IMPROVED REGIONAL YIELD PREDICTION BY CROP GROWTH MONITORING SYSTEM USING REMOTE SENSING DERIVED CROP PHENOLOGY

V.K. Sehgal¹, D.R. Rajak, K.N. Chaudhary and V.K. Dadhwal⁶

Crop Inventory and Modelling Division, Agricultural Resources Group, Space Applications Centre (ISRO), Ahmedabad 380015 INDIA – (vksehgal@yahoo.com)

⁶Commission VII, WG VII/6

KEY WORDS: Crop Simulation Model, Growth Profile, Crop Phenology, Wheat, GIS, Remote Sensing, Yield Prediction

ABSTRACT:

Dynamic process-based crop simulation models are useful tool in predicting crop growth and yield in response to environmental and cultural factors but are constrained by lack of availability of the required large number of inputs when applied for regional studies. In this study we report (a) development of a prototype Crop Growth Monitoring System (CGMS) for wheat using WTGROWS simulation model on a 5’X5’ grid in GIS environment for generating daily crop growth maps and predicting district-wise grain yield, (b) demonstration of a technique for estimating date of sowing (DOS) using RS-derived spectral-temporal crop growth profiles and CGMS simulation capability and (c) evaluation of the capability of CGMS for spatial yield mapping and district level yield prediction for Haryana State during 2000-01 crop season. The technique for estimating district-wise DOS matched the RS-derived date of peak NDVI (from multi-date WiFS sensor aboard IRS-1D satellite) to date of peak LAI simulated in CGMS for a range of plausible dates of sowing. The peak date of NDVI was computed by fitting Badhwar model to the multi-date NDVI values. The CGMS performance was evaluated by incorporating RS-derived date of sowing in predicting district level wheat yields with and without use of district-wise N fertilizer application rate computed from district-wise fertilizer consumption statistics. The correlation between district yield simulated by CGMS and official State Department of Agriculture (SDA) estimates was only 0.163 when constant median/mean inputs of DOS, N fertilizer and irrigation application were specified for all the districts. The correlation increased to 0.52 when RS-CGMS-derived district-wise DOS was used as input and further increased to 0.74 when information from consumption statistics of N fertilizer use was additionally specified. Thus, the study demonstrated the derivation of crop sowing dates from RS inputs using crop simulation based CGMS and its incorporation into CGMS framework along with spatial variation in other crop management inputs for improving yield prediction.

1. INTRODUCTION

Monitoring of crop growth and yield assessment at regional scales need assimilation of crop environment information from various sources spatially and a crop growth monitoring system (CGMS), linked to GIS, provide such a framework. At the core of CGMS is a process based crop simulation model which simulates the crop growth and development at regular time steps in response to the changes in crop environment. The simulation models generally run at point scale and require detailed inputs on weather, soil, crop management etc. The CGMS permits spatial analysis of crop growth by running it at a regular geographical grid. The grid-wise simulation model inputs are generated from spatial layers of weather variables, soil properties, crop distribution, cultivar parameters and crop management practices in a GIS environment. Depending upon the study objectives and spatial resolution of inputs, the CGMS framework can be applied to provide qualitative and quantitative crop status spatially at regular time interval from a field to nation scale. Thus, CGMS is a powerful tool for agricultural managers and planners to visualize the complex interaction between crop and various environmental and management factors spatially and temporally to plan better crop management strategies and predict crop yield at different scales. A CGMS based on WOFOST model is used for operational yield forecasting of important crops grown in 12 member states of European Union (Meyer-Roux and Vossen, 1994). Satya Priya et al. (1998) have also demonstrated a CGMS based on EPIC model working at two different grid resolutions of 50 km and 10 km for predicting maize, wheat and rice yields in India at national and regional scales.

One of the major constraints of implementing CGMS for crop yield prediction is the non-availability of information on spatial variability in crop management practices adopted by farmers in a region. The dominant yield determining crop management practices are: (i) date of sowing (Saini et al, 1988), (ii) amount, type and method of N fertilizer application and, (iii) number and amount of irrigation (Sehgal, 2001), besides, weeding and cultivar type. In a given region and a season, these factors result in large field-to-field yield variability. The multi-date satellite data offers the opportunity to identify crop in a region and monitor its phenological development (Dadhwal et al., 2002; Rajak et al., 2002). For the irrigated spring wheat grown in northern India, variations in date of sowing within a season as well as variations in the rate of nitrogen fertilizer application are the two dominant factors causing field-to-field variations in crop growth and yield in a region.

For this study, (a) a prototype Crop Growth Monitoring System (CGMS) for wheat was developed which is based on WTGROWS simulation model (Aggarwal et al., 1994), (b) a technique for estimating dates of sowing using multi-date remote sensing (RS) derived (IRS-WIFS) spectral growth profile within a CGMS framework was demonstrated, and (c) CGMS performance was evaluated by incorporating RS-derived date of sowing in predicting district level wheat yields with and
without use of district-wise N fertilizer application rate computed from district-wise fertilizer consumption statistics. This study was carried out for northern Indian state of Haryana for the crop season 2000-01.

2. METHODOLOGY

2.1 CGMS Prototype

The CGMS developed for this study consisted of four components, namely, (a) inputs assimilated in GIS, (b) a relation database management system (RDBMS), (c) a two-way linking shell between RDBMS and crop model and (d) crop simulation model WTGROWS. The framework has been implemented on MS Windows NT™ platform on a personal computer. The MS ACCESS™ software has been used as RDBMS while AGROMA™ IP/GIS software has been used for image processing and GIS functions. All the spatial layers for the study area were geo-referenced in UTM projection Zone 43 North with Indian Datum. The generation of input spatial layers and CGMS sub-system functions are described below.

2.1.1 Grid Layer: A 5°X5° polygon vector grid layer was generated for the state of Haryana in GIS. Each grid cell represents one simulation model run and hence all the other inputs were assimilated/aggregated at grid cell level in the RDBMS though they were having information at different spatial scales. The serial number i.e. identifier, area and central latitude of each grid cell was generated and stored in RDBMS.

2.1.2 Administrative Boundary Layer: Vector layer of district boundaries were digitized from 1:250,000 scale Survey of India (SOI) maps. The district boundary layer was overlaid on grid layer and each grid was assigned a district code depending on the maximum district area in the grid.

2.1.3 Soil Properties Layers: The soil resource map of Haryana produced by NBSS&LUP (Sachdev et al., 1995) at 1:250,000 scale was digitized. Soil depth and soil texture layers were generated after reclassifying the soil-mapping units according to their attributes of depth and particle size, respectively. Each grid cell was assigned average soil depth and dominant soil texture class. To account for soil fertility, soil organic carbon raster map was produced by interpolating from 72 point data collected from literature using inverse square distance interpolation. An average soil organic carbon content was calculated for each grid cell and stored in RDBMS.

2.1.4 Weather Surfaces: The daily weather data (Rainfall, Maximum and minimum temperature, Wind speed, and Relative Humidity) of 21 surface observatories in and around Haryana State were entered as table in RDBMS. A weather data interpolation program was written in “Visual Basic™”. The program read daily weather data of observatories with their locations from the database tables and generated daily surface of each weather parameter at 5°X5° resolution using inverse square distance interpolation. Boring through the daily weather surfaces resulted in grid-wise daily weather data file in the format of WTGROWS. Due to the non-availability of daily solar radiation for most of the observatories, Hergreve’s method of estimating daily solar radiation from temperature range was adopted. Nain and Dadhwal (2001) have derived coefficients of Hergreve’s equation for various stations in wheat belt of India. Using the geographical coordinates of such stations in and around Haryana State, a thiessen polygon surface was generated and each grid cell was assigned dominant polygon’s Hergreve’s coefficients.

2.1.5 Crop Model: The WTGROWS (Aggarwal et al., 1994) is a detailed production level-3 mechanistic model, which simulate the potential production, phenology, soil water balance, soil and plant nitrogen balance and water and nitrogen stress on plant growth and development. It has limitation that it does not simulate the effect of biotic stresses (pests and diseases) on crop growth and development. It requires inputs on site data, daily weather data, soil characteristics and crop management data. The model has been well calibrated for Indian wheat cultivars. In this study, the standard values of genetic constants for a semi-dwarf medium duration high yielding wheat cultivar were adopted (Aggarwal et al., 1994). The model, written in PCSMP (Personal Computer Continuous System Modeling Program by IBM, 1975), runs on IBM compatible PC under MS-DOS.

2.1.6 CGMS Shell: For interfacing the spatial inputs generated as grid attribute table to WTGROWS model, the “linking” strategy described by Hartkemp et al. (1999) was adopted. The linking shell was written in C language, which read the grid attribute table and generated the required input parameters for the model for each grid having wheat area. It also copied the daily weather file for each grid as the current weather file. Pedo-transfer functions were incorporated into the shell to generate volumetric soil-water constants from the grid cell textural class. The pedo-transfer function coefficients were generated from the experimental soil dataset of twenty locations in Haryana published by Komos et al. (1979). The organic nitrogen in soil was initialized for each grid cell from organic carbon content by assuming a C:N ratio of 10:1. The shell also initialized soil moisture at sowing as 75 percent of field capacity to simulate a pre-sown irrigation which is common adopted practice in the State. The shell ran the model for each of the grid and model outputs were written back into the grid attribute table in the RDBMS to be visualized as grain yield and biomass maps in GIS. The error trapping was also built into the shell to know if model simulation could not be accomplished for any of the grid cell.

2.2 RS-data Analysis

2.2.1 Wheat distribution Layer: Eleven IRS-WiFS images acquired between 28-Oct-2000 and 22-Apr-2001 were registered, georeferenced and radiometrically normalized. Hierarchical decision rule based classification (Oza et al., 1996) was carried out to discriminate wheat from other categories resulting in wheat distribution map. Fraction of wheat area to total grid area was computed for each grid cell and stored in grid attribute table. This fraction was used as weight in computing weighted average district yield from grid values.

2.2.2 Estimating Wheat Phenology: The technique for estimating wheat phenology including dates of sowing is based on the premise that in a season, date of peak NDVI is very distinctive and corresponds to the date of peak leaf area index (LAI) of the crop for a given set of soil, weather and cultivar type. So, if we iteratively vary only date of sowing and simulate such a LAI profile by crop model whose date of peak LAI matches with the date of peak NDVI, then that is the represented value of date of sowing. Using wheat distribution image and district boundary vector, district-wise mean wheat NDVI was computed for each date. A functional form of
Badhwar model (Equation 1) was fitted to the mean NDVI temporal set to develop district-wise wheat spectral profiles. Badhwar (1980) described wheat growth profile by the following functional form:

\[ G(t) = \begin{cases} G_0 \cdot \exp \left[-\beta \left(t - T_0\right)^2\right] & \text{for } t > T_0 \\ G_0 & \text{for } t \leq T_0 \end{cases} \]  

(1)

where \( G(t) = \) wheat NDVI at time \( t \)
\( G_0 = \) soil NDVI at \( T_0 \)
\( T_0 = \) spectral emergence day
\( \alpha, \beta = \) crop specific constants (\( \alpha > 0 \) and \( \beta > 0 \)).

By analysing the image, a value of 0.1 NDVI was found for \( G_0 \) in this study. Using the profile parameters \( \alpha \) and \( \beta \), the day of peak NDVI, Tmax is estimated as:

\[ \text{Tmax} = \left(\frac{\alpha}{2\beta}\right)^{1/2} \]  

(2)

The profile parameters (Alpha (\( \alpha \)) and Beta (\( \beta \))) and phenology indicators (\( T_0 \) and Tmax) were computed for all the 16 districts. In order to save on the computation time spent in iterative running of CGMS with dates of sowing varying by one day interval, the CGMS was run for five dates of sowing (310, 320, 330, 340 and 350 Julian day) only. A constant N fertilizer application rate of 120 kg/ha applied in three splits and four irrigations each of 60mm to all the districts were specified. For each date of sowing, the district-wise daily average LAI profile was computed from grid simulated daily LAI profile belonging to that district. The two consecutive dates of sowing DOS1 and DOS2 which had a lower date of peak LAI (DLL) and a higher date of peak LAI (DHL) with respect to Tmax were identified. The district-wise date of sowing (DOS) was computed by linear interpolation as given below:

\[ \text{DOS} = \text{DOS1} + (\text{Tmax} - \text{DLL}) \times \frac{\text{DOS2} - \text{DOS1}}{\text{DHL} - \text{DLL}} \]  

(3)

2.3 CGMS Performance Evaluation

The CGMS performance in terms of district level yield prediction was evaluated under three scenarios of specification of management inputs. Scenario-1: a constant date of sowing (330 Julian day), constant fertilizer application rate (120 N kg/ha) and constant number of irrigations (4 each of 60 mm) for all the grids. Scenario-2: district-wise RS-derived date of sowing and constant fertilizer and irrigation application. Scenario-3: district-wise spectral growth profile derived date of sowing, district-wise estimated fertilizer application rate computed from consumption statistics (Anonymous, 2001) and a constant irrigation application. For CGMS evaluation, district level observed wheat yields were obtained from the State Department of Agriculture of Haryana.

3. RESULTS

Haryana was covered in 708 grid cells of 5’X5’ and the area of grid cells varied between 73.54 and 75.94 km². The soil texture classes were dominated by loamy class (338 grid cells), followed by sandy (192 cells) and coarse loamy (156 cells) classes. In the case of average organic Carbon, 493 grid cells had values between 0.2 – 0.4 percent, 162 between 0.4 – 0.6 percent while 24 grid cells had above 0.6 and 29 below 0.2 percent. The classified image generated from multi-date WiFS data indicated that 35 percent of grid cells had wheat area fraction of less than 0.25, 21 percent between 0.25 – 0.50, 24 percent between 0.50 – 0.75, and 20 percent above 0.75.

The spectral growth profile parameters fitted to district-wise mean NDVI were found to be statistically significant. The values of profile parameters Alpha, Beta, \( T_0 \) and Tmax ranged from 30.2 to 79.8, 9.074E-5 to 2.300E-4, 320.6 to 359.8 Julian days, and 41.8 to 53.8 Julian days, respectively (Table 1). The \( T_0 \) showed a variation of 40 days whereas Tmax showed a variation of only 12 days.

<table>
<thead>
<tr>
<th>DISTRICT</th>
<th>ALPHA (*1.0E-4)</th>
<th>BETA (*1.0E-4)</th>
<th>T0 (Julian day)</th>
<th>Tmax (Julian day)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ambala</td>
<td>47.006</td>
<td>1.40</td>
<td>335.4</td>
<td>43.7</td>
</tr>
<tr>
<td>Bhiwani</td>
<td>69.508</td>
<td>2.01</td>
<td>353.7</td>
<td>49.7</td>
</tr>
<tr>
<td>Faridabad</td>
<td>79.410</td>
<td>2.30</td>
<td>356.2</td>
<td>49.4</td>
</tr>
<tr>
<td>Gurgaon</td>
<td>71.790</td>
<td>2.12</td>
<td>350.3</td>
<td>45.1</td>
</tr>
<tr>
<td>Hissar</td>
<td>79.826</td>
<td>2.27</td>
<td>359.8</td>
<td>53.8</td>
</tr>
<tr>
<td>Jind</td>
<td>75.680</td>
<td>2.17</td>
<td>355.5</td>
<td>51.6</td>
</tr>
<tr>
<td>Kaithal</td>
<td>63.537</td>
<td>1.86</td>
<td>345.9</td>
<td>47.5</td>
</tr>
<tr>
<td>Karnal</td>
<td>56.335</td>
<td>1.68</td>
<td>339.0</td>
<td>43.8</td>
</tr>
<tr>
<td>Kurukshetra</td>
<td>48.789</td>
<td>1.47</td>
<td>333.3</td>
<td>41.8</td>
</tr>
<tr>
<td>Mahendragarh</td>
<td>53.605</td>
<td>1.60</td>
<td>339.7</td>
<td>43.8</td>
</tr>
<tr>
<td>Panipat</td>
<td>63.979</td>
<td>1.88</td>
<td>346.1</td>
<td>46.6</td>
</tr>
<tr>
<td>Rewari</td>
<td>60.219</td>
<td>1.80</td>
<td>343.0</td>
<td>43.4</td>
</tr>
<tr>
<td>Rohtak</td>
<td>73.171</td>
<td>2.11</td>
<td>355.1</td>
<td>50.3</td>
</tr>
<tr>
<td>Sirsa</td>
<td>77.910</td>
<td>2.22</td>
<td>357.3</td>
<td>52.4</td>
</tr>
<tr>
<td>Sonipat</td>
<td>69.561</td>
<td>2.02</td>
<td>351.6</td>
<td>48.5</td>
</tr>
<tr>
<td>Yamunanagar</td>
<td>30.223</td>
<td>0.91</td>
<td>320.6</td>
<td>42.1</td>
</tr>
</tbody>
</table>

\( T_0 \) is days from 1-Jan-2000; Tmax is days from 1-Jan-2001
districts. In these districts the other dominating crop is sugarcane during the season and due to the 188 m ground resolution of WiFS, the NDVI profile contained the combined response of wheat and sugarcane crops. These two districts were dropped from further analysis which resulted in an increase in correlation coefficient from 0.81 to 0.93 (Figure 3).

The difference between T0 and computed DOS ranged from 4 to 16 days with a mean value of 11 days.

For all the parameters, variability captured under scenario-3 was highest followed by that under scenario-2. Under scenario-1 the spatial variability captured is the result of variations in weather and soil characteristics only, whereas, under scenario-2 and scenario-3, the spatial variability captured is due to the variations in management practices besides variations in weather and soil. The variability captured by CGMS in grain yield across the State under scenario-3 is shown in Figure 5.

<table>
<thead>
<tr>
<th>SCENARIO</th>
<th>PARAMETERS</th>
<th>RANGE</th>
<th>MEAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCENARIO-1</td>
<td>ANTHD (days)</td>
<td>85 – 97</td>
<td>90.50</td>
</tr>
<tr>
<td></td>
<td>GFD (days)</td>
<td>27 – 33</td>
<td>29.84</td>
</tr>
<tr>
<td></td>
<td>LAI_ANT</td>
<td>3.58 – 4.46</td>
<td>3.89</td>
</tr>
<tr>
<td></td>
<td>TDM (t/ha)</td>
<td>7.79 – 10.56</td>
<td>9.39</td>
</tr>
<tr>
<td></td>
<td>YLD (t/ha)</td>
<td>3.29 – 4.86</td>
<td>4.06</td>
</tr>
<tr>
<td>SCENARIO-2</td>
<td>ANTHD (days)</td>
<td>81 – 97</td>
<td>87.75</td>
</tr>
<tr>
<td></td>
<td>GFD (days)</td>
<td>25 – 33</td>
<td>28.42</td>
</tr>
<tr>
<td></td>
<td>LAI_ANT</td>
<td>3.73 – 4.53</td>
<td>3.95</td>
</tr>
<tr>
<td></td>
<td>TDM (t/ha)</td>
<td>6.89 – 10.12</td>
<td>8.49</td>
</tr>
<tr>
<td></td>
<td>YLD (t/ha)</td>
<td>2.80 – 4.67</td>
<td>3.72</td>
</tr>
<tr>
<td>SCENARIO-3</td>
<td>ANTHD (days)</td>
<td>80 – 99</td>
<td>87.43</td>
</tr>
<tr>
<td></td>
<td>GFD (days)</td>
<td>25 – 33</td>
<td>28.39</td>
</tr>
<tr>
<td></td>
<td>LAI_ANT</td>
<td>2.78 – 5.05</td>
<td>3.88</td>
</tr>
<tr>
<td></td>
<td>TDM (t/ha)</td>
<td>6.33 – 11.83</td>
<td>8.42</td>
</tr>
<tr>
<td></td>
<td>YLD (t/ha)</td>
<td>2.80 – 5.48</td>
<td>3.69</td>
</tr>
</tbody>
</table>

ANTHD: Pre-anthesis duration; GFD: Post-anthesis duration; LAI_ANT: LAI at anthesis; TDM: Above ground total dry matter; YLD: Grain yield

Table 4: Range and mean of grid-wise simulated crop growth and development parameters under three scenarios of input specification for Haryana (2000-01 season).

The CGMS was run under three different scenarios of crop management input specification as described in methodology section. Under all the scenarios, CGMS generated grid-wise daily outputs of crop growth and development parameters which were visualized in GIS as maps. The spatial variability of these parameters captured by CGMS is summarized in Table 4.
In the present study, CGMS integrates RS information in two ways. Grid-wise crop distribution map derived from RS data is used as weight for computing district-level average yields. Also, the RS-derived spectral-temporal profile based phenology indicators were coupled to CGMS for estimating dates of sowing and its spatial variability. The specification of RS-CGMS-derived dates of sowing for improving CGMS performance is similar to the model “re-initialization” strategy (Moulin et al., 1998). The CGMS framework can also use other RS-derived information such as spatial inputs on agro-meteorological parameters (rainfall, radiation, temperature etc.) and crop biophysical parameters (LAI, fAPAR etc.). The agrometeorological parameters can be directly used as model inputs or for accurate interpolation of point-wise meteorological data. The biophysical products, such as LAI, can be used for in-season model calibration / course correction to simulate accurate crop growth. With the availability of some of these products from MODIS sensor such a possibility is very real though these products need to be validated for their accuracy at multiple sites and time (Pandya et al., 2002).

While in this study, the NDVI as well as LAI were aggregated to district level for the purpose of estimating DOS, it is also possible to apply this technique at individual grid level for capturing within district spatial variability in DOS. The study has also highlighted the need for accurate crop identification as profiles and estimated DOS were sensitive to errors resulting from mixed wheat – sugarcane cropping pattern in two districts of Ambala and Yamunanagar. Improved spatial resolution from AWiFS (50m) is expected to resolve this problem.

The CGMS described in this study predicts actual yield in contrast to potential and water-limited yield as implemented in MARS program. This has been made possible due to (a) use of a production level-3 model which incorporates water and N stress, and (b) specification of district-level variable crop management inputs.

**SCENARIO-1**
\[
y = 0.0617x + 3827.5 \\
r = 0.163 \\
rmse = 404
\]

**SCENARIO-2**
\[
y = 0.347x + 2348.8 \\
r = 0.53 \\
rmse = 472
\]

**SCENARIO-3**
\[
y = 0.9129x + 1.0579 \\
r = 0.74 \\
rmse = 470
\]
4. CONCLUSIONS

A prototype CGMS was developed which could assimilate spatial inputs of weather, soil, crop management, RS-derived crop distribution and link it with wheat simulation model WTGROWS to generate crop parameters and yield maps on a Windows platform. The crop simulation was carried out at each grid, in this case 5’X5’ size.

A procedure for using RS derived crop spectral profile and CGMS derived LAI profiles for estimating the important crop management input of date of sowing at district level was demonstrated. The RS-CGMS-derived phenology, its variation amongst districts, and derived sowing dates were consistent with results from various field studies. The incorporation district-wise dates of sowing and N fertilizer application improved yield prediction.

The simulation model was able to capture the variability in wheat yield within and across districts and the grid-wise weights derived from wheat distribution map allowed for estimation of average district level yields.

The developed CGMS could be an important tool in crop monitoring and yield prediction at different spatial scales. At very fine resolution, CGMS can capture spatial variability in yield at farm / group of farms level for implementing precision crop management and at very coarse resolution it can forecast change in crop growth patterns due to global climate change.

REFERENCES


ACKNOWLEDGEMENTS

The authors wish to acknowledge Dr N. Kalra and Dr P.K. Aggarwal of Centre for Applications of System Simulation, Indian Agricultural Research Institute, New Delhi for providing the model WTGROWS and sharing their expertise in its implementation. A great deal of encouragement and guidance was provided by Shri J.S. Parihar, Group Director, Agricultural Resources Group in carrying out this study.