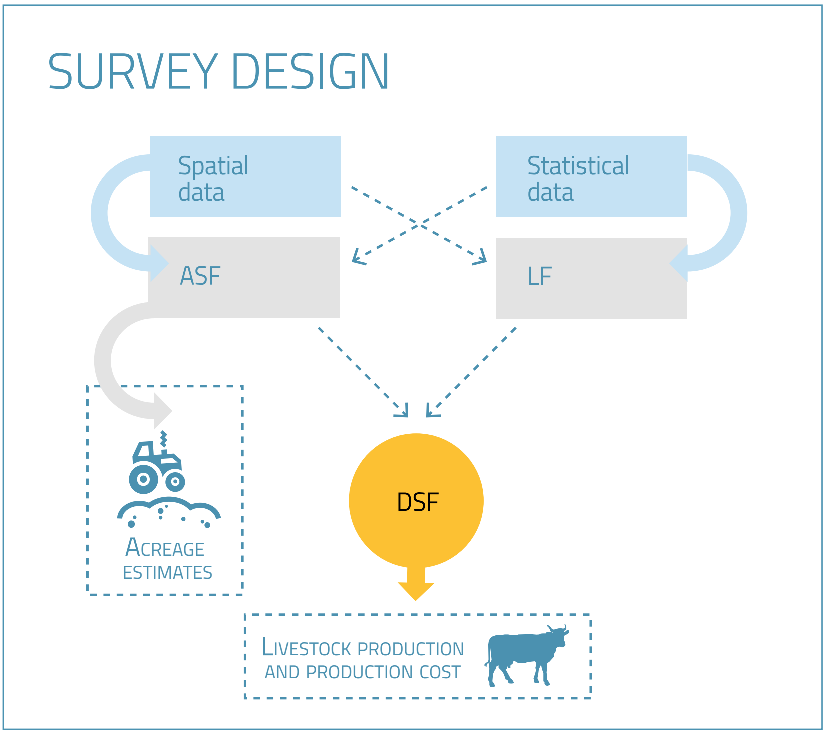


FIGURE 46 - Conceptual presentation of the survey design (after Bouzaffour and Sharifi, 2006)



The ASF uses recent satellite data to stratify and directly identify and measure the area covered by each crop in each selected segment in the ground, and estimate the acreage of each crop. The list frame is used to collect data from special holdings to make a significant contribution to the total estimate of some important survey variables such as the livestock. This represents a complementary list of special holdings, such as those who own largest total area, the holdings with the largest area for a given crop or those with the largest number of livestock and poultry. The dual frame combines the ASF and LSF. The data from the list complements those gathered from the area frame for each variable.

As compared to either of the lists and area frames, this approach can provide interesting advantage. For example, because of the check with data of ASF, influences of subjectivity and the resulting errors in the list frames can be reduced. In addition, list frames are limited to special holders and collection of detailed data of list frames is not required. Also, variables such as the livestock or production costs which are not measurable by the area frames, can be effectively estimated. Two approaches of closed and weighted segments have been used to associate segments (sampling units) with the reporting units (holdings or tracts) and define values of the survey variables for each segment.

For the closed segment method, the value of a variable in a segment is simply the sum of values in each of the tracts of the segment. For a given variable and segment, data on the totality of a holding is not needed – except for holdings totally included in the segment. The closed segment method requires the collection of data (by direct observation or by personal interview with the holders) corresponding to agricultural activities that are found physically within the boundaries of the segment. If information on land use is required, data is collected on land use within the boundaries of the segment. Effective use of the closed segment depends on having photographic enlargements of a known scale or scale drawings of the segment to control data collection.

Therefore, response and coverage errors are relatively low. Sample estimate of a survey variable y, in closed segment is calculated as:

$$Y_c = \sum_{h \in s} \sum_{j \in B_h} \sum_{k \in G_{hj}} e_{hjk} \sum_{m \in T_{hjk}} t_{hjkm}$$

where:

Y_c is the sample estimate of a total for the survey variable y

S is the set of all land-use strata

B_h is the set of all substrata in stratum h

G_{hj} is the set of all segments in substratum j of land-use stratum h

T_{hjk} is the set of all tracts in segment k of substratum j or land-use stratum h

e_{hjk} is the expansion factor for all tracts in segment k

t_{hjkm} is the tract value for the variable y associated with tract m.

The closed segment is the normally accepted method to estimate planted crop areas and has the major advantage (compared with the weighted and open segment methods) of being independent of the definition of a holding so that it eliminates the ambiguities in ascertaining the land contained within a holding. However, for most variables of a multiple-purpose survey, the closed segment method is not applicable since the tract is not the appropriate reporting unit.

For the weighted segment method, the variable in each tract is defined as the value of the variable in the holding multiplied by a factor equal to the ratio between the area of the tract divided by the area of the holding. Then, the value of a variable in a segment is the sum of the variable in each of its tracts. For the weighted segment method, data is collected for the entire holding associated with any land area in the sample segment. The data collected from each such holding is weighted by the proportion derived by dividing the area of the holding that is within the segment by the entire area of the holding (both inside and outside the segment). Sample estimate for variable of interest is computed as:

$$Y_w = \sum_{h \in s} \sum_{j \in B_h} \sum_{k \in G_{hj}} e_{hjk} \sum_{m \in T_{hjk}} w_{hjkm} y_{hjkm}$$

where:

Y_w is the sample estimate of a total for variable y

S is the set of all land-use strata

B_h is the set of all zones in stratum h

G_{hj} is the set of all segments in zone j of land-use stratum h

T_{hjk} is the set of all tracts in segment k of zone j of land-use stratum h

e_{hjk} is the expansion factor for all tracts in segment k (inverse of the probability of selection)

W_{hjkm} is the weight used to prorate Y_{hjkm} (usually area of tract m divided by the total holding area for tract m)

Y_{hjkm} is the value for the variable Y associated with tract m.

The effect of this method is that data for each holding is pro-rated among the segments in which it is located. For instance, if ten percent of a holding area is included in the segment, ten percent of the total holding data recorded on the questionnaire will be assigned to the segment and thus to the summarization process. In contrast with the closed segment, the weighted segment estimator can be used for all survey variables like livestock, costs and yield, since the holding is the reporting unit required.

Crop area estimation as reported by Bouzaffour, and Sharifi (2006) and Sharifi (2000) included five processes, namely:

- stratification of the region
- preparation for establishing primary sampling units
- identification of the selected primary sampling units
- identification of secondary sampling units (SSUs)
- area estimation through remote sensing

where the last process (area estimation through remote sensing) was implemented in the first area frame test. These processes are briefly described below.

15.1 Stratification of the region

Identification of the homogenous areas in term of use types and patterns of agricultural fields have been the main objectives of the stratification process.

Stratification is included the following activities:

- study and examine the existing land use maps as developed by ASID in 1999, and ITC in 1997. This was carried out by super imposing the land use maps over the 1/100 000 orthophoto Landsat ETM image of the province acquired in 2002. Results were further compared with the photo mosaic of the SPOT-5 images of the province (with spatial resolution of 10 meters) acquired during the study in 2005;
- enlargements of the SPOT images were also used for validation. As these processes did not lead to satisfactory results, new stratified map was developed as follows: Manual image interpretation of IRS–WIFS (with spatial resolution of 180 meters) acquired in September 2005, for delineation of the following major strata:

    1. irrigated agriculture
    2. rainfed agriculture
    3. orchards
    4. rangelands
    5. non-agriculture
    6. urban

- super-impose the resulting stratified map on top of the Landsat ETM orthophoto map, and manual improvement of the stratification through on-screen visualization and digitization;
- print the stratified map, control and improve it with the help of higher-resolution satellite images (SPOT-5, 2005 enlargements);
- calculate the area covered by each stratum.

15.2 Preparation for establishing primary sampling units

Primary sampling units were defined through the following activities:

- preparing the base data for manual delineation of PSUs. These data were provided by a print-out of the Landsat ETM 1/100 000 scale orthophoto-map containing the stratified boundaries, as well as the administration boundaries;
- estimating the number of PSUs in each stratum. This was carried out by deciding on the size of secondary sampling units (SSU). Then assuming around 10–20 SSU in each PSU, the number of sampling units for each stratum was estimated. Results of this process are presented in Table 9;
- preparing the required material, and training three local teams to delineate PSUs by considering the natural boundaries in each stratum. In this process, administration boundaries as well as the stratum boundaries were assumed to coincide with the physical boundaries;

- delineation of PSUs on the stratified map, which was super-imposed on Landsat ETM 1/100 000 scale orthoimage. During this process, the stratification was revisited and improved. Resolution of Landsat ETM orthoimage was relatively suitable for completion of this process in the rangelands, rainfed agriculture and irrigated lands;
- but because of the heterogeneity and complexity of the structure of orchards, data with higher resolution were required. Therefore, the 1/10 000 data provided by the SPOT-5 image was used for delineation of PSUs of orchards (Figure 33);
- transfer the PSUs on the digital stratified map by scanning the stratified maps and on-screen digitization of all PSUs;
- check the quality of the transfer (PSUs from map to stratified digital map) and correct where necessary. Prepare the required information as presented in Table 9 for sampling from PSUs.

Sample selection: The objective of a sample survey design is to provide estimators with small variances at the lowest cost. The following three factors are important for effective design and sample selection:

- the importance of each stratum (total number of elements)
- the variability within each stratum
- the cost of sampling and observations in each stratum.

Since the number of elements in each stratum affects the quantity of information in the sample, a large sample size has been assigned to strata containing large numbers of elements. The variability has also been considered, because a larger sample is needed to obtain good estimates of population parameters when the observations are less homogeneous. To solve the problem, a specific judgment has been made to determine the sample size and allocate them to different strata. In this process, the importance of each stratum has been based on the main variables such as the annual crops and the survey costs.

Because there were no reliable estimates on the population variance and it was difficult to estimate in advance, a rule of thumb was used to determine the percentage of the study area for sampling (JRC draft papers I and II). On this basis, the sample size and sampling rates were specified. The sample size and rate in each stratum are highlighted in Table 10. In this process, trade-offs (ratio) between the inputs (cost and resources) and outputs (accuracy) has played a major role. As a result, a compromise was found and the final decision was to take 181 sample units in different strata (Table 14). Variability and optimal sample size for each stratum could be calculated from the survey data.

To distribute the selected number of samples in the study area, a replicate sampling method was applied. This approach was used because of ease of modification in future; and replicate sampling removes estimators' bias. The use of replicated sampling method also allows the overall sample size to be increased or decreased by adding or deleting replicates in order to achieve a desired precision. The replicate number is coordinated with the desired sample rotation.

Table 14 -The area of each stratum and its related total and sampled number of SSU's (Bouzaffour and Sharifi 2006)

Stratum	Area (ha)	No. SSU's	SSU size (ha)	Sample rate	Number of selected SSU's
Rangelands	428 085	4 281	100	0.004	16
Rainfed agriculture	814 256	16 285	50	0.005	81
Irrigated	376 137	15 045	25	0.004	64
Orchards	50 498	2 520	20	0.008	20
No agriculture	352 450	0	0	0	0
Total	**2 021 426**				**181**

15.3 Identification of the selected primary sampling units

Primary sampling units included the following activities:

- choose the sampling rate as discussed above for each stratum;
- draw samples in a way that is evenly distributed over the stratum (as discussed above);
- identify the selected PSUs in each stratum;
- check distribution of samples through visual inspection of the printed map, and improve as needed;
- transfer boundaries of selected PSUs on the recent geometrically corrected 1/10 000 scale (2005) SPOT-5 images with resolution of 10 meters; for orchards, the 1/5 000 enlargement of IRS–Pan image was used;
- adjust boundaries of the PSUs to the natural boundaries on the ground for some samples of irrigated agriculture and orchards.

FIGURE 47 - An example of PSU and the related SSUs in irrigated areas

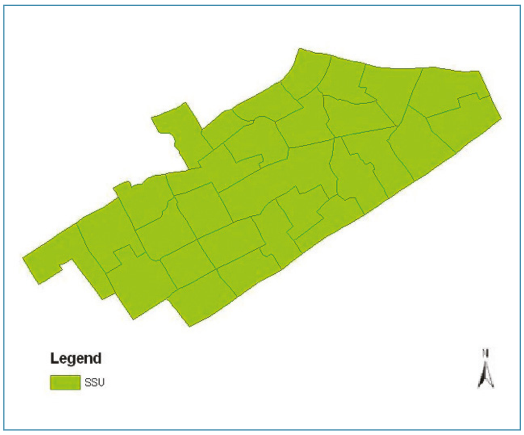
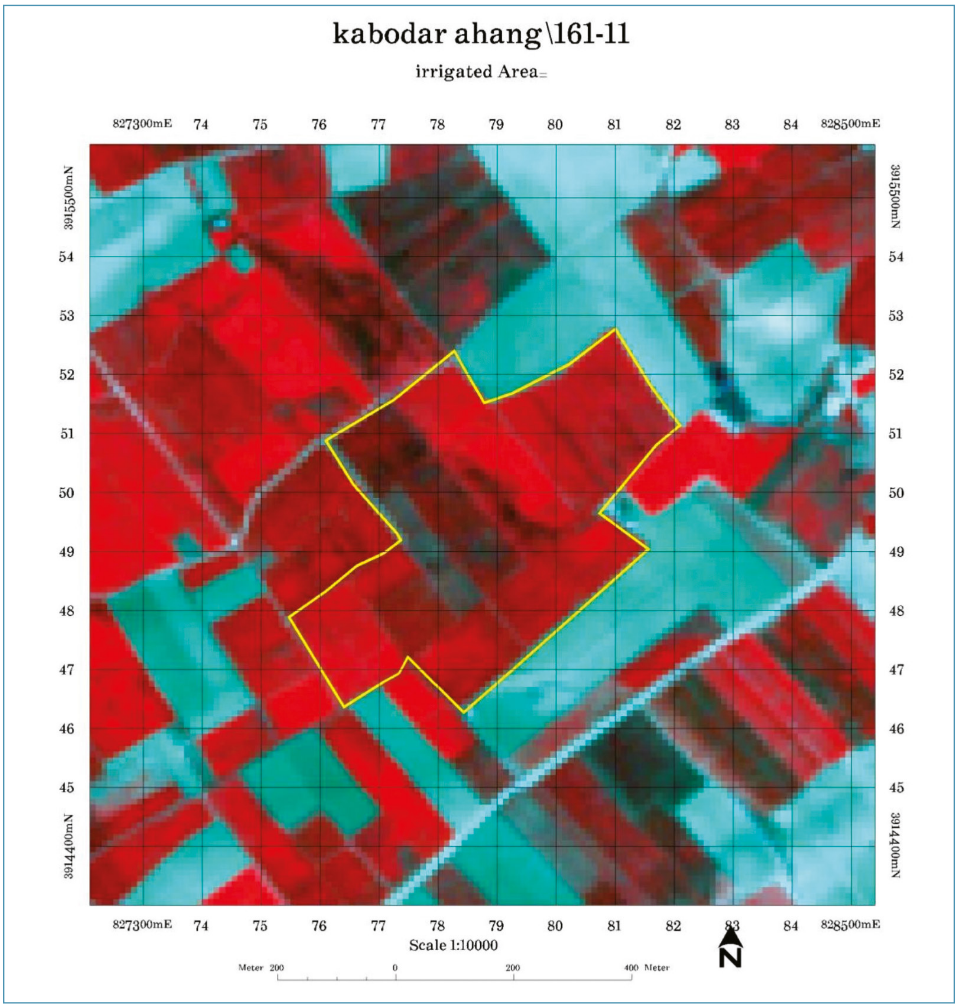


FIGURE 48 - An example of selected sample SSU in irrigated areas



15.4 Identification of secondary sampling units (SSUs)

In this process, each selected PSU has been divided into 10 to 20 SSUs as described below (details in Table 14):

- transfer the boundary of selected PSUs to the related IRS–Pan or SPOT–Pan images, with 5 m or 2.5 m resolution. Adjust the boundary where necessary;
- delineate SSUs in each PSU by considering the natural boundaries;
- select an SSU for field data collection by numbering the SSUs and random selection of one out of all;
- document each selected SSU and prepare for field work.

The stages for sampling design above are valid for both area frame tests. But minor differences include the use of square segments measuring 500 m–700 m in the first study (Sharifi, 2000) as shown in Figure 49; and using irregular segments in the second study as shown in Figure 50 (Bouzaffour and Sharifi, 2006).

Although reasons for a change of the employed segment shapes from regular to irregular were not explained (Bouzaffour and Sharifi, 2006) it may be attributed to better matching of regular shapes with the remotely sensed data in the first study; and the easier location and survey of the irregular shapes with identifiable boundaries in the second study. Easy location of the segments and SSUs are of special importance, especially when enumerators do not have the required experiences and skills.

However, these differences may be justified by considering the first study was concerned with the integration of remote sensing and AFS, whereas main focus of the second project was the combination of list and area frames.

FIGURE 49 - Example of square sample segments

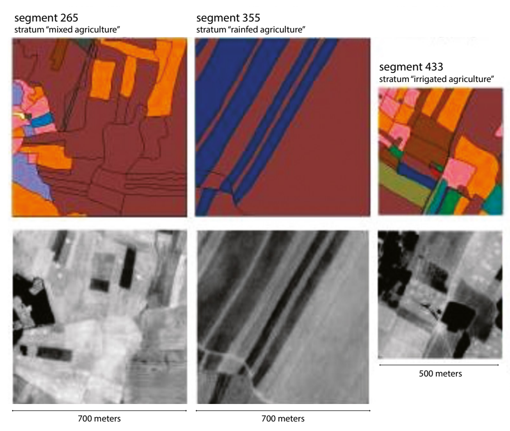
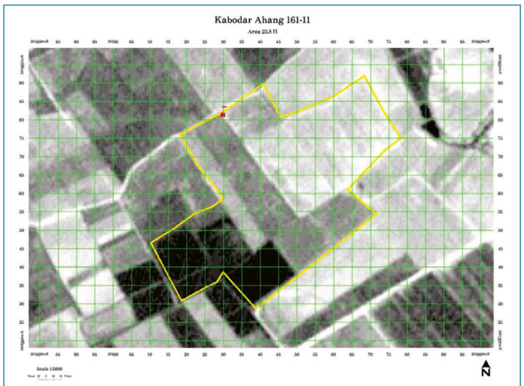


FIGURE 50 - Example of irregular sample segments



Figures 51 to 57 provide detailed data of the location of PSUs and SSUs in the second area and dual frame study (Bouzaffour and Sharifi, 2006).

15.5 Area estimation by remote sensing

This process was based primarily on relating the remote sensing and area frame results, using the regression approach.

After pre-processing and radiometric and geometric correction of satellite data, these data were classified using the sample segments as the training data. Results of the area frame sampling and the image processing were then statistically related and used to derive an improved area estimate per crop in each stratum. By overlaying the strata-based crop area maps with maps of the lowest administrative units, maps of crop areas in different administrative units were produced.

FIGURE 51 - Rangeland PSU's superimposed on ETM Ortho map of 2002 (Hamedan Province)

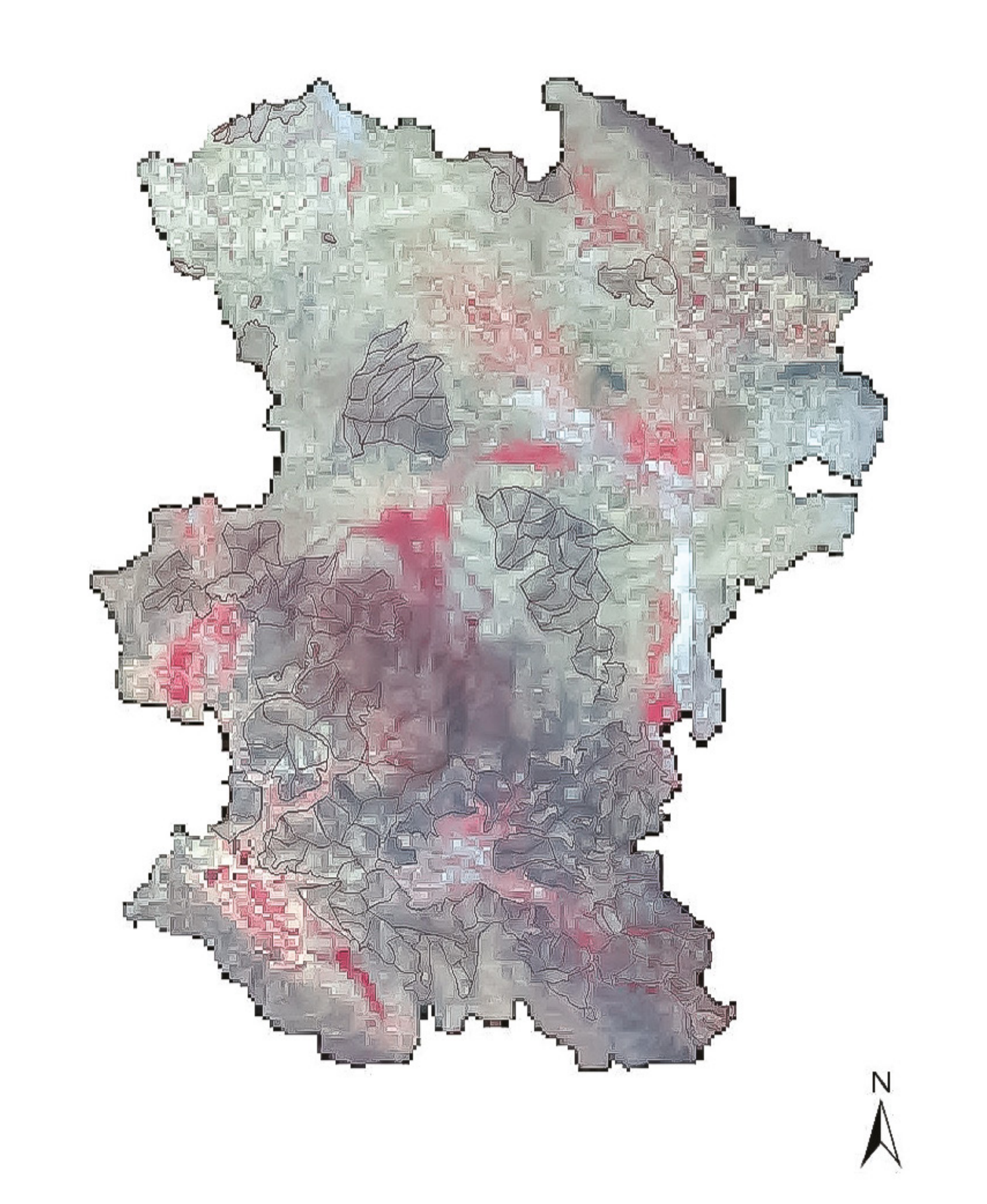


FIGURE 52 - Rainfed PSU's superimposed on ETM Ortho image of 2002 (Hamedan Province)

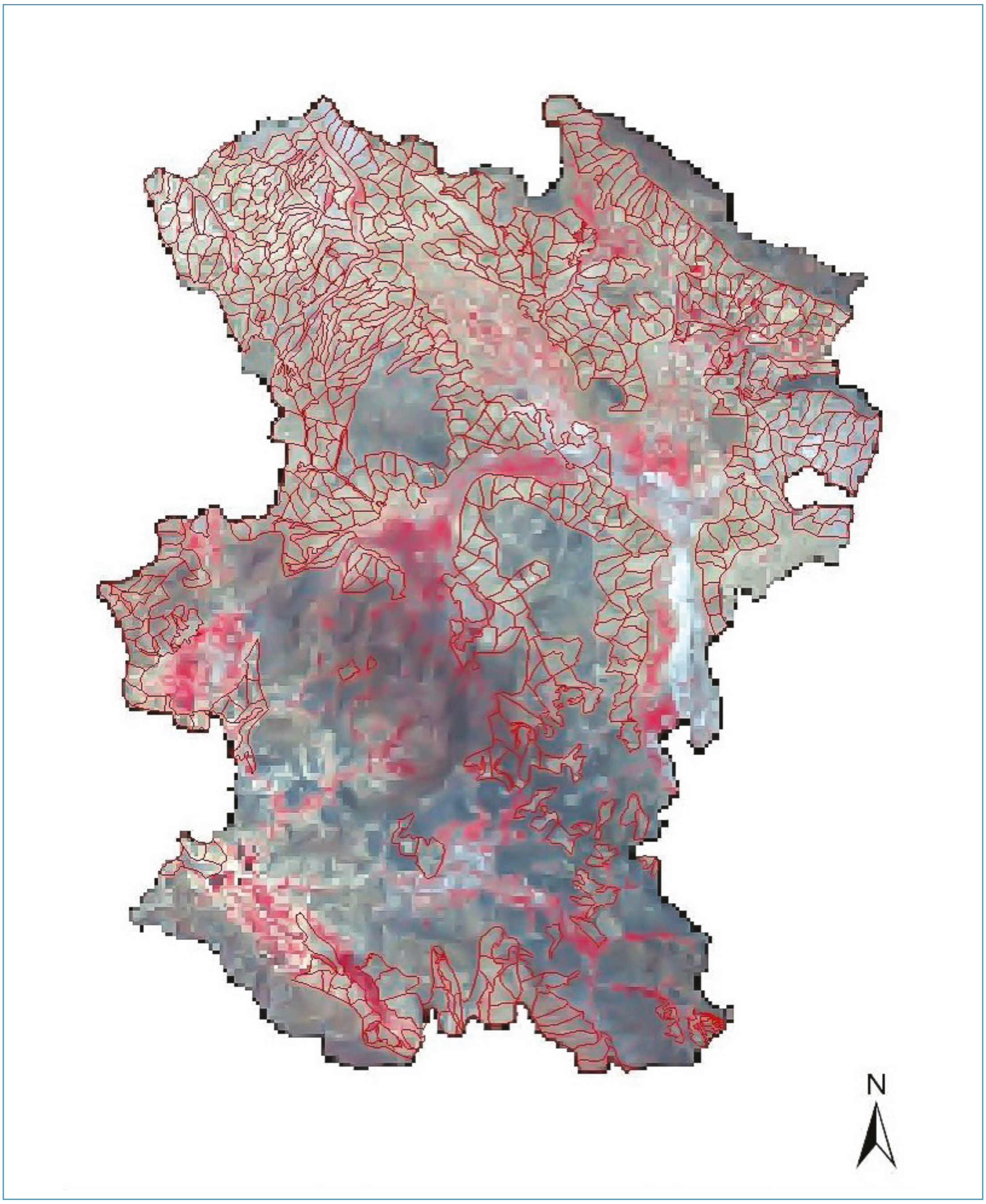


FIGURE 53 - . Irrigated agriculture PSUs superimposed on ETM Ortho map of 2002 (Hamedan Province)

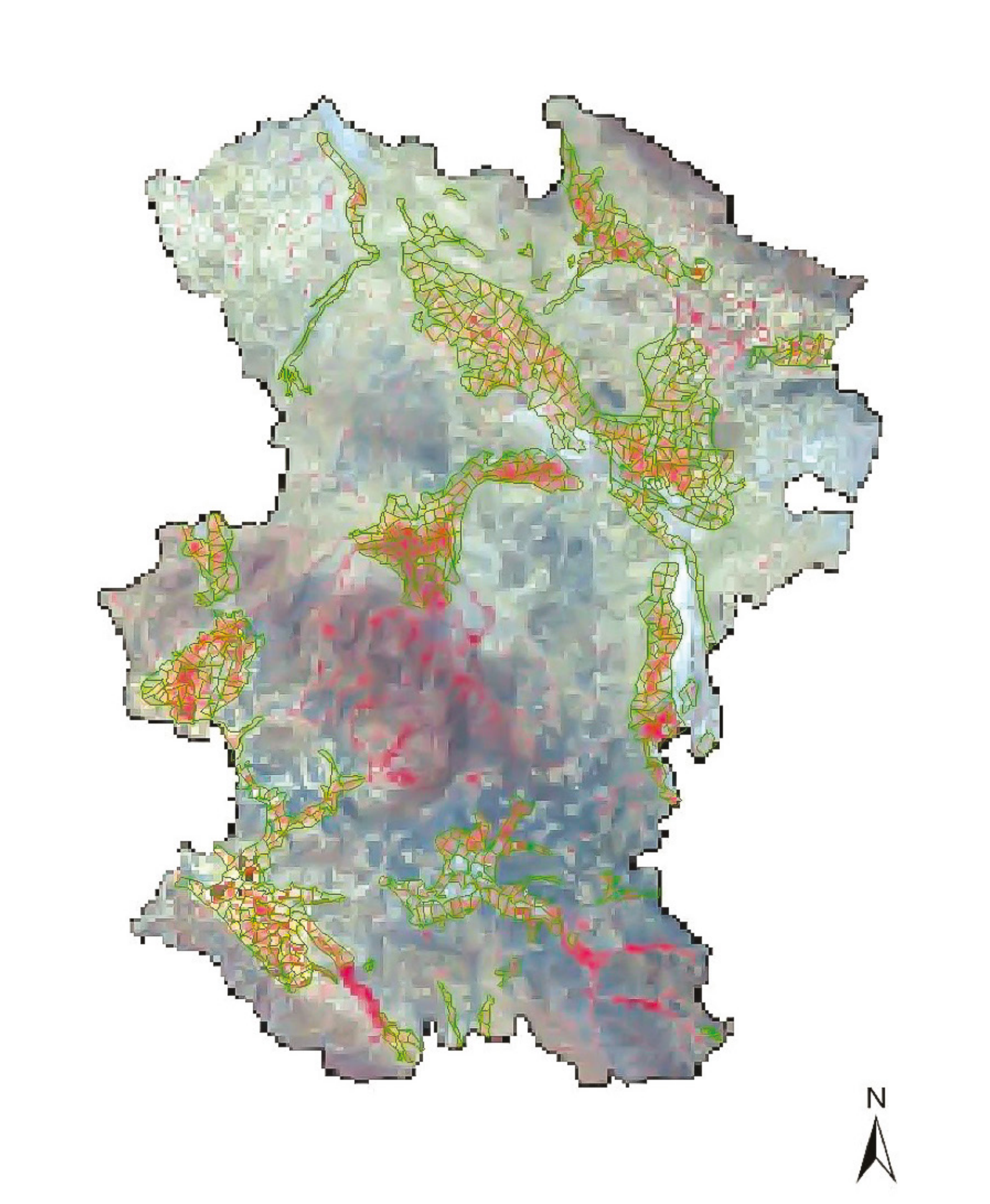


FIGURE 54 - Orchard startum superimposed on ETM Ortho map of 2002 (Hamedan Province)

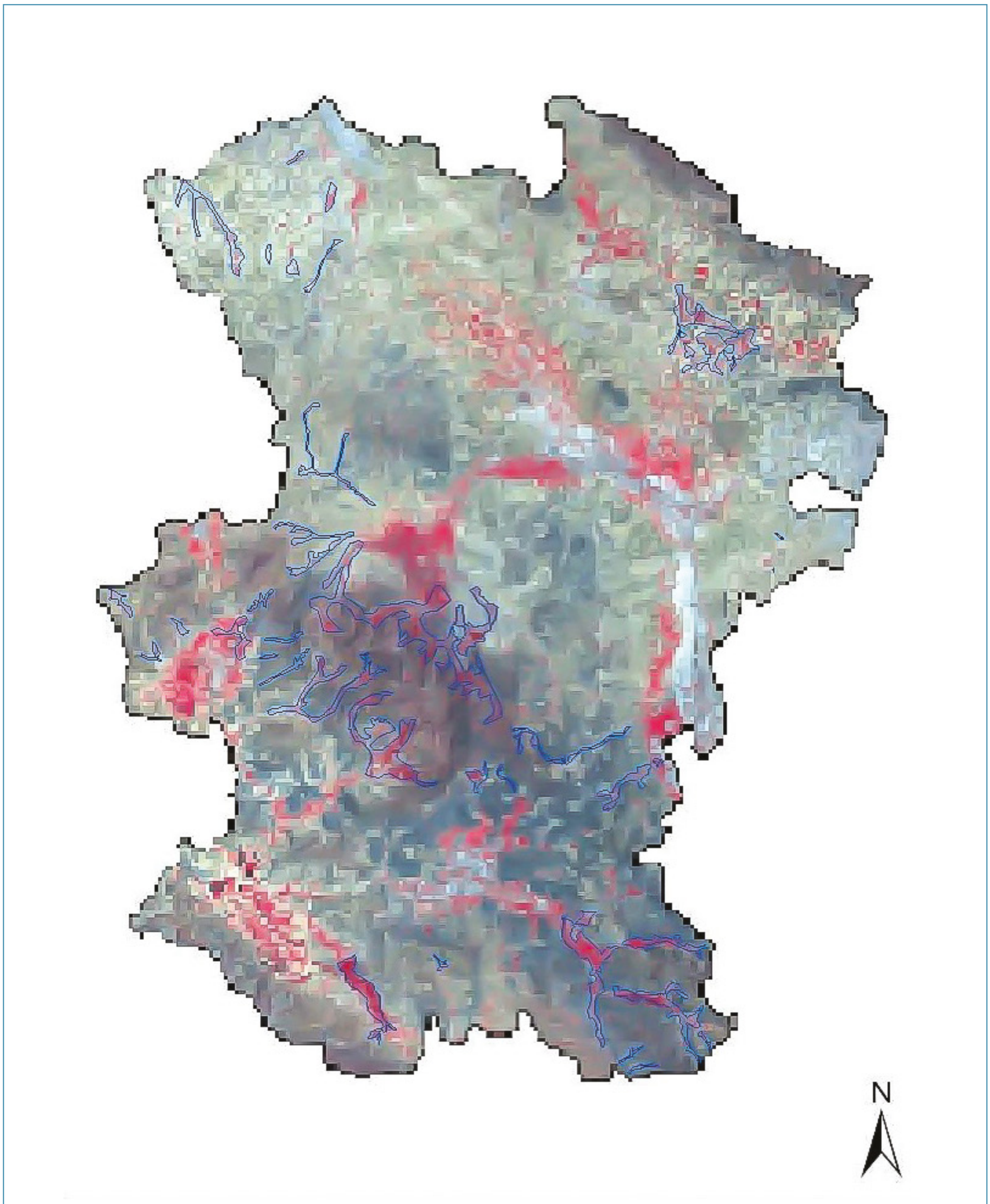


FIGURE 55 - Identified PSUs superimposed on top of the ETM Ortho image of 2002 (Hamedan Province).

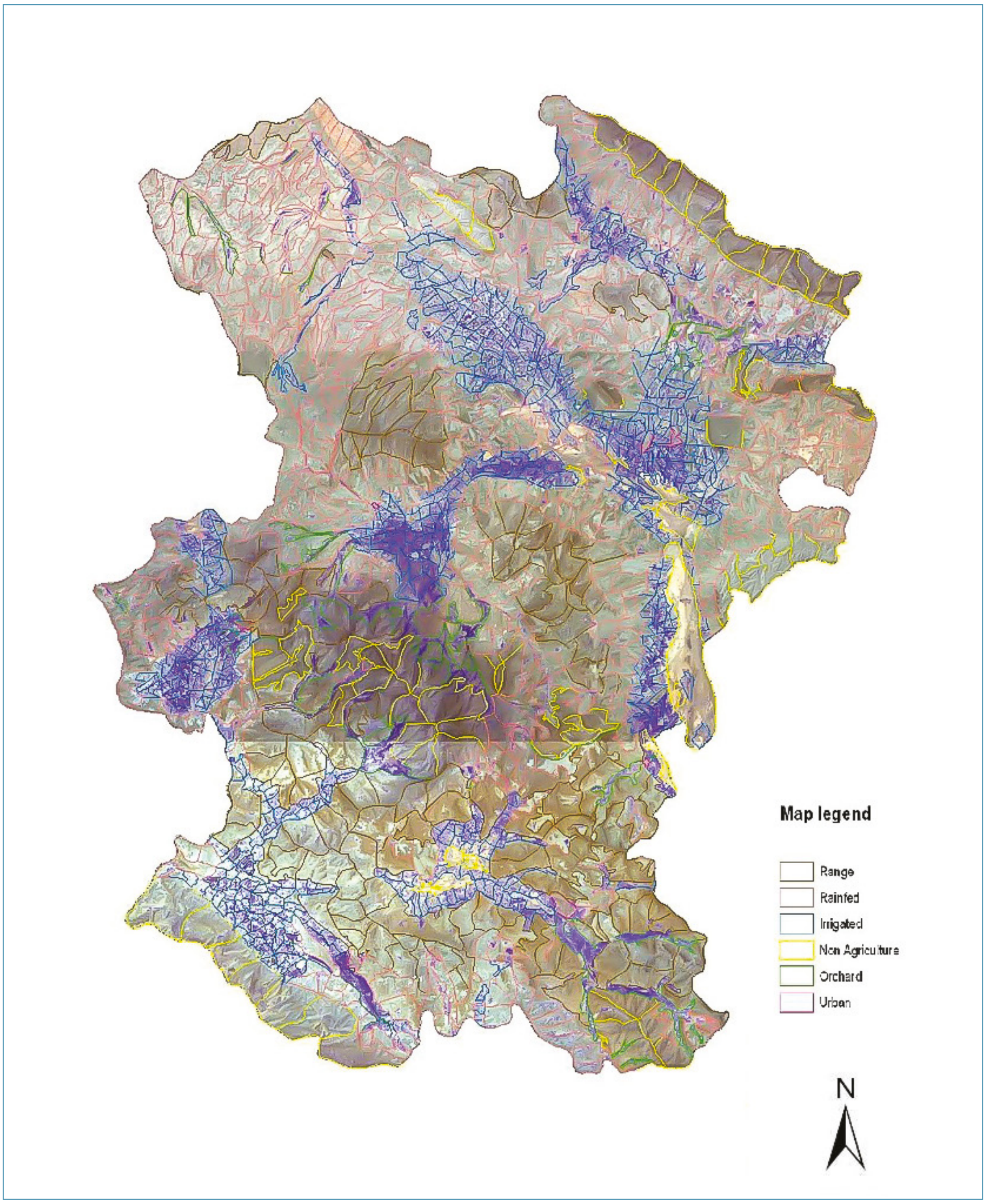


FIGURE 56 - Stratified map, PSUs and selected PSUs in each stratum (Hamedan Province)

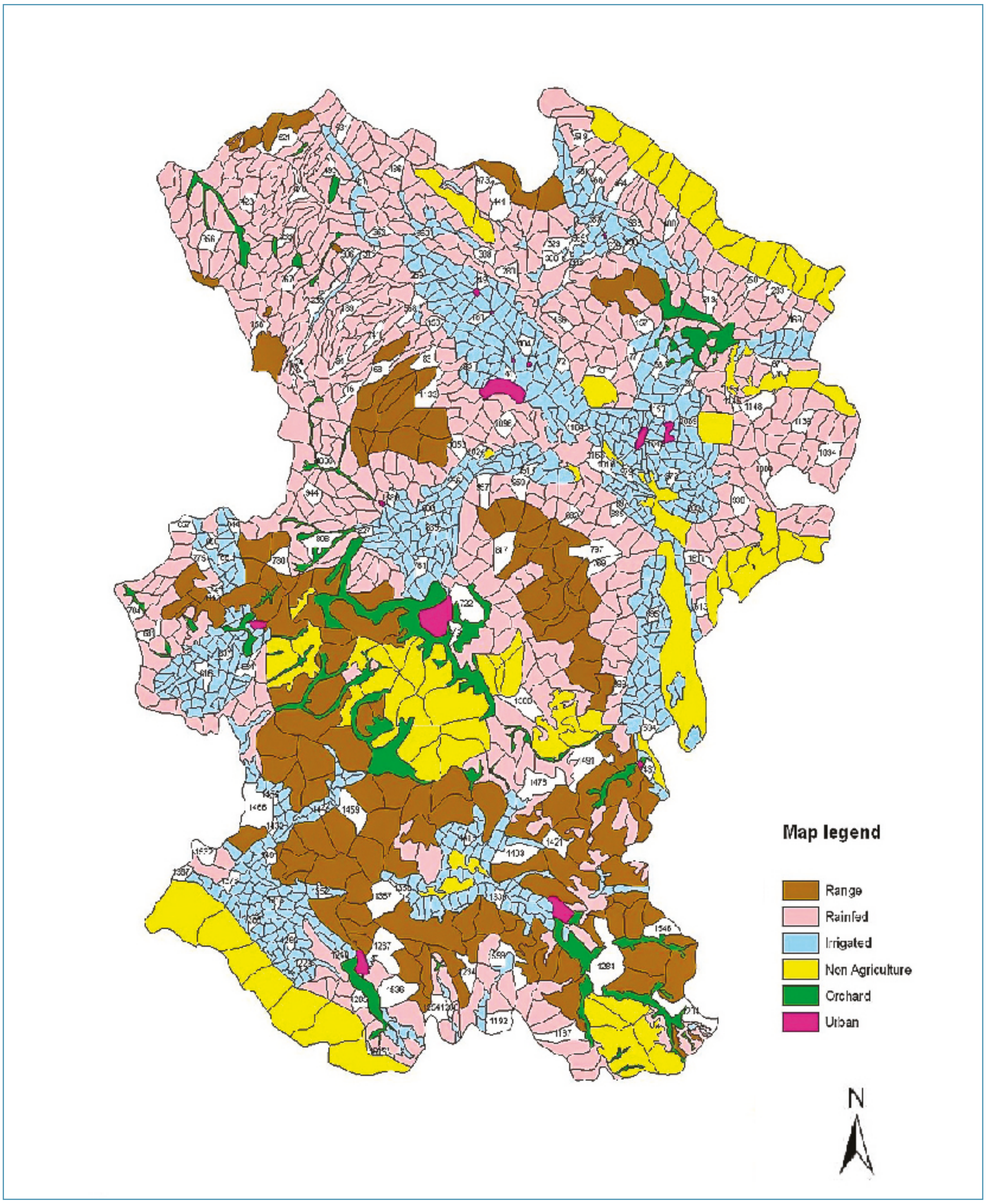
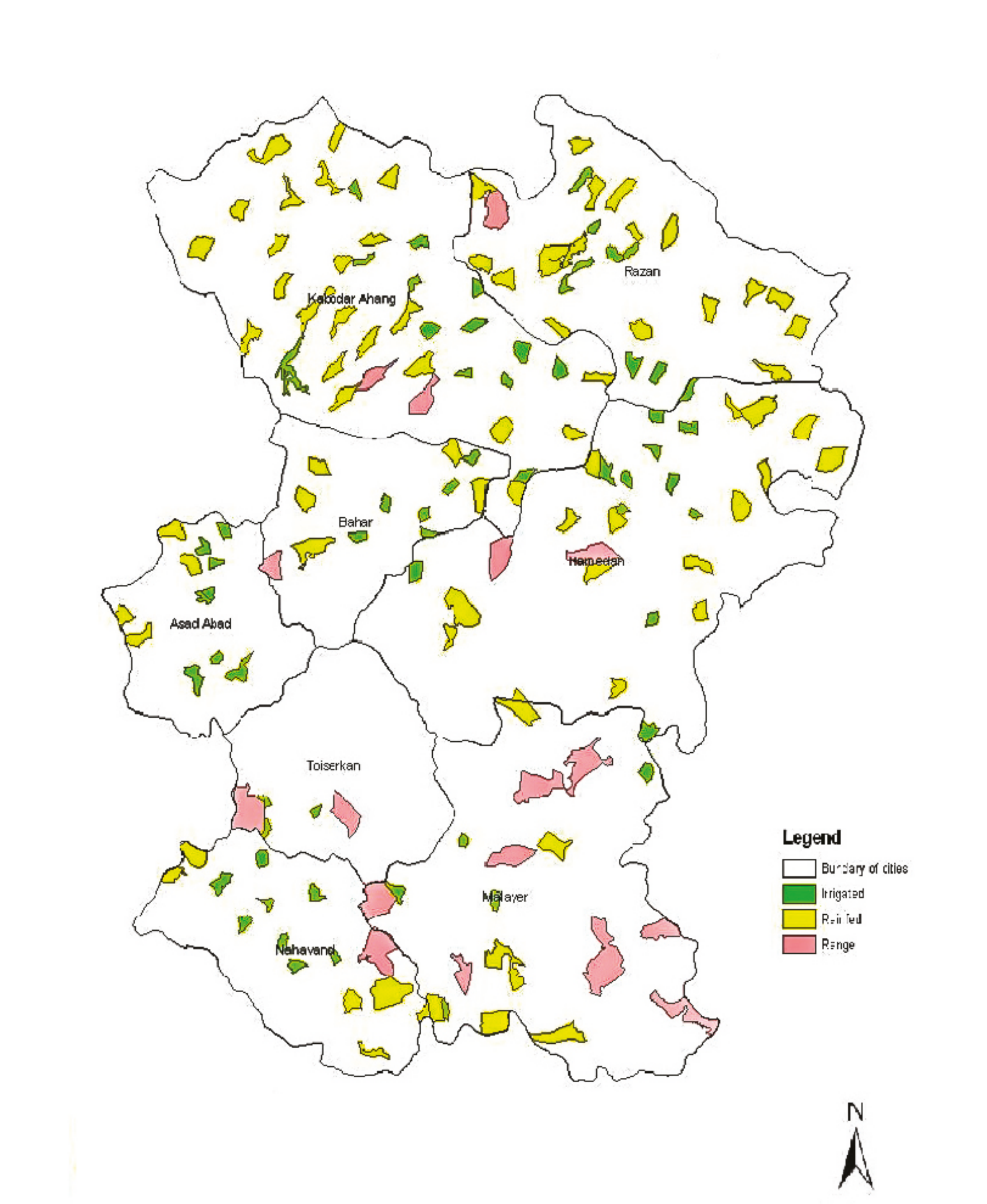


FIGURE 57 - Distribution of all PSUs in Hamedan Province



15.6 Sampling and survey methodologies adopted for yield estimation

Yield estimation has been limited to the most important crops in the region, including the wheat, barley and potato. It has been mainly based on the use of a crop growth simulation model called CGMS (crop growth monitoring system), requiring detailed data of soil, weather, crop physiology, crop management and historical production data. This model simulates the potential yield rather than the actual yield.

Therefore, it has been assumed that there is a strong correlation between the yield potential and the actual yield of crops in the region. Implementation of the model has been based on the homogeneous areas called the "simulation units" (SUs) defined by combination of spatial data of crop type, soil physical group (SPG), rooting depth (RD) and weather characteristics.

The calibration and validation of this model requires intensive data on crop phenological/genetical properties, crop management such as the crop calendar, as well as the historical daily weather data, physical, and chemical properties of soil and the actual historical yield in different years. Because of the lack or deficiencies of most of these data, best estimates and best expert (subjective) judgments have been used instead.

Using the historical daily weather data from past five to ten years, for two different scenarios including the "potential" (ideal conditions with no water limitations) and "water-limited" yields of each crop have been derived by running the model for all SUs. Results have been aggregated to the required administration units.

In the next step, correlations between different indicators as the outputs of the forecasting model in two scenarios and the actual historical productions at the same years have been investigated. These indicators included potential biomass, potential yield storage, water limited biomass, water limited yield storage, potential leaf area index, water limited leaf area index, development stage of crop, soil moisture, water consumption and water requirement for each SU. The forecasting indicator showing stronger correlations has been selected and used for crop yield estimation in varying administration units. Aggregated data of crop area and crop yields for similar administration units have been then used to compute the crop yields in different administration units.

15.7 Sampling and survey methodologies adopted for socio-economic variables

Socio-economic variables are not sampled and they are mainly based on the complete census data with update frequencies of 10 years. Therefore, validity of these data would depend on the gaps between the collection and use times of the data, with the least reliable results being a year before the new complete census time.

15.8 Sampling and survey methodologies adopted for agri-environmental parameters

As with the socio-economic variables, the agri-environmental parameters have not been surveyed but have been extracted from the existing data sources such as the land use, soil maps and measured data in meteorological stations.

15.9 Use of geospatial information at the design level

Geospatial data – particularly, the existing land cover/use maps, aerial photos and satellite data (Landsat and IRS) were used for land cover mapping and stratification of the area (Hamedan Province). These data together with data of topography, watersheds and data of the natural and physical boundaries were also used for definition of PSUs, SSUs and sampling segments.

15.10 Use of geospatial information at the estimator level

Aerial photos, maps and remotely sensed data together with the GPS were used for precise location of segments. Remotely sensed data were then used for regression modelling of relations between the estimated and measured crop areas on the satellite data in different strata and administration levels.

15.11 Use of geospatial information for yield estimation

As previously mentioned, yield estimates were mainly based on the availability of detailed agri-environmental variables which were not available and remotely sensed data were not effectively used for yield estimation.

15.12 Various kinds of administrative data used for the area frame test's design

Similar to report in Part 1, the use of the administrative data[7] has been limited to the maps of land use/cover, topography and political boundaries in varying levels (mainly province, townships and villages).

15.13 Quality analysis of data collected through the area frame test

Due to use of area-based data rather than interviews, risks of sampling errors are considerably higher than those of the list frames.

However, main sources of sampling and non-sampling errors may be named as follows:

- errors in data used for the stratification (land use maps, topography, aerial photos and satellite images);
- the methodology employed for the stratification and definition of sample size for different strata;
- geometric errors in defining the segment boundaries during the field works;
- uncertainties related to the imagery and ancillary data used for the stratification.

15.14 Difficulties faced in the area frame test's design and implementation

Because of the high speed of change of the land cover/use in Iran, acquisition of up-to-date data is of primary importance in both the design and implementation stages of land cover related surveying.

Unfortunately, recent aerial photos of the study area have not been available and the existing employed land use maps and aerial photos which are more than 10 years old have been used.

For these reasons, use of such an old data can lead to serious errors in stratification, matching of different data sources and proper location and sampling of the segments. However, difficulties caused by the positioning errors and access to the target sampling units have been highlighted as the most important problems during the implementation.

[7] Including register data

16. GAPS AND LIMITS OF THE AREA FRAME TEST

16.1 Improper choice of the crop forecasting algorithm

The selected crop forecasting algorithm requires spatially detailed data about soil, weather, crop physiology, crop management and historical production data. Most of these are not available in Iran. Therefore, employment of the "best guess" by experts and also default values for model run, have led to unreliable and non-realistic predictions of the crop yields.

16.2 Limitation of the study to a single province

Selected case studies for area frame tests have been limited to Hamedan Province. Because of the high variability of the climate, topography, agricultural systems and other ecological and socio-economic conditions of Iran, a single province does not well represents the country. Therefore, results of the test may be of limited use and they may not be so applicable to the other areas.

16.3 Use of old aerial photography and low resolution satellite data

Availability of suitable data sources plays an important role in the success of many field sampling activities. Because of the heterogeneity and high dynamics of the agricultural units, recent aerial photography and very high-resolution satellite data with spatial resolutions of 1 m to 5 m, are needed for precise selection, location and survey of the area frame samples. The available and unavoidably employed data showed significant deficiencies. As an example, the employed land use maps and satellite data do not meet the required resolution, and display age and scale differences. Yet the land use maps and aerial photos that were used were more than 10 years old and availability of the high resolution and up to data satellite data (e.g. IRS–Pan) was limited to small areas.

16.4 Limitations caused by organization and human resources

Because of the lack of previous experience, limitations caused by the experience and skills of the human resources, such as the experience and skills in agricultural area frame sampling are more serious than that of the current methods.

16.5 Limited use of the geospatial data potentials

Although remotely sensed data have been used in different stages of the sampling process, there are considerable unused potentials of these data. Careful selection and processing of the geospatial data, can lead to further improvements in the design and implementation processes. As an example, consideration of the spatial structure and autocorrelations in different locations and strata could lead to more effective sampling with reduced costs. Also, by effective use of the remotely sensed and GIS capabilities, significant improvements in the stratification, definition and allocation of sample and segment size, planning and collection of field data and database management processes could be achieved.

REFERENCES

Bouzaffour, S., Sharifi, A. 2006. *Multiple Frames for Agricultural Survey, Mission Report (Agricultural Improvement Program).* Bureau of Statistics and Information Technology, Iranian Ministry of Agriculture, Iran.

Cochran, W.G. 1977. *Surveying Techniques. 3rd ed.* New York, Wiley.

Cotter, J.J., Tomczak, C.M. 1994. *An Image Analysis System to Develop Area Sampling Frames for Agricultural Surveys. Photogrammetric Engineering & Remote Sensing,* 60, 299–306.

Dalenius, T., Hodges, J.L.Jr. 1959. *Minimum variance stratification.* JASA, 54, 88–101.

FAO. 1996. *Current Surveys Based on Area and List Sampling Methods, Multiple Frame Agricultural Survey.* FAO Statistical Development Series, Volume 1.
www.fao.org/3/ca6106en/CA6106EN.pdf

Madana, M.H.B.P.H. 2002. *Improving Land Use Survey Method using High Resolution Satellite Imagery,* ITC.
webapps.itc.utwente.nl/librarywww/papers/msc_2002/nrm/madana.pdf

Sharifi, M.A. 2000. *Crop Inventory and Production Forecasting Using Remote Sensing and Agro-meteorological Methods: The Case of Major ,Agricultural Commodities in Hamedan Province, Iran, International Archives of Photogrammetry and Remote Sensing,* Volume XXXIII, Part B7, 1364–1372.
www.semanticscholar.org/paper/CROP-INVENTORY-AND-PRODUCTION-FORECASTING-USING-AND-Sharifi/4dfaa2c1e22906fb9a9174096667192becc605d0

FOOD AND AGRICULTURE ORGANIZATION
OF THE UNITED NATIONS

www.fao.org/in-action/improve-agriculture-monitoring-systems-for-iran

ISBN 978-92-5-131965-9

9 789251 319659

CA7020EN/1/11.19