CROP INVENTORY AND PRODUCTION FORECASTING USING REMOTE SENSING AND AGROMETOROLOGICAL MODELS: THE CASE OF MAJOR AGRICULTURAL COMMODITIES IN HAMADAN PROVINCE, IRAN

M. A. SHARIFI
International Institute for Aerospace Survey and Earth Sciences
Social Sciences Division
Enschede, The Netherlands
Alisharifi@itc.nl

Working Group VII/2

KEY WORDS: Crop inventory, Remote sensing, Agrometorological models, Production forecasting, Monitoring

ABSTRACT

This paper summarises the result and findings of a research project, that was aiming at design and development of an information system which supports crop forecasting of major agricultural commodities in Iran. In this system, crop area estimates are generated through area frame sampling and processing of satellite data, and yields are derived through the application of crop growth simulation model and statistical techniques. The applicability of some of the presents techniques in the context of developing countries are commented. The components that could be operationally applied as well as the major bottlenecks and difficulties that are limiting the application of remote sensing in mapping of the land cover, and crop growth simulation for yield estimation, especially in the context of crop forecasting are identified.

1. INTRODUCTION

Recent developments in aerospace survey technology, digital image processing, modeling of crop production process, and geographic information systems has created promising opportunities for upgrading the agriculture statistical systems. The technology in the related disciplines is well developed, many experimental project have demonstrated the efficiency and effectiveness of crop inventory programs that are based on the application of remotely sensed data. For yield assessment, several agrometerological models have been developed and proved their potential in simulating the behavior of various agricultural production systems. Geographic information systems have created a great potential to bring various forms of information from many different sources together and relate them through a common spatial basis. All these techniques are available and have worked well, when applied individually in many pilot and some operational projects. In a few crop inventory and crop monitoring programs considerable efforts have been put on integration of all these elements in a system. This paper reports on some aspects of another integration effort that tries to design and develop an information system which supports crop forecasting of major agricultural commodities in Iran.

In this system, crop area estimates are generated through area frame sampling and processing of satellite data, and yields are derived through the application of crop growth simulation model and statistical techniques. All the components are integrated into an information system to support timely generation of reliable information on the area and production of the major agricultural commodity in a pilot area. This study has been carried out in the framework of a research program between ITC and the Iranian Ministry of Agriculture and has been carried out during 1995-1999, by ITC and the Agricultural Statistics and Information Department "ASID" of the Iranian Ministry of Agriculture (Sharifi, et., al 1999).

This paper summarises the result and findings of the project in terms of operationality of the existing methodologies for area estimation and yield forecasting in general, and for developing countries in particular. It presents the components that could be operationally applied as well as major bottlenecks and difficulties that are limiting the application of remote sensing in mapping of the land cover, and crop growth simulation for yield estimation, especially in the context of crop forecasting.

2. OBJECTIVES AND DEVELOPMENT APPROACH

The main objective of this project was to develop an integrated method and procedure that can be used by ASID to derive reliable, cost-effective, timely, and repeatable information on agricultural production of the major commodities (wheat, barely and potato) at regional level prior to the harvesting date. This method was developed for Hamadan province with the intention of extension to other provinces and finally the entire Iranian territory. On this basis the main objectives of the project was formulated as follows:

- **Development of crop inventory method and procedures:** To distinguish, identify, measure and map the area of crops of significant importance in the study area (crop inventory).
- **Development of Yield forecasting method:** To derive a reliable estimate of yields (production per hectare) of the major agricultural commodities that are growing in the study area.
- **Development of crop forecasting method:** To derive a reliable estimate on the production of major agricultural commodities prior to the harvesting date of each crop.
- **Development of a conceptual model for the crop forecasting system:** To integrate all the required processes and their respective thematic and spatial data in an operational method.

Development approach: Development of a crop forecasting system for a large area such as this is a complex process which includes development, calibration, validation and experimentation with a number of models of different principles. Development of these models involves extensive development, realisation, experimentation and evaluation activities which can only be controlled when it is concentrated on a smaller area. In the context of decision science, this can be considered as a decision problem, in which the decision maker (the Project team in this case), would like to consider the state of the arts technology and develop and choose a method and procedure which is best suited to the requested local conditions. To carry out this task a systematic approach based on the bounded rationality (sharifi, 1999) the following steps was defined and implemented:

- Design an operational crop forecasting method based on existing technology and capacity of the involved organisations (preliminary study of state of the arts in the related technologies as well as understanding the capacity of the organizations involved in Iran)
- Apply the method in a pilot area (one district)
- Evaluate and modify the original method (discover the shortcoming of the first concept and modify them accordingly to develop the improved method)
- Apply the improved method in the provincial level
- Evaluate and, modify the method (discover the shortcoming of the improved method and modify them accordingly to develop the final method)
- Document the final method

3. OVERALL SYSTEM DESIGN

The crop forecasting system as designed here is composed of three main processes: area estimation, yield estimation, and production forecasting (Figure 1).

Process 1. Area estimation (Crop inventory): This process makes use of different techniques and data which are coming from variety of sources such as remote sensing, field observation and historical data to derive area estimates of the major agricultural commodities.

Process 2. Yield forecast: This process makes use of crop growth simulation models and detailed data on soil, weather, crop physiology, crop management, and historical production data to derive periodical estimates on yield of various agricultural commodities.

Process 3. Production forecasting: This process makes use of area and yield estimates derived from processes 1. and 2., and calculates the periodical production forecasts for the lowest administration units. This unit is then aggregated to the higher levels of administration.

In the following each of these processes are considered as one subsystem and further detailed.

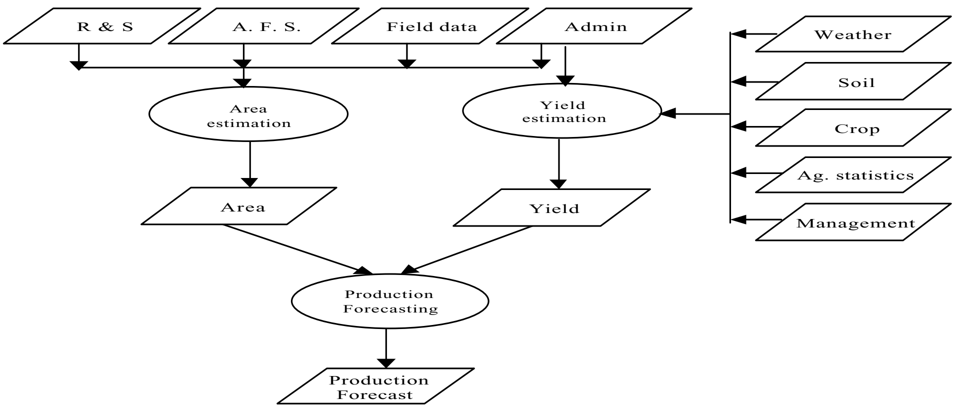

Figure 1. Overall design of the crop forecasting system

3.1: Crop Area Estimation Subsystem

Crop inventory subsystem includes four main processes for identification of crops, estimation of their area and mapping their distributions. These are: stratification, area estimates through area frame sampling, area estimates through remote sensing, area estimates through combination of the two sub-processes and aggregation/ desegregation to different administration levels (Figure 2.). For details see Sharifi, et al., 1999.

Stratification of the region: This process makes use of a number of data sets such as current land use map, soil map topographic map and Landsat, and Spot satellite data to stratify the region into homogenous areas in term of use-type and pattern of agricultural fields.
Area frame sampling: Is a well established techniques used in agricultural statistics based on sample parcel survey (for detail information see Sharifi & Abkar, 1999; Gallego 1995; SRS 20, 1975). This technique has proven itself in terms of reliability and accuracy. However, to achieve a reliable estimate it requires rather large number of samples and field measurements every year. As a result, costs are high and the procedure is error prone, and the product is an estimate of the total area of each crop per stratum without providing their distributions as a map. The application of remote sensing techniques (images) can provide a total coverage of the area and reduces the total number of segments to be surveyed while keeping the accuracy and reliability of the estimates.

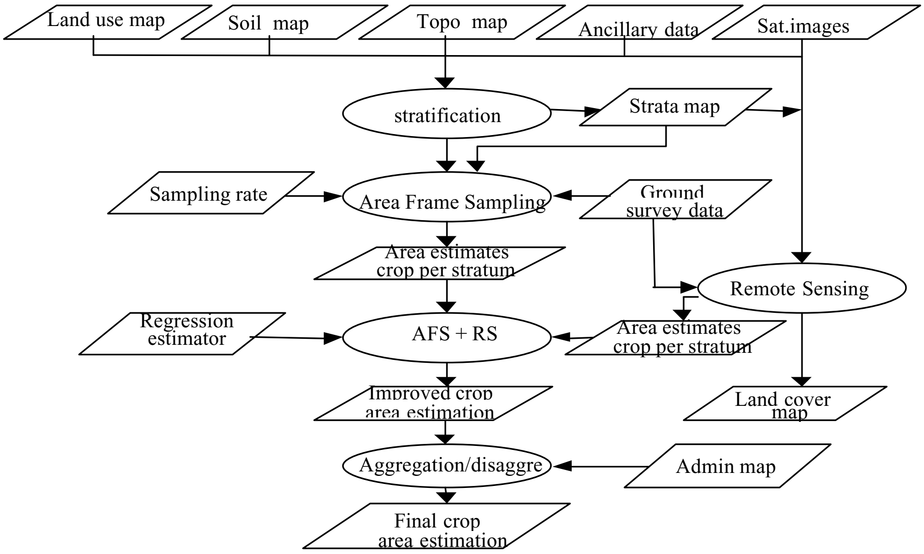

Figure 2. Top level design of Crop Area Estimation subsystem

Area estimate through remote sensing: One of the operational method of applying remote sensing in agricultural statistics is so called "regression estimator" that has been recommended by JRC as a results of their 5 years experimentation in Europe (Gallego, 1995, MARS, 1994; Gonzalez and Cuvas 1993). The process includes preprocessing of satellite data to remove the radiometric and geometric errors in the data-set and their classification through supervised routines which include training of the classifiers by sample segments. In this method the result from the area frame sampling and the image processing are statistically related and used to derive an improved area estimate per crop in each stratum.
Aggregation/Desegregation to various administration levels: The area estimate derived through the above method is at the stratum level which, physically do not exist. To transform that into the required administration units, the result is overlaid with the lowest administration maps and later on aggregated to the required level.

3.2: Yield Forecasting Subsystem

Yield is a function of biophysical (crop, soil and weather characteristics) and management (various inputs, quality and timeliness of the crop husbandry operations) factors. The climatic factors from the biophysical sets and most of the management factors are changing from field to field and year to year. This makes the yield forecasting very difficult task. Fortunately agro-ecological models have been developed to simulate the behaviour of various agricultural production systems considering biophysical factors and effects of some elements of management such as time and amount of irrigation, fertiliser and seeds. In these models, the relevant crop- environment interactions are described quantitatively in a set of simulation models based on understanding of the constituent processes of the system and their impact on system behaviour. These models when properly calibrated and validated can be used reliably, to estimate the production potential of a crop with clearly defined properties in a well- defined aerial environment (mapping unit). The aerial environment is characterised by the relevant soil and weather characteristics. For details see Sharifi, et. al., (1999). This implies that the simulation models can not predict the actual yield of certain crop, as many factors that are affecting yields are not considered in the simulation. However, it can produce reliable estimates on potential yield considering a number of biophysical and some managerial factors. Following Van der Wall and van Dieppen (1995) we have assumed that there is a strong correlation between the yield potentials and the actual yield of crops in the region. Based on this assumption the yield-forecasting model is formulated as shown in the Figure 3

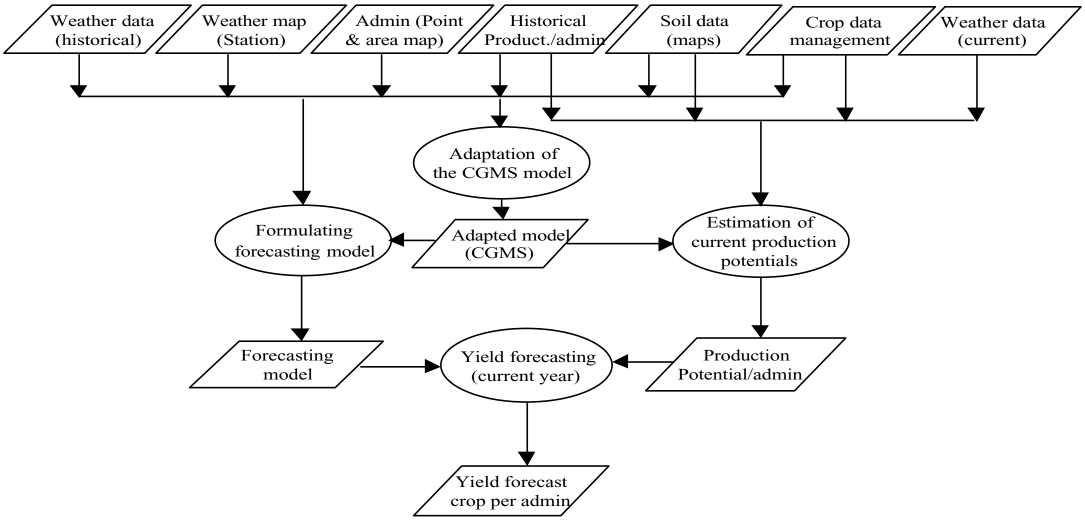

Figure 3. Yield forecasting subsystem

Formulation of forecasting model: In this process, first a crop growth simulation model is calibrated and validated for the major crops grown in the region. In this study the Crop Growth Monitoring System “CGMS” (Perdiago and Supit 1999, Supit, et. Al 1994) which has been developed by Joint Research Center of European Commission has been selected and used. The calibration and validation requires data on crop Phenological/genetical properties, crop management such as crop calendar, amount of seeds as well as the historical daily weather data, physical, and chemical properties of soil and the actual historical yield in different years. Next the model is run for all the Simulation Units “SU”, which are defined as homogenous units in terms of soil and weather characteristics, see Figure 4.

Using the historical daily weather data of past 5-10 years, the potential and water-limited yields of each crop were derived by running the model for all SU’s, and aggregated to the required administration units “Shahrestan”. In the next step the actual historical productions were correlated with the potential yield of the same crops at the same years to formulate the forecasting model. The correlation is established between different indicators of yields derived from simulation and the actual historical yields and select the one which shows stronger correlation as the forecasting model.

Estimation of current production potentials: Once the forecasting model is established, the potential and water limited production of each crop in the current year in different administration units are estimated and fed in to the model. This requires preparation of soil, weather and crop information for each SU which are suitable for production of the designated crop in each administration units. As a first approximation (in the beginning of the season) the long-term average weather data is used to derive the potential productions. This average data are substituted by the actual measured daily weather data as the crop cycle goes on. Naturally the estimates are becoming more accurate as the crop cycle advances.

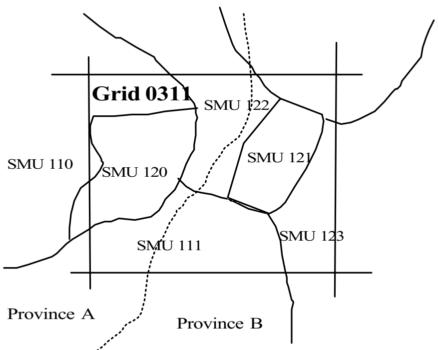

Figure 4: Example of a Simulation Unit

Yield forecasting (current year): Applying the potential and water limited production (or its relevant indicators) of the current year in the forecasting model will produce an estimate for the yield in the current growing period, in the specified administration unit.

3.3: Crop Production/Forecasting Subsystem

Crop production is the product of the two independent factors of crop area and yield estimates. The main inputs are area estimate and corresponding yield forecast of each crop per administration unit derived through subsystems 1 and 2. The main processes here are used to establish linkage between different results of the above subsystems and aggregation of the production to the higher level of administration units. Further, for better presentation and communication of the results to the decision makers it provides facilities to combine land cover or any types of maps which can be produced in the system with the final or any of the intermediate results. This subsystem is graphically shown on Figure 5.

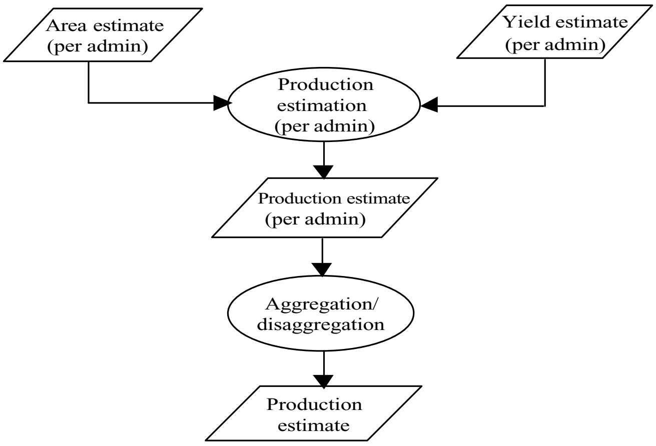

Figure 5 Production Forecasting Subsystem

4. Experimentation results

4.1 Study Area:

The study area was the province of Hamadan, Islamic Republic of Iran. It is located in the Western part of the country and covers 1,902,500 hectares, of which 562,000 ha. is agricultural land. The area consists of high (3500 m.) mountains, hilly areas and plains. In the plains irrigated agriculture is predominant, though there are substantial salinised areas. In the hilly areas dry farming (mainly wheat and barley) is practised and also grassland and natural vegetation is present. The mountains are mainly covered by natural vegetation and grassland with (small) valleys where agriculture is practised (mostly irrigated). Wheat and barley which are covering over 74 % of the area are considered as the major agricultural commodities in the region. Since potato strategically is a very important crop based on the request of the client was included in the project. As a result yield estimation and crop forecasting was basically covering wheat, barley and potato. In this paper, the inventory results of the whole province for wheat and the yield performance results of Razan District will be presented. Razan district is located at the northern part of Hamadan province (34^o-35^o 45´N, 46^o 45´-49^o 30´E) of Iran, covers about 134000 hectares of which 50% is agricultural land. The study area is a shahrestan (a district), which is subdivided into seven Dehstans (sub-district) and contains 157 villages (Sharifi, et al 1999). The area consists of high (3500m) mountains, hilly areas and plains and has a semiarid climate with mild summers and very cold winters. The mean annual rainfall ranges between 320 and 350mm, and mean monthly temperature varies from -5^oC in January to 24^oC in July. The soils of Hamadan province are predominantly of clayey texture and the litology is determined by calcareous schist, limestone. In the plains irrigated agriculture is predominant, though there are substantial salinised areas (Sharifi et. al., 1999). The major crops in the area are wheat, barley, alfalfa, potato and beans. Wheat and barley that cover 80% of the area are considered as the major crops in the region.

4.2 Area estimate

The area frame sampling included 181 field sample segments of 500*500 and 700*700 meters, which corresponds to an overall sampling rate of 0.4 percent for the province. Samples were distributed over different strata based on their probabilities. The samples distribution together with some example of them is given in Figures 6 and 7. The summary result of the inventory for irrigated wheat , rainfed wheat, the total agricultural land as well as the total man-month efforts which was put for data collection and calculation in comparison with the traditional statistical method is presented in Table-1. The field work was carried out within a period of 2 months, and calculation of the estimated areas took about 1.5 months, which can be easily reduced, in the next exercises. This meant that the estimates became available during the second half of October, which is before the harvest of some crops but after the harvest of grains. The system is capable of estimating the areas of grains before harvest, but only if either the summer crops are discarded or the segments are visited twice. At the time the AFS results became available, the traditional statistics had not been published yet. From the table-1, it can be concluded that the proposed method can produce the same types of with considerable less amounts of efforts in a more timely fashion, even though the data which is used for comparison, is related to the first exercise which normally takes more time, cost more and produce less reliable results. Naturally by gaining more experience in field measurement and calculation the efficiency and effectivity of the method will increase.

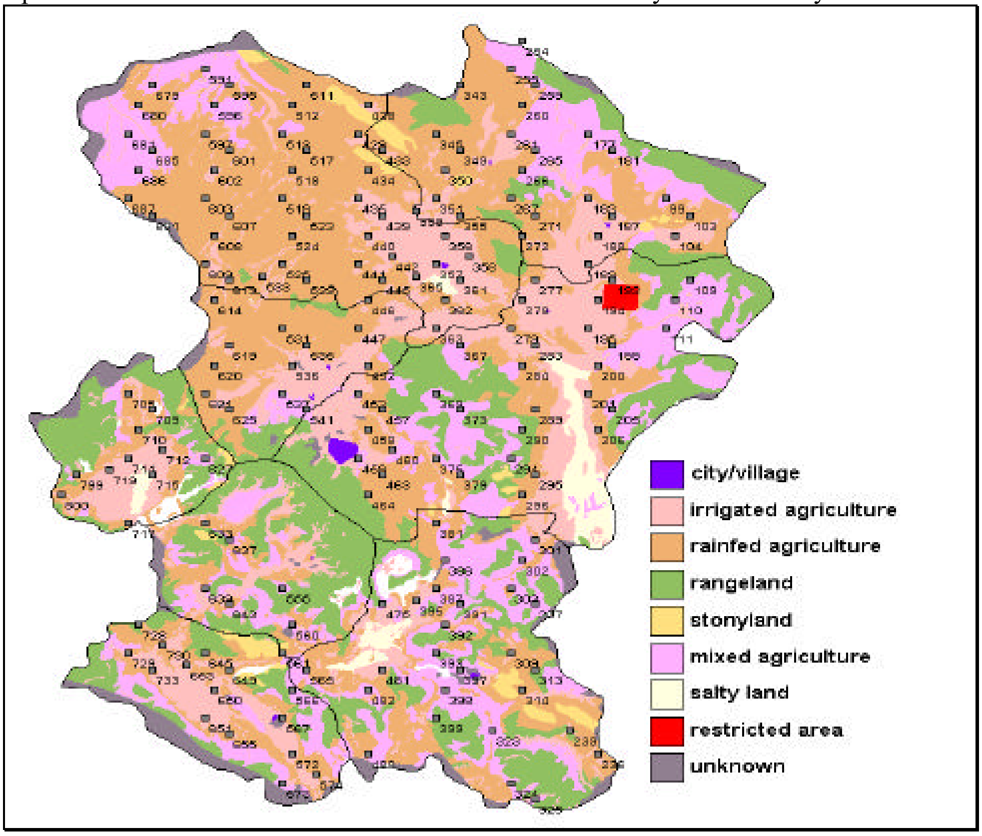

Figure 6. Sample segments distributions over stratified land use map of province

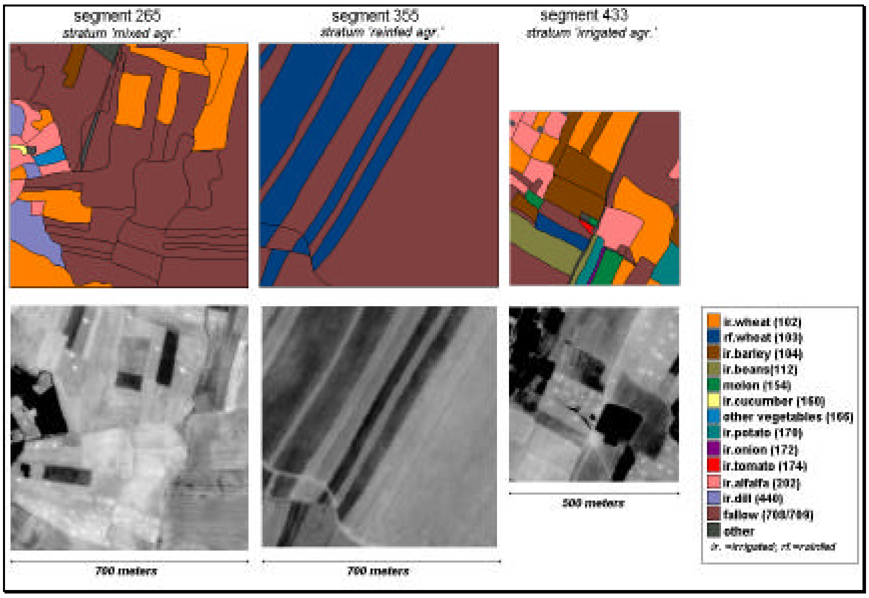

Figure 7. examples of sample segments and their identified fields

year	Wheat-irrigated	Wheat - rainfed	Total wheat	Total agriculture	Total man-month
1997-traditional survey	109,724	297,301	407,025	588,879	76
1997- area fram.samp	126,192	239,504	365,696	564,247	52
Difference in Ha	16,486	57,797	41,329	26,432	24
Difference in %	15	19	10	4	31

Table 1. Summary of the inventory result for wheat in Hamadan Province

4.3 Yield performance results

The simulations were executed for Navid wheat variety for each Simulation Unit. Simulation unit were defined as a combination of a crop number (which determines the crop parameters), a grid number (which determines the weather data) and a combination of the soil parameters, Soil Physical Group (SPG) and Rooting Depth (RD). The SPG together with the RD determined the necessary soil parameters for a simulation run. This implies that all soils that are suitable for a crop having the same SPG/RD combinations will have the same yield for the same crop within the same grid.

Dry matter accumulation and its distribution of leaves, stems and grains are simulated from sowing to maturity on the basis of physiological processes as determined by the crop's response to daily weather (water availability, solar radiation, maximum and minimum temperature and air humidity), soil moisture status and management practices. The soil water balance was calculated without ground water influence and was driven only by rainfall, evapotranspiration and possible surface storage. Calculation included the processes of infiltration, soil water retention, percolation and loss of water beyond the maximum root zone (Supit et.al., 1994). The calculation of yield estimation produced results which 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 simulated results were checked before the aggregation to find whether any unsuitable unit was included in the simulation or not. Then the weighted average of the dry matter and grain weight (yield) was calculated. The potential and water limited yields (kg/ha) were aggregated separately. A criterion has been fixed for suitability check. The SU, which produced potential yield, lower than 10 kg/ha for every year were excluded in calculating weighted average. These SUs were considered as the unsuitable for growing the irrigated Navid variety. The same procedure has been followed for the aggregation of the water limited simulated results. The simulated aggregated yields and the aggregated official historical yields from 1988 to 1996 were used to construct yield forecast model. Therefore, the SU based simulation results were aggregated to the shahrestan (district) level. The actual yields (kg/ha) of irrigated and rainfed wheat for the corresponding years was separately aggregated also at the shahrestan level. These data were used to establish the relationship between the actual yields and simulated yields. The simulated yields for 1996 was used for the prediction of the actual yields with the developed yield models. These data were also aggregated on the basis of the threshold values. The SU, which produced potential yield higher than 500 kg/ha and water limited yield higher than 100 kg/ha only included in the aggregation. Then the weighted average was calculated for dry matter and yield for both the potential and water limited situations.

A rising trend yields (kg/ha) of the irrigated and rainfed wheat of the study area was observed for the period of 1988 to 1996 (Figure 8). This rising trend is the result of improved farming practices like the introduction of the new wheat varieties, higher application rates of fertilizers and more intensive control of weeds, pests and diseases (Hooijer and van der Wall, 1994). A smooth trend of any type over a large number of years assumes a continuity which might be unrealistic (de Koning et.al., 1993). According to these authors, the predictor should only be based on data from recent past. The length of the series should nevertheless be long enough to give a sufficient number of degree of freedom in the regression analysis (Supit, 1997). The yield models are developed for the period of 1988 to 1995 and the established regression coefficients are subsequently used for the prediction of yield of 1996. Considering the trend, and regression between simulated potential and water-limited grain, total biomass and the actual yield of irrigated and rainfed wheat in the district a number model were developed and tested. Models with only simulation indicators as the only predictor did not performed well, however models witch included trends are rather robust and performing well (correlation coefficient over R^2=0.8). Better prediction accuracy can be obtained using the developed models for the irrigated wheat. But the models for the rainfed wheat appeared to be less robust. Table 2 shows the results of the selected model and their performances for irrigated and rainfed yield in Razan district during 1996 growing period.

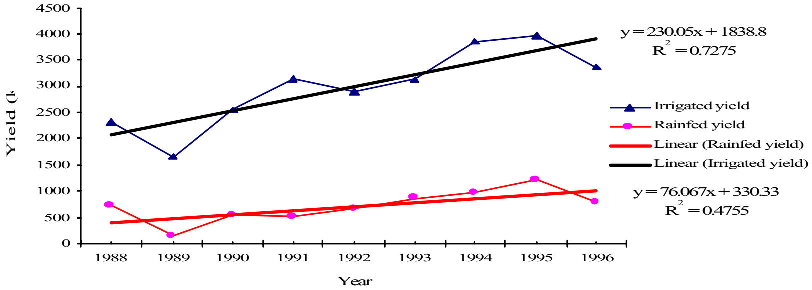

Figure 8 Actual yield of the irrigated and rainfed wheat in Razan District (1988-1996)

Crop	Predictor	Correlation value (R^2)	Predicted yield (kg/ha)	Actual yield (kg/ha)	Prediction error (%)
Irrigated wheat	Water limited yield	0.91	2596	3372	-23.0
	Water limited yield + trend yield	0.94	3162	3372	-6.1
Rainfed wheat	Potential yield + potential biomass + trend yield	0.86	712	775	-8.1

Table 2: Comparison between the predicted and actual yields for irrigated and rainfed wheat, Razan - 1996

5. Concluding remark

The proposed methodology is based on the current state of the art and two years of experimentation at different levels (district and provincial level). Functionally, the system includes three main components mainly, crop area estimation, and yield estimate and production forecasting. The overall evaluations of each function/sub-systems are as follows:

Crop area estimation is making use of area frame sampling and remote sensing techniques. The area frame sampling performed quite well in all aspects e.g., it cost less, it is more timely and produces a better results. As it is based on measurements rather than interviews; technically it was also well understood and could be easily operationalized at different extent. Although its operationalization at higher level requires more training for field staff. More training produces more reliable and timely information. Availability of recent aerial photography or very high-resolution satellite data (1-5 meter resolution) was considered as a limiting factor for the expansion of area frame sampling. With the existence of a current plan to cover the whole Iranian territory with 1/40,000 black and white aerial photography, which already covers a good portion of the country (including Hamadan province), and recent development in the satellite data collection, this problem if not resolved (IRS-D pan with 6 meters resolution already in space and existing Russian satellite data with 5 meter resolution) will be soon resolved by the new series of satellite which collect world wide high resolution data in the order of one meter resolution, e.g., IKONOS-1 satellite. The full application of remote sensing technique (digital image processing of current satellite data) although it was technically feasible, it was not practically applicable. This was due to the difficulty of obtaining the timely receiving and delivery of satellite data as well as technical difficulty in its processing. For detailed analysis of the existing bottlenecks of crop inventory based on remote-sensing see Sharifi & Abkar 1999. If the operational problems are removed in the future, remote sensing can help to increase the accuracy as well as reducing the amount of field work and costs. If an information system is developed to formalise and support all the activities of the area frame sampling certainly it can further reduce the time and costs of operation.

The yield estimate is making use of crop growth simulation models which make use of an extensive data sets on crop phonology, physiology, soil, weather and management practices. In the absence of the real data for many of the required parameters a reasonable estimates as a default value is applied in the simulation procedures. As time goes by and new data set becomes available the default values should be replaced by the actual values to improve the quality of the results. The yield estimation/forecasting is based on the relationship between the simulated yields/indicators of crop in different year and their corresponding actual historical yields. Since in this process the simulated yield is used as an indicator of the crop performance in each year, changes in their absolute values will not affect the final yield estimate, therefore very exact calibration and validation of the respective models were not absolutely necessary. The yield model for wheat, which includes trend yield is quite robust and produces a good results. If historical data can be made available, it can be extended to the other area and applications. Due to its information requirement it can serve as a comprehensive framework for collection, and organisation of large data sets which can be used for many other applications. It will also set a framework for new data collection, research, and development. As the area will not change dramatically from year to year, yield alone can be used to derive a good outlook of the agricultural production at any time in the growing period.

Derivation of a reliable, timely and cost-effective estimates on crop area and their related productions does not concern a simple transfer of an established operational method or techniques especially for yield forecasting and image processing portion. Such methods and techniques include many advanced concepts, which require further understanding and development. This can only be achieved through further research and experimentation in crop growth simulation, information system development, and image processing, to allow further understanding of the processes, gaining the required knowledge and completion of the required data sets.

ACKNOWLEDGEMENTS

The author would like to acknowledge the valuable contributions of the project team at ITC, Wageningen and the counterpart team in the Department of Agricultural Statistics and Information in the Iranian Ministry of Agriculture, both in Tehran, and Hamadan Province.

REFERENCES

de Konning, G. H. J., Jansen, M. J. W., Boons-Prins, E. R., van Diepen, C. A., Penning de Vries, F. W. T., 1993: Crop Growth simulation and Statistical Validation for Regional Yield Forecasting Across The European Community. Simulation Reports CABO – TT, No. 31, AB-DLO, Wageningen, The Netherlands

Gallego, F. J., 1995. Sampling Farme of Square Segments. JRC Publication.
MARS 1994. Proceedings of the conference on The MARS Project: Overview and Prospective, 17 18 November 1993, Italy.

Gonzalez, F., and Cuevas, J. M., 1993. Crop area estimation by regression estimators and confusion matrices, In Proceeding of International Symposium Operationalization of Remote Sensing, Vol 2. 19-23 April, Enschede the Netherlands

Perdiago V., Supit I., 1999. An Early crop warning estimation method for Finish conditions: the crop grrowth monitoring system of the joint research center with and without remotely sensed and other additional input data . JRC publication EUR 18975 EN.

Sharif M. A., Abkar, A., 1999. Crop inventory through remote sensing: problems of operationalization. Paper presented to the second International Symposium on operationalization of remote sensing, 16-20 August 1999, Enschede

Sharifi, M. A.., Driessen, P., Van Keulen H., Bronsveld, M. C., Clavaux, M. B. W, and Lenaard, J., 1999. Development of crop inventory and forecasting system for the major agricultural commodities in Hamadan Province Islamic Republic of Iran. Final Report, on the contract research for the Agricultural statistics and Information Department, of the Iranian Ministry of Agriculture

SRS 20, 1975. Area Frame Sampling in Agriculture. Statistical Reporting Service. U.S. Department of Agriculture.

Supit, I. 1997: Predicting national Wheat Yields using Crop Simulation and Trend Models. Agricultural and Forest Meteorology 88, 199- 214

Supit I., Hooijer A.A., van Diepen C.A.., 1994. System Description of the WOFOST 6.0 crop simulation model implemented in CGMS. JRC Publication EUR 15956 EN.

Wal van der T., and van Diepen, C. A., 1995. Crop Growth Monitoring System CGMS. paper presented in “ From Soil Survey to Sustainable Farming” the Conference of 35th Anniversary of the soil Fertility Research Institute Bratislava, 3-5 October 1995, Bratislava.