MONTHLY PROGRESS REPORT
January 6, 1975
Contract NAS 9-14006
Period covered by report – December 1, 1974 to January 1, 1975

CROP STATUS EVALUATIONS AND YIELD PREDICTIONS

Principal Investigator: J. R. Haun
Department of Horticulture
Clemson University, Clemson, S. C.

A. Significant activities and results –

1. A model for predicting the day 50% of the wheat crop is planted in North Dakota has been developed. This model incorporates location (LOC) as another independent variable. The addition of this variable with its transformations brought the total number of independent variables to 49. The Julian date when 50% of the crop was planted for the 9 divisions of North Dakota for 7 years was then regressed on the 49 variables through the step-down multiple regression procedure. This procedure begins with all of the independent variables and sequentially removes variables that are below a predetermined level of significance after each step. The Julian date of 50% planted for the years 1967-72, 74 was obtained as explained in the July 8 and August 5 reports. The validity of this procedure was confirmed by a recent USDA report, N. D. Wheat Historic Estimates 1955-1970, Ag. Statistics No. 33. 1973 data was omitted for use in testing the model. 1974 data were used in the analysis since the year was very atypical and therefore broadened the scope of data for the analysis. The following basic independent variables were used: running three, six, and nine day sums of average minimum and maximum temperature values \((C^0)\) (N3S, N6S, N9S, X3S, X6S, X9S), estimated soil moisture \((E0\) in \%), preseason precipitation \((PP\) in cm), and location \((LOC)\). The basic variables were used as additional independent variables in the form of the following transformations: square and cube; and cross-products of all basic variables. Table 1 shows the statistics for the final step of the analysis, with Fig. 1 showing the relationship of the predicted values to the actual values.

The prediction equation \((\text{Predicted planting date} = 136.7 + 0.0055 (E0 \times N9S) - \text{etc.})\) was tested on daily data not used in the analysis for 1973. An example of the predicted values for one division (represented by dots) is presented in Fig. 2. The curved line is the line of best fit (considering only 1st and 2nd order equations) through the points. The straight line \((45^\circ)\) passes through points where actual and predicted values are equal. The objective of the procedure is to find the date where the value predicted by the equation is equal to the actual Julian date. Thus, the point where the regression line crosses the \(45^\circ\) line is chosen as the predicted planting date. For the Northwest division of North Dakota, the actual date was missed by 8 days using this method. Results for all divisions are as follows:

---

<table>
<thead>
<tr>
<th>Division</th>
<th>Predicted Julian Date</th>
<th>Actual Julian Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Northwest</td>
<td>123</td>
<td>131</td>
</tr>
<tr>
<td>North Central</td>
<td>138</td>
<td>133</td>
</tr>
<tr>
<td>Northeast</td>
<td>124</td>
<td>127</td>
</tr>
<tr>
<td>West Central</td>
<td>116</td>
<td>124</td>
</tr>
<tr>
<td>Central</td>
<td>138</td>
<td>118</td>
</tr>
<tr>
<td>East Central</td>
<td>124</td>
<td>111</td>
</tr>
<tr>
<td>Southwest</td>
<td>126</td>
<td>126</td>
</tr>
<tr>
<td>South Central</td>
<td>100</td>
<td>116</td>
</tr>
<tr>
<td>Southeast</td>
<td>126</td>
<td>112</td>
</tr>
<tr>
<td><strong>Avg.</strong></td>
<td><strong>124</strong></td>
<td><strong>122</strong></td>
</tr>
</tbody>
</table>

The accuracy of this model is considered satisfactory for the purpose of finding the historic dates (i.e. years when no USDA reports were available) on which to initiate our yield prediction model.

2. Growth rate prediction models have been developed for spring wheat for the first two weeks of the 1974 season. These were developed from analysis of 1974 data from Dickinson, Williston (2 locations), Minot, North Dakota and Clemson, South Carolina. Since observations were not begun on the exact date of emergence at each location, it was necessary to establish a stage of development in common for all locations on which to base the subsequent weekly increments of data for analysis. The stage 2.4 was used as the point in common to represent the end of the first week of development. Successive 7-day increments will be used for the subsequent weekly analyses. Each analysis will combine the current week’s data with all previous weeks.

Growth rates were regressed on the following independent variables: maximum air temperature \((C^0)\), minimum air temperature \((C^0)\), precipitation (cm.), estimated soil moisture (100% by Thornthwaite method), solar radiation (Langleys); at 0, 1, 2, and 3 day lag periods; squared, cubed; and 45 selected cross products of the basic lagged variables; for a total of 93, by the step-up multiple regression procedure. This procedure selects the most highly correlated individual independent variable in the first step, and in successive steps variables are added or removed to maintain significance at a designated level (considering all possible combinations). Tables 2, 3, 4, and 5 present statistics of the steps selected for use as growth prediction models.

B. Overall status and problem areas -

Development of prediction models for spring wheat should be completed shortly and then the models will be subjected to tests on domestic data. Several inquiries have been made without success to obtain weather and yield data on the foreign countries for which there is interest in trial predictions. We have on hand a limited amount of Canadian and USSR data which may be used. Unless a source is found soon for data from China, Argentina, Australia and India it is doubtful that tests will be accomplished on these countries before April 1, 1975.
C. Expected accomplishments during January 1975 -
   Analyses will continue on North Dakota spring wheat data and the resulting
   prediction equations will be applied to test data.

D. Recommendations and summary outlook for future work -
   No new recommendations at this time.

E. Travel summary and plans -
   A work conference with NASA personnel and other collaborating research parties
   is scheduled at JSC the week of January 23.
Table 1. Analysis of variance, regression coefficients, and statistics of fit for the dependent variable, day of 50% planted in North Dakota, 1967-72, 1974.

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>Prob&gt;F</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>4</td>
<td>1918.5994</td>
<td>59.7372</td>
<td>0.0001</td>
<td>0.8129</td>
</tr>
<tr>
<td>Error</td>
<td>55</td>
<td>32.1173</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected total</td>
<td>59</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Partial regression coefficients

|                  |      | Student's t | Prob.>|t| |
|------------------|------|-------------|--------|
| Intercept        | 136.6887 | 54.99      | 0.0001 |
| Product of E0 and N9S | 0.0055  | 12.84      | 0.0001 |
| Product of LOC and X9S | -0.1177 | -7.54     | 0.0001 |
| Product of LOC and X3S | 0.2777  | 6.24       | 0.0001 |
| Product of PP and X3S | -0.0138 | -4.71     | 0.0001 |

N9S - Running nine day sum of average minimum temperature values (°C)
X3S - Running three day sum of average maximum temperature values (°C)
X9S - Running nine day sum of average maximum temperature values (°C)
LOC - Location (north, central, south)
PP - Preseason precipitation
E0 - Estimated soil moisture
Fig. 1. Relationship of predicted date on which 50% of wheat was planted to actual date for the period 1967-72, 1974, based on equation from Table 1.
Fig. 2. Predicted planting date (PPLDATE) model applied to all dates from April 1, to June 15 in 1973 for the Northwest Division of North Dakota.
Table 2. Analysis of variance, regression coefficients, and statistics of fit for the
dependent variable, growth rate of spring wheat 1974, one week (with Langley's)

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>Prob. &gt; F</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>5</td>
<td>0.005458</td>
<td>18.252</td>
<td>0.0001</td>
<td>0.843</td>
</tr>
<tr>
<td>Error</td>
<td>17</td>
<td>0.000299</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected total</td>
<td>22</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Partial regression coefficients

|                      | Student's t for H₀:B = 0 | Prob. > |t| |
|----------------------|---------------------------|---------|---|
| Intercept            | 6.971 X 10⁻²              | 3.784   | 0.0015|
| TN0                  | 5.575 X 10⁻³              | 5.108   | 0.0001|
| P1                   | 7.630 X 10⁻²              | 3.537   | 0.0025|
| PSE02                | -8.460 X 10⁻⁶             | -3.102  | 0.0065|
| LSE02/1000           | 3.262 X 10⁻⁵              | 7.601   | 0.0001|
| L3E02/1000           | -6.099 X 10⁻⁵             | -6.389  | 0.0001|

The first number following each basic variable indicates the lag period. The second
number indicates that the variable is squared (2) or cubed (3). Two variables combined
indicate a cross product.

PS - sum of P0, P1, P2, and P3
LS - sum of L1, L2, and L3
Table 3. Analysis of variance, regression coefficients, and statistics of fit for the dependent variable, growth rate of spring wheat 1974, one week (without Langley's)

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>Prob. &gt; F</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>8</td>
<td>0.003593</td>
<td>13.862</td>
<td>0.0001</td>
<td>0.888</td>
</tr>
<tr>
<td>Error</td>
<td>14</td>
<td>0.000259</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected total</td>
<td>22</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Partial regression coefficients | Student's $t$ | Prob. > |t| |
|----------------------------------|---------------|---------|
| Intercept                        | -2.252 X 10^{-1} | -2.44 | 0.0285 |
| TX2                              | 1.719 X 10^{-2} | 6.01   | 0.0001 |
| TX3                              | 1.172 X 10^{-2} | -7.02  | 0.0001 |
| TX12                             | -3.526 X 10^{-4} | -5.44 | 0.0001 |
| TX22                             | 1.345 X 10^{-4} | 3.69   | 0.0024 |
| TN0                              | 9.307 X 10^{-3} | 7.17   | 0.0001 |
| P02                              | -5.701 X 10^{-1} | -4.46 | 0.0005 |
| E3                               | 5.340 X 10^{-3} | 3.23   | 0.0060 |
| E13/1000                         | 2.332 X 10^{-4} | -2.49 | 0.0259 |

The first number following each basic variable indicates the lag period. The second number indicates that the variable is squared (2) or cubed (3). Two variables combined indicate a cross product.
Table 4. Analysis of variance, regression coefficients, and statistics of fit for the dependent variable, growth rate of spring wheat 1974, two weeks (with Langley's)

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>Prob. &gt; F</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>9</td>
<td>0.01587</td>
<td>25.983</td>
<td>0.0001</td>
<td>0.839</td>
</tr>
<tr>
<td>Error</td>
<td>45</td>
<td>0.00061</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected total</td>
<td>54</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Partial regression coefficients | Student's t for H0:B = 0 | Prob. > |t|
|---------------------------------|--------------------------|---------|
| Intercept                       | 6.017 X 10⁻²              | 2.61    | 0.0123  |
| TX33/1000                       | -2.573 X 10⁻³             | -4.62   | 0.0001  |
| TX0P0                          | 1.638 X 10⁻³              | -4.61   | 0.0001  |
| TXOL1                          | 4.660 X 10⁻⁶              | 4.48    | 0.0001  |
| TX2EO                          | 8.180 X 10⁻⁵              | 6.31    | 0.0001  |
| TNOP2                          | 2.280 X 10⁻²              | 9.21    | 0.0001  |
| TN1P2                          | 4.464 X 10⁻²              | -12.28  | 0.0001  |
| P33/1000                       | 3.971 X 10⁰               | -3.39   | 0.0015  |
| P1LI                           | -6.300 X 10⁻⁵             | -3.43   | 0.0013  |
| P2L1                           | 5.605 X 10⁻⁴              | 10.66   | 0.0001  |

The first number following each basic variable indicates the lag period. The second number indicates that the variable is squared (2) or cubed (3). Two variables combined indicate a cross product.
Table 5. Analysis of variance, regression coefficients, and statistics of fit for the dependent variable, growth rate of spring wheat 1974, two weeks (without Langley's)

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>Prob. &gt; F</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>10</td>
<td>0.01354</td>
<td>23.013</td>
<td>0.0001</td>
<td>0.839</td>
</tr>
<tr>
<td>Error</td>
<td>44</td>
<td>0.00059</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected total</td>
<td>54</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Partial regression coefficients | Student's t for $H_0:B = 0$ | Prob. > |t| |
|---------------------------------|-----------------------------|---------|---|
| Intercept                       | $2.334 \times 10^{-1}$     | 4.42    | 0.0001|
| $E1$                            | $-2.394 \times 10^{-3}$    | -2.53   | 0.0150|
| $E_{03}/1000$                   | $2.223 \times 10^{-4}$    | 3.93    | 0.0003|
| $P_{22}$                        | $-8.396 \times 10^{-2}$   | -8.92   | 0.0001|
| $P_{33}/1000$                   | $-5.747 \times 10^{0}$    | -4.37   | 0.0001|
| $P_{1E0}$                       | $-3.145 \times 10^{-4}$   | -3.29   | 0.0020|
| $TX_{33}/1000$                  | $-2.270 \times 10^{-3}$   | -3.91   | 0.0003|
| $TX_{0P2}$                      | $2.089 \times 10^{-2}$    | 10.22   | 0.0001|
| $TX_{0TX2}$                     | $1.708 \times 10^{-4}$    | 6.47    | 0.0001|
| $TN_{1P0}$                      | $-5.762 \times 10^{-3}$   | -5.13   | 0.0001|
| $TN_{1P2}$                      | $-3.392 \times 10^{-2}$   | -9.52   | 0.0001|

The first number following each basic variable indicates the lag period. The second number indicates that the variable is squared (2) or cubed (3). Two variables combined indicate a cross product.