Forecasting the Harvest Date and Yield of Sweet Corn by Complex Regression Models

L.W. Lass, R.H. Callihan, and D.O. Everson

Department of Plant, Soil, and Entomological Sciences, University of Idaho, Moscow, ID 83843

Abstract. Predicting sweet corn (Zea mays var. rugosa Bonaf.) harvest dates based on simple linear regression has failed to provide planting schedules that result in the uniform delivery of raw product to processing plants. Adjusting for the date that the field was at 80% silk in one model improved the forecast accuracy if year, field location, cultivar, soil albedo, herbicides, kernel moisture, and planting date were used as independent variables. Among predictive models, forecasting the Julian harvest date had the highest correlation with independent variables \(R^2 = 0.943\) and the lowest coefficient of variation \(cv = 1.31\%\). In a model predicting growing-degree days between planting date and harvest, \(R^2\) (coefficient of determination) = 0.85 and cv = 2.79%. In the model predicting sunlight hours between planting and harvest, \(R^2 = 0.88\) and cv = 6.41%. Predicting the Julian harvest date using several independent variables was more accurate than other models using a simple linear regression based on growing-degree days when compared to actual harvest time.

Forecasting the harvest date from carefully scheduled planting date is becoming an important tool for providing uniform delivery of raw product in the vegetable processing industry. Accurately estimating the amount of raw product ready for harvest and processing each day would reduce extra labor needs during peak production and result in more efficient equipment use. However, accurately scheduling planting and harvesting and predicting yield depend on quantifying the relationships among environmental variables, biological processes, and phenological events.

Réaumur (1735) recognized that basking plant development rate on heat-unit accumulation is preferable to using growth days. Hanna (1925) observed a high correlation between accumulated mean temperature and corn growth. Others have modified Réaumur's original thermal-constant equation, and the most commonly used growing-degree-day equation for sweet corn is: \(\text{minimum daily temperature} + \text{maximum daily temperature}/2\) - the base temperature of 9.44C.

This growing-degree equation may be modified to have the maximum in the optimum range set between 27.7 and 32.2C, where temperatures above the maximum are artificially set equal to the maximum for use in the equation (Aung et al., 1967; Davitaya, 1964; Gilmore and Rogers, 1958; Hortik and Arnold, 1965).

Gilmore and Rogers (1958) found that the growing-degree-day equation was inadequate for measuring sweet corn maturity in Texas because of the cool nights, and they suggested a modified system of heat units for cool nights, based on degree-hours, using the equation: \([(\text{minimum temperature} + \text{maximum temperature for a 3-h period})/2] - 10\text{C}, \text{ with an optimum temperature of 30C}.

Degree-day equations have been modified further to substitute soil temperatures for air temperatures in calculating accumulating heat units between planting and seedling emergence. Most research on soil-degree-day equations originated in the midwestern United States, where soils warm rapidly in the spring. Aung et al. (1967) observed that the soil temperature during the first 2 weeks of plant development made little difference in the time of sweet corn maturity. These observations may hold true for corn planted in soil > 15C, the lowest temperature Aung and coworkers used, but not necessarily for lower planting temperatures.

Sweet corn phenological events are forecast using simple or complex multiple regression models. Simple models use a single factor from the planting date to the development date of harvestable ears. The single factor is usually a measure accumulated over time (e.g., heat units, days, or sunlight). Multiple-regression systems predict the time expected to elapse from planting date to the development of harvestable ears, after adjusting for a specific growth stage (e.g., silk date).

In comparing degree-day equations for predicting sweet corn maturity in the midwestern United States, Arnold (1974) found that the growing-degree-day equation was simple to calculate, yielding an accuracy level similar to the Gilmore and Rogers (1958) equation and the Ferguson complex-regression system (Arnold, 1960).

Our research objectives were: 1) to analyze environmental factors and management variables that affect sweet corn yields and maturity rates, 2) to balance the importance of environmental and management factors in regulating and predicting harvest schedules, and 3) to correlate maturity and yield forecasts with actual harvest values.

Materials and Methods

The sweet corn harvest readiness date was predicted from five types of phenological values accumulated during seven growing seasons (1979-1985) at Buhl, Idaho. These values were 1) growing-degree days between planting date and harvest (from 66% to 77% kernel moisture), 2) sunlight hours between planting date and harvest, 3) the Julian date when 80% of the field plants have silked, 4) the Julian harvest date, and 5) accumulated analog heat units between planting date and harvest for 2 years (1984 and 1985). Growing-degree days were calculated from temperatures measured with a 7-day circular recorder with air and soil probes (Taylor Instrument, Sybron Corp., Arden, N.C.). Growing degrees were calculated according to the Barger equation: \([(\text{daily minimum temperature} + \text{daily maximum temperature})/2] - 9.4\text{C}, \text{ with maxima } > 30\text{C considered equal to } 30\text{C and minima } < 9.4\text{C consid-}
ered equal to 9.4°C (Arnold, 1974). Temperatures used to calculate analog heat units were determined from the resistance of an external 10-V thermocouple probe (Omni Data, Logan, Utah).

Analog heat units were calculated at 10-min intervals using the sum of the growing-degree equation: \[ \text{temperature for } 24 \text{ h at 10-min intervals}/144 \times 9.4°C = \text{maxima >30°C changed to 30°C, and minima <9.4°C changed to 9.4°C.} \]

Sunlight hours were accumulated from a National Weather Service sunrise-sunset table (National Weather Service, Kimberly, Idaho), which lists the daylight hours between planting date and harvest. Sunlight-hour calculations did not account for cloudy days; however, <15% of the days are cloudy in this desert region during the growing season (Environmental Science Services Administration, 1968).

Seven weather stations within the study area were used to record temperatures for heat-unit calculations. Weather station sites were located in garden spots of growers and represented the areas in production. The air-temperature probes at six stations were located in stationary, wooden shelters (Loomis, 1890). Instruments and air-temperature probes at one weather station were shaded by a white, wooden Stevenson meteorological shelter (Geddes, 1930). The air-temperature probes for the circular-chart recorder and the analog computer were 122 cm above the ground. The soil probes from the circular-chart recorder and the analog computer were placed next to each other in bare soil at a 7.5-cm depth. Soil moisture was maintained by weekly flood irrigation. From 3 to 4 cm of soil over the probes was loosened with a shovel after each irrigation to reduce soil albedo and soil dryness around the probes.

The model components for forecasting sweet corn harvest dates and yields were established from a commercial processor's 7-year data base containing production information from 3029 sweet corn fields. The independent variables of the semiempirical models were year, field location, cultivar, planting date, soil albedo weighted by planting date, sweet corn maturity stage at harvest, and herbicide family weighted by planting date.

Year, as a model component, was represented by a categorical coded value for each of the seven growing seasons (1979-1985) of the historical data base. Variability among field sites was reduced by clustering the field locations into 80 clusters of four sections (≈2.59 km²) or into seven large clusters based on field proximity to the nearest weather station. The field location was described by a categorical coded number indicating either the township, range, and the field's coded cluster number, or the weather station number. A township is a U.S. public lands survey unit of measure denoting a north-south column of six sections, each of which is usually ≈2.59 km², but which can vary to account for planetary sphericity. A range denotes an east-west row of six sections. The Boise base line and meridian were used for the reference lines of the townships and ranges in this study. The coded cluster number is a grouping of the 36 sections of a township into nine clusters of four sections each.

Sweet corn cultivars were coded as categorical values. Although the data base history covered 7 years, only cultivars 3 and 4 were grown in all 7 years, because new cultivars were being introduced. New cultivars 1, 2, and 6 were grown for only 3 of the 7 years, and cultivar 5 was grown for 5 years. The planting date was the date planting was completed in the field. Sweet corn maturity was estimated from the kernel moisture level using an optical refractometer.

The model's herbicide component was represented by a numeric categorical code for the herbicide family (i.e., thiocarbamate, phenoxy, triazine, or no herbicide) used on the sweet corn for weed control (Anderson, 1983). weighted to stratify by planting date in 10-day groups using the second digit of the Julian planting date. For example, if the herbicide code was 3 and the planting date was 142, the code stratifying herbicide by planting date is 34.

Models forecasting yields used prior crop and plant population as additional independent variables. The prior crop component was a numeric categorical value coded for the crop grown in the field the previous year. Plant populations were determined at the seeding stage.

Soil albedo as a categorical value was represented by the percent of reflected light from an average of four dried soil samples from each section, weighted by planting date in 10-day groups using the second digit of the Julian planting date. For example, if the albedo in two fields was 16% and the Julian planting dates were 132 and 151, the codes stratifying albedo by planting date were 163 and 165. The soil was sampled four times to a depth of 20 cm near each of the four corners in each section using a 5-cm-diameter soil auger, and the samples were air dried. Samples were not screened, but large clods were broken apart. The apparatus used to measure the albedo consisted of four 150-W incandescent light bulbs, a 20-cm round aperture in a 50 × 40 cm metal sheet painted flat black (Krylon, Borden, Columbus, Ohio), and a quantum radiometer (Model LI 190, LI-COR, Lincoln, Neb.) measuring light in the photosynthetic range. The radiometer's sensor was shaded by placing it in a black plastic 35-mm film canister with the bottom removed to allow passage of the instrument wire. The canister was taped to a brazing rod holder that was painted black. The shielded sensor was positioned 23 cm above the aperture and held in place by the sensor holder. The light source was 6 cm above the sensor with the four bulbs pointing toward the sensor. The four bulbs were positioned to make a hallo-like square of light measuring 30 × 30 cm around the sensor. The apparatus was placed in a semidark room. Incoming light at the sample surface was 256 µmol·m⁻²·s⁻¹. The dry soil was spread in a 1.75-cm-thick layer and leveled over a 25-cm-diameter white plastic disposable dinner plate. The plate, which was placed directly under the aperture, touched the aperture metal.

The five dependent variables were y (Julian harvest date), y' (Julian date of 80% silk), y'' (analog heat units from planting date to harvest date), y''' (accumulated heat units from planting date to harvest date), and y'''' (accumulated sunlight hours to harvest date).

Two generalized least square models (one simple and one complex) were defined for each of the five dependent variables. The simple model was: \[ y_{ijk} = \mu + A_i + B_j + C_{ij} + e_{ijk}, \] where \( \mu \) is overall mean, \( A_i \) is effect of \( i^{th} \) field location (i = 1, 2, 3, ..., 150), \( B_j \) is effect of \( j^{th} \) cultivar (i.e., 1, 2, 3, ..., 8), \( C_{ij} \) is effect of \( k^{th} \) year (k = 79, 80, 81, ..., 84), \( D_{il} \) is effect of \( m^{th} \) plant altitude - planting date code (l = 11, 12, ..., 116), \( E_{im} \) is effect of \( n^{th} \) herbicide, \( X_{ik} \) is Julian planting date, and \( X_{ij} \) is Julian harvest date, and \( e_{ijk} \) is model error.

Additionally, for Julian harvest date, a regression coefficient for Julian 80% silk date was included.

Models forecasting yields contained prior crop and plant population as additional independent variables. The prior crop component was expressed as an arbitrary numeric value coded for the crop grown the previous year. Sweet corn yield as the dependent variable in metric tons per hectare was: \[ y_{ijklm} = \mu + A_i + B_j + C_{ij} + D_{il} + E_{im} + F_{ln} + G_{ijlm} + H_{ijlm}, \] where all variables are defined as before and \( F_{ln} \) is effect of the \( n^{th} \) crop grown the previous year.
previous crop \( (n = 1, 2, 3, \ldots, 6) \), \( \beta_i X_i \) = Julian harvest date, and \( \beta_j X_j \) = sweet corn plants per hectare.

**Results**

The simple models were distinctly less accurate than the complex models in predicting events and values for sweet corn growth, harvest, and heat units (Table 1). For example, with the simple model predicting days from planting to 80% silking date, \( R^2 = 0.695 \) and \( cv = 3.63 \), whereas with the complex model, \( R^2 = 0.933 \) and \( cv = 1.56 \). The model's \( R^2 \) indicates the variability in the dependent variable that can be accounted for by the model. The \( cv \) describes the relative variation of the estimated dependent variable (SAS, 1985). Yield, however, could not be predicted by either model (Table 1).

Some variability within the dependent variable was reduced by clustering fields into regional groups around weather stations or into small groups of four sections. The reduced variability of the dependent variable by clustering fields into regional groups and small groups improved the accuracy of both models (Table 2). The improved accuracy was greatest in the heat-unit models. In the simple heat-unit model without clustering, \( R^2 = 0.299 \), but it increased to 0.685 when four-section clusters were used. Regional clustering increased \( R^2 \) and decreased \( cv \), but never as much as the four-section clusters. Changing cluster size also reduced the simple heat-unit model \( cv \) from 6.56 for nonclustered data to 4.48 for the four-section clusters. The \( R^2 \) of complex models improved from 0.560 for nonclustered data to 0.733 for the four-section clusters.

The simple model's lower \( R^2 \) and higher \( cv \) resulted in large part from the variability of the dependent variables, which were not consistent from year to year. For example, the heat units for cultivar 2 ranged from 938 to 1096, with \( sd \) (standard deviations) from 7 to 75. The heat units from planting to harvest over 7 years ranged from 965 to 1065 for cultivar 3, with \( sd \) from 37 to 103 (Table 3). The lack of a consistent number of heat units accumulated between planting date and harvest was shown by mean differences between cultivars. For example, the differences between cultivars 1 and 3 ranged from -33 to 70 heat units (Table 3), and between cultivars 5 and 3 from 46 to 256.

Table 1. Accuracy of simple and complex models used to predict sweet corn phenological events and yields.

| Forecasted events and values | Model type | Predictability |  
|-----------------------------|------------|----------------|
|                             | \( R^2 \)  | \( cv \)        |
| 80\% silk date              | Simple     | 0.695          | 3.63          |
|                             | Complex    | 0.933          | 1.56          |
| Harvest date                | Simple     | 0.660          | 3.60          |
|                             | Complex    | 0.919          | 1.58          |
| Harvest date including partial regression of 80\% silk date | Simple | 0.940 | 1.51 |
|                             | Complex    | 0.943          | 1.31          |
| Heat units                  | Simple     | 0.685          | 4.48          |
|                             | Complex    | 0.733          | 3.56          |
| Analog heat units           | Simple     | 0.669          | 4.39          |
|                             | Complex    | 0.854          | 2.97          |
| Sunlight hours              | Simple     | 0.853          | 7.20          |
|                             | Complex    | 0.888          | 6.41          |
| Yield                       | Simple     | 0.123          | 21.1          |
|                             | Complex    | 0.191          | 20.2          |

Table 2. Effects of field clustering and cluster size on the predictability of sweet corn phenological events.

| Model type and clustering procedure | Predictability |  
|------------------------------------|----------------|
|                                    | \( R^2 \)      | \( cv \)        |
| 80\% silk date                     |                |                |
| *Simple model*                     |                |                |
| Not clustered by location          | 0.565          | 4.25           |
| Locations in multisection clusters | 0.613          | 4.01           |
| Locations in four-section clusters | 0.695          | 3.63           |
| *Complex model*                    |                |                |
| Not clustered by location          | 0.923          | 1.66           |
| Locations in multisection clusters | 0.924          | 1.65           |
| Locations in four-section clusters | 0.933          | 1.56           |
| Harvest date                        |                |                |
| *Simple model*                     |                |                |
| Not clustered by location          | 0.524          | 4.18           |
| Locations in multisection clusters | 0.575          | 3.95           |
| Locations in four-section clusters | 0.660          | 3.60           |
| *Complex model*                    |                |                |
| Not clustered by location          | 0.904          | 1.70           |
| Locations in multisection clusters | 0.905          | 1.68           |
| Locations in four-section clusters | 0.919          | 1.58           |
| Harvest date including partial regression 80\% silk date |                |                |
| *Simple model*                     |                |                |
| Not clustered by location          | 0.933          | 1.56           |
| Locations in multisection clusters | 0.935          | 1.54           |
| Locations in four-section clusters | 0.940          | 1.51           |
| *Complex model*                    |                |                |
| Not clustered by location          | 0.937          | 1.36           |
| Locations in multisection clusters | 0.939          | 1.35           |
| Locations in four-section clusters | 0.943          | 1.31           |
| Heat units                          |                |                |
| *Simple model*                     |                |                |
| Not clustered by location          | 0.299          | 6.56           |
| Locations in multisection clusters | 0.528          | 5.39           |
| Locations in four-section clusters | 0.685          | 4.48           |
| *Complex model*                    |                |                |
| Not clustered by location          | 0.560          | 4.50           |
| Locations in multisection clusters | 0.654          | 4.00           |
| Locations in four-section clusters | 0.733          | 3.56           |
| Analog heat units                  |                |                |
| *Simple model*                     |                |                |
| Not clustered by location          | 0.364          | 5.88           |
| Locations in multisection clusters | 0.423          | 5.61           |
| Locations in four-section clusters | 0.609          | 4.39           |
| *Complex model*                    |                |                |
| Not clustered by location          | 0.716          | 3.98           |
| Locations in multisection clusters | 0.787          | 3.47           |
| Locations in four-section clusters | 0.854          | 2.97           |
| Sunlight hours from planting date to harvest |                |                |
| *Simple model*                     |                |                |
| Not clustered by location          | 0.817          | 7.86           |
| Locations in multisection clusters | 0.818          | 7.83           |
| Locations in four-section clusters | 0.853          | 7.20           |
| *Complex model*                    |                |                |
| Not clustered by location          | 0.876          | 6.64           |
| Locations in multisection clusters | 0.878          | 6.59           |
| Locations in four-section clusters | 0.888          | 6.41           |

452 J. Amer. Soc. Hort. Sci. 118(4):450-455. 1993.
Simple models using days between planting and 80% silk or harvest also showed year-to-year inconsistency (Table 3). The largest variation from planting to 80% silk was with cultivar 3 (71 to 88 days) and from planting to harvest with cultivars 3 (103 to 119 days) and 4 (95 to 133 days). The largest deviation from cultivar 3, in days from planting to harvest, was 22 in cultivar 4 (Table 3). Other cultivar differences were smaller.

The variability of means and SD of phenological measurements of sunlight hours, heat units accumulated with analog equipment, and days from planting date to 80% silk are representative of those found in other models (Table 3). The mean differences among all of the tested phenological measurements of cultivars, when compared to cultivar 3, did not remain constant from year to year.

Predicted values were subtracted from actual values to determine model accuracy. Model accuracy of forecasted values shows that the error in forecasts was random for models based on days.

### Table 3. Average phenological values from five predictive models \( Y \) for six cultivars (cv) of sweet corn compared with cultivar 3.

| Year | Days until 80% silk | Days until harvest | Heat Units until harvest | Heat Units until harvest | Sunlight hours until harvest |
|------|---------------------|--------------------|--------------------------|--------------------------|-----------------------------|
|      | Deviation from cv 3 | Mean dev from cv 3 | Deviation from cv 3      | Mean dev from cv 3       | Deviation from cv 3         |
|      |                     | Mean dev from cv 3 | Deviation from cv 3      | Mean dev from cv 3       | Deviation from cv 3         |
| 1983 | 82 ± 4              | 103 ± 1            | 1006 ± 50                | 1116 ± 159               |
| 1984 | 89 ± 6              | 112 ± 6            | 1135 ± 85                | 1660 ± 86                |
| 1985 | 74 ± 7              | 106 ± 5            | 968 ± 107                | 1544 ± 87                |
| All years | 82 ± 6       | 107 ± 4            | 1036 ± 81                | 1569 ± 111               |

\[ \text{Cultivar 1} \]

| Year | Days until 80% silk | Days until harvest | Heat Units until harvest | Heat Units until harvest | Sunlight hours until harvest |
|------|---------------------|--------------------|--------------------------|--------------------------|-----------------------------|
| 1983 | 86 ± 1              | 104 ± 1            | 938 ± 7                  | 1263 ± 180               |
| 1984 | 91 ± 6              | 114 ± 5            | 1096 ± 42                | 1658 ± 85                |
| 1985 | 84 ± 5              | 114 ± 4            | 1004 ± 75                | 1662 ± 64                |
| All years | 87 ± 4       | 111 ± 3            | 1013 ± 42                | 1649 ± 74                |

\[ \text{Cultivar 2} \]

| Year | Days until 80% silk | Days until harvest | Heat Units until harvest | Heat Units until harvest | Sunlight hours until harvest |
|------|---------------------|--------------------|--------------------------|--------------------------|-----------------------------|
| 1979 | 82 ± 7              | 111 ± 6            | 1044 ± 50                | 1622 ± 100               |
| 1980 | 85 ± 9              | 119 ± 9            | 995 ± 39                 | 1682 ± 147               |
| 1981 | 88 ± 8              | 109 ± 9            | 1040 ± 37                | 1560 ± 147               |
| 1982 | 84 ± 7              | 106 ± 8            | 965 ± 40                 | 1078 ± 100               |
| 1983 | 84 ± 5              | 108 ± 6            | 992 ± 63                 | 1083 ± 105               |
| 1984 | 81 ± 9              | 103 ± 8            | 1065 ± 82                | 1488 ± 142               |
| 1985 | 71 ± 4              | 106 ± 5            | 1001 ± 103               | 1526 ± 94                |
| All years | 82 ± 7       | 109 ± 7            | 1014 ± 59                | 1427 ± 119               |

\[ \text{Cultivar 3} \]

| Year | Days until 80% silk | Days until harvest | Heat Units until harvest | Heat Units until harvest | Sunlight hours until harvest |
|------|---------------------|--------------------|--------------------------|--------------------------|-----------------------------|
| 1979 | 82 ± 4              | 108 ± 4            | 1056 ± 116               | 1616 ± 62                |
| 1980 | 85 ± 3              | 113 ± 5            | 1082 ± 73                | 1661 ± 68                |
| 1981 | 84 ± 3              | 112 ± 5            | 1015 ± 51                | 1642 ± 71                |
| 1982 | 84 ± 3              | 110 ± 2            | 914 ± 28                 | 895 ± 32                 |
| 1983 | 71 ± 2              | 95 ± 1             | 987 ± 13                 | 815 ± 23                 |
| 1984 | 90 ± 7              | 113 ± 8            | 1191 ± 78                | 1677 ± 118               |
| 1985 | 76 ± 4              | 100 ± 3            | 976 ± 21                 | 1498 ± 48                |
| All years | 82 ± 4       | 107 ± 4            | 1031 ± 54                | 1575 ± 60                |

\[ \text{Cultivar 4} \]

| Year | Days until 80% silk | Days until harvest | Heat Units until harvest | Heat Units until harvest | Sunlight hours until harvest |
|------|---------------------|--------------------|--------------------------|--------------------------|-----------------------------|
| 1979 | 78 ± 3              | 105 ± 3            | 1098 ± 167               | 1570 ± 45                |
| 1980 | 83 ± 5              | 111 ± 4            | 1251 ± 24                | 1623 ± 55                |
| 1981 | 82 ± 2              | 107 ± 2            | 1004 ± 55                | 1575 ± 32                |
| 1982 | 78 ± 3              | 103 ± 3            | 981 ± 136                | 798 ± 39                 |
| 1985 | 71 ± 2              | 94 ± 2             | 955 ± 49                 | 1413 ± 34                |
| All years | 78 ± 3       | 104 ± 3            | 1058 ± 86                | 1316 ± 41                |

\[ \text{Cultivar 5} \]

| Year | Days until 80% silk | Days until harvest | Heat Units until harvest | Heat Units until harvest | Sunlight hours until harvest |
|------|---------------------|--------------------|--------------------------|--------------------------|-----------------------------|
| 1982 | 81 ± 4              | 106 ± 2            | 897 ± 68                 | 838 ± 34                 |
| 1984 | 87 ± 1              | 110 ± 0            | 1030 ± 6                 | 1636 ± 6                 |
| 1985 | 75 ± 1              | 95 ± 2             | 946 ± 20                 | 1430 ± 22                |
| All years | 81 ± 2       | 104 ± 1            | 960 ± 31                 | 1185 ± 21                |

\[ \text{Cultivar 6} \]

### Notes

3Heat unit constants are accumulated growing degree days in Celsius. Some varieties were not grown in all years and are indicated by ---.
errors in predicted values for the harvest date model, including partial regression of 80% silk date, were evenly distributed among independent variables. The $P$ values for individual variables were 0.999 for cluster location, 0.98 for year, 0.66 for sweet corn cultivar, 0.990 for soil albedo, 0.780 for kernel moisture, 0.270 for herbicide family, and 0.111 for planting date (Table 4). The prediction error for the heat-unit model was correlated with the independent variables of the model ($R^2 = 0.61$ and $P = 0.00001$). The independent variables that contributed to the correlation were cluster location ($P = 0.0022$); year, kernel moisture, and herbicide family ($P = 0.0001$); soil albedo ($P = 0.065$); and cultivar ($P = 0.0003$). The only independent variable that did not contribute significantly to the prediction error of the heat-unit model was planting date.

**Discussion**

Simple models using year, cultivar, and field location failed to accurately predict harvest dates in this test. The lack of a reliable simple model to predict harvest dates is reflected in low $R^2$ and high cv values. Complex multiple regression models increased the correlation of the dependent variable with independent variables, and decreased variation. The complex model’s independent variables accounted for more of the variability of the phenological constants than did those of the simple models.

Models predicting sweet corn maturity based on days had greater consistency among years (7% to 15%) and smaller deviations within any given year (2% to 8%), when compared to other models. Models using temperature to predict sweet corn maturity had the greatest variability among years, with constants ranging from 9% for cultivar 3 to 27% for cultivar 5. The $R^2$ of the temperature-based constant within a single year ranged from 7% to 12%. The general errors can be attributed to among-year and within-year variability and must be accounted for in all models before forecasting accuracy can be increased to more accurately reflect silk dates, harvest dates, and yields. Errors resulting from nutrients and watering differences appeared to only affect the model when severely deficient and probably would not occur in most fields under irrigated production (data from unpublished experiments, not shown).

Models predicting harvest dates must compensate for the among-year and within-year deviation for each cultivar. It is more important for the time between planting date and harvest to be consistent within a year than between or among years. Delaying all cultivars in the harvest schedule by 5 days would be compensated by starting harvest 5 days later, but a 5-day variation within a year would provide more sweet corn at some time during harvest than the factory could process, and would require below-capacity operation at other times. Model adjustment for among-year and within-year variation assumes that phenological time differences are constant among cultivars. An example of a constant phenological time difference is when a cultivar’s harvest date bears a constant temporal relationship to that of another cultivar. The assumption that cultivars bear a constant phenological relationship to each other was not true in these studies. The heat-unit differences between cultivar 3 and the other cultivars was 102 for cultivar 1, 84 for cultivar 2, 175 for cultivar 4, 301 for cultivar 5, and 40 for cultivar 6. The difference in days between cultivar 3 and the other cultivars was 12 days for cultivar 1, 13 days for cultivar 2, 21 days for cultivar 4, and 11 days for cultivars 5 and 6. Adjusting a model by the phenological time difference would add an error, because phenological time differences are not constant among fields.

The basic premise of the heat-unit theory, as outlined by Arnold and other investigators, suggests that the same number of heat units should accumulate between planting and harvest, independent of the field location (Arnold, 1960 and 1974; Aung et. al., 1967; Davitaya, 1964; Gilmore and Rogers, 1958; Hanna, 1925; and Réamur, 1735). However, models based on heat units accumulated between planting and harvest showed the greatest improvement in accuracy of all models tested when fields were clustered in four-section groups. The simple heat-unit model, with clusters of four sections, doubled correlation values and decreased the cv by one-third. The correlation between planting and harvest and the complex heat-unit model increased 30% when the cluster size was reduced to four sections. The correlation of simple models forecasting heat units accumulated by analog equipment increased 80% and the cv decreased 25%. The improvements in $R^2$ and cv stemming from smaller clusters suggest that, although heat-unit accumulation is the major determinant of phenological events in sweet corn, other factors influence these events significantly. Testing models by forecasting the difference between predicted and actual heat units from planting to harvest showed that the cluster location significantly improved model correlation with the independent variables. The error in heat

### Table 4. Source and significance of errors in sweet corn harvest and 80% silk date $Y$ forecasting models.

| Type of forecasting model | Model statistic ($Pr > F$) | Cluster location ($R^2$) | Year | Variety | Albedo | Kernel moisture | Herbicide | Planting date |
|---------------------------|---------------------------|--------------------------|------|---------|--------|----------------|-----------|---------------|
| 80% silk date             | 0.99 0.046                | 0.9088                   | 0.9690 | 0.9800 | 0.9800 | ---             | 0.9900    | 0.7100        |
| Harvest date              | 0.86 0.050                | 0.9999                   | 0.9999 | 0.6666  | 0.9900 | 0.7800          | 0.2700    | 0.1110        |
| Harvest date including partial regression of 80% silk date | 0.99 0.043 | 0.9999 | 0.9800 | 0.6600 | 0.9900 | 0.3900 | 0.3000 | 0.0900 |
| Heat units (GGD equation) | 0.0001 0.610             | 0.0022                   | 0.0001 | 0.0003 | 0.0651 | 0.0001          | 0.0001    | 0.3960        |
| Heat units (analog)       | 0.0001 0.430             | 0.0001                   | 0.0001 | 0.4700 | 0.2600 | 0.1500          | 0.9800    | 0.3800        |
| Sunlight hours            | 0.0001 0.300             | 0.9965                   | 0.0001 | 1.0000 | 0.0005 | 0.8290          | 0.0820    |

*Independent variables introduce error when probabilities for differences between predicted and actual forecasted values are significant at $P = 0.05$. Probability of receiving a greater $F$ value is indicated by ($Pr > F$).*
units forecasts accumulated by analog equipment was attributed to cluster location and year, since the other independent variables were not significant contributors ($P > 0.10$). The errors of models predicting heat units from the growing-degree equation were due to year, cluster size and location, kernel moisture, herbicide, cultivar, and soil albedo as independent variables. The higher correlation and the significance of the model testing errors suggest that temperature alone fails to account for all factors influencing number of heat units between planting and harvest.

Reducing the cluster size for models based on days from planting to harvest resulted in a slight increase in $R^2$ and a decrease in cv. When testing error by forecasting the difference between predicted and actual values of models based on day, the cluster size did not significantly contribute to the correlation for independent variables. In fact, the correlations for the model testing this error were low ($R^2 < 0.05$, $P = 0.86$ to 0.99) and indicate that none of the independent variables were major contributors to the correlation ($P = 0.09$ to 0.99). The low correlation and the lack of significance of the model testing forecasting error indicate that the differences between predicted and actual values were evenly distributed within each independent variable.

The model based on sunlight hours between planting date and harvest also had the highest correlation and lowest cv when the cluster size was four sections. This suggests that the sunlight hours between planting date and harvest could not account for all variability within a region. The error of the predicted values was not related specifically to a single location, but was distributed over all clustered field locations. A uniform error in the model testing forecasted values indicates that another independent variable may improve the forecasting accuracy of sunlight hours between planting and harvest. The errors in sunlight-hour models come from cultivar, location, and kernel moisture.

The most accurate models have some built-in errors in a few of the forecasted values because of variability in farming and environmental conditions within and among years. Although the independent variables of models based on days do not account for all factors, errors in predicted values were distributed evenly among the independent variables. The errors of temperature-based prediction models were correlated with specific independent variables that reduced data base variability.

The best forecasting method to provide an even flow of sweet corn into the factory requires that the independent variables of the model account for all factors that influence rates of phenological development. Forecasts with a complex model (planting date to harvest date including partial regression of 80% silk date), based on days, were the most accurate. The correlation for the planting date-harvest date model, including partial regression of the 80% silk date, was 0.94, with cv = 1.31. The lower model correlation based on heat units only from the growing-degree-day equation or heat units accumulated with analog equipment models, when compared to models based on days, suggests that the regression equations based on heat units alone do not account for as many factors as do models based on days.

### Literature Cited

Anderson, W.P. 1983. Weed science principles 2nd ed. West Publishing Co., St. Paul, Minn.
Arnold, C.Y. 1960. Maximum-minimum temperature as a basis for computing heat units. J. Amer. Soc. Hort. Sci. 76:682-492.
Arnold, C.Y. 1974. Predicting stages of sweet corn development. J. Amer. Soc. Hort. Sci. 99:501-505.
Aung, L.H., F.G. Teubner, and J.O. Young. 1967. Effect of temperature on maturity of sweet corn *Zea mays rugosa*. J. Amer. Soc. Hort. Sci. 92:516-522.
Davitaya, F.F. 1964. A method of predicting heat units and duration of the growth period. Agr. Meteorol. 2:109-119.
Environmental Science Services Administration. 1968. Climatic atlas of the United States. U.S. Department of Commerce, Washington, D.C.
Geddes, A.E.M. 1930. Meteorology. Blackie & Son Limited, London.
Gilmore, E.C. and J.S. Rogers. 1958. Heat units for measuring maturity in corn. Agron. J. 50:611-615.
Hanna, W.F. 1925. The nature of the growth rate in plants. Sci. Agr. 5:133-138.
Hortik, H.J. and C.Y. Arnold. 1965. Temperature and rate of development of sweet corn. J. Amer. Soc. Hort. Sci. 87:303-312.
Loomis, E. 1890. A treatise on meteorology. Harper & Brothers, New York.
Réaumur, R.A.F. De. 1735. Observations du thermométre, faites à Paris pendant lannée 1735 comparées avec celles qui ont été faites sous la ligne, à lIsle de France, à Alger et en quelques-unes de no isles de lAmérique. Paris Memoirs, Acad. des Science, Année.
SAS Institute. 1985. SAS/STAT guide for personal computers (version 6 ed.). SAS Institute, Cary, N.C.