Preliminary Report on
The Use of Time Related Growth Models in
Forecasting Components of Corn Yield

By

WENDELL W. WILSON

Research and Development Branch
Research Division
Statistical Reporting Service

May 1974

Introduction

In 1973, a preliminary investigation of time related growth models was conducted. The objectives of the investigation were: (1) to determine if biological laws relating corn growth to various time scales exist, and (2) if they exist, what utility they may have in forecasting important components of corn yield. The investigation was a continuation of earlier research efforts to develop new techniques for forecasting weight per ear during the earlier part of the growing season. The overall goal is to improve methods of forecasting the amount of dry matter deposited in kernels. Forecasts of dry kernel weight can be readily expressed in terms of two important yield components; (1) kernel weight per ear at 15.5 percent moisture, and (2) weight per kernel at 15.5 percent moisture.

The Sample

The basic population sampled in the 1973 study included all corn plants in three central Iowa fields. The three fields were purposely selected to be different. However, they were located in the same general area. Three row segments of 400 plants each were randomly located within each field. An initial or zero growth time based on the event "silks starting to dry" was observed for most of the plants. Domains of plants for which this time variable was defined were identified in each row segment. Random samples of plants were selected from each domain to be observed on each of ten specific field visits during the season. The plants were destructively sampled by removing the ears. Ears were then subsampled during the laboratory phase of the investigation by randomly selecting two kernel rows per ear.

Data Collection

Data collection involved intensive observation and sampling of plants in the three central Iowa fields. The preliminary investigation required frequent and detailed observation of various phenological events. These events generally related to various phases of the kernel fertilization period. The observations were made at four day intervals from July 25 through August 14. The phenological observations were followed by periodic visits tied to estimating dry weight per ear and per kernel as the season progressed. These field visits were made every eight days starting on August 14 and ending on October 25. Determining dry grain weight required some rather detailed sampling and laboratory operations.

The Iowa State Statistical Office provided supervision and assistance in all data collection phases. Outstanding effort and interest by Duane Skow and his staff helped insure high quality work. The input of views and ideas by Dick Knight, George Hanuschak and Glenn Wassom was especially helpful.

The Growth Model

The general form of the logistic growth model utilized in this study is:

$$\hat{Y}_t = \frac{1}{\hat{\alpha} + \hat{\beta}(\hat{\rho})^t}$$

This is basically a non-linear model where t is the independent time variable. Y_t is the dependent variable, and α, β and ρ are the parameters which can be estimated from data sets of the form:

$$\begin{array}{ll} t_1 & Y_{t_1} \\ t_2 & Y_{t_2} \\ \vdots & \vdots \\ t_n & Y_{t_n} \end{array}$$

In this application, $\hat{Y}_t$ is the estimated mean dry weight per ear or per kernel at time t. The variable t is the time (days) after one of the phenological events; tassel emerged, silk emerged, silk starting to dry, silk continuing to dry, silk finishing or finished drying or the "time" variable dry matter percentage of the grain when sampled.

The basic model uses repeated observations from the current year to estimate the parameters needed to predict the dependent variable (dry weight per ear or per kernel) at maturity. The model may be updated at various times during the growing season as more data becomes available and as data becomes available for later stages of growth. That is, the same type of model would be used each year, but the parameters derived from the data would relate to: (1) the current year, and (2) a given cut-off time within the growth period.

The role of the three parameters in the growth model can be described in terms of various phases of growth.

1. The initial phase or base weight is at t = 0. Since $\hat{\rho}$ (whatever its value) raised to the power t = 0 is 1, $\hat{Y}_0 = \frac{1}{\hat{\alpha} + \hat{\beta}}$ estimates the base weight.

2. The final phase or mature forecast weight of the dependent variable is the most important in terms of forecasting final corn yield and production. Assuming that $0<\hat{\rho}<1$, we see that the forecast harvest

weight is $\hat{Y}_m = \lim_{t\to\infty} \hat{Y}_t = \frac{1}{\hat{\alpha}}$. That is, for large values of t, $\hat{Y}_t$ depends primarily upon $\hat{\alpha}$. Therefore the parameter α is termed the primary parameter.

3. The intermediate phase of the growth period follows the initial phase and continues until maturity, when the large values of t are reached. The value of $\hat{\rho}$ reflects the rapidity of the weight increase from $\hat{Y}_0$ to $\hat{Y}_m$ as t increases. For $0<\rho<1$ the model is indeed a growth model and ρ can be termed the rate of growth parameter. If $\hat{\rho}$ is near zero the growth is very rapid. If $\hat{\rho}$ is near unity, growth at a gradual rate is indicated. The ratio of $\frac{\hat{Y}_m}{\hat{Y}_0} = 1 + \frac{\hat{\beta}}{\hat{\alpha}}$ determines the range of the Y_t scale.

The program utilized to derive the parameters from the 1973 data required approximate starting values. For the dependent variable dry weight per ear, the values were $\hat{\alpha} = .006$, $\hat{\beta} = .08$, and $\hat{\rho} = .87$. For Y_t = dry weight per kernel at time t, the values used were $\hat{\alpha} = 3.8$, $\hat{\beta} = 130$, and $\hat{\rho} = .87$.

Two Examples

Each set of parameter estimates defines a specific model at a given time. For example in 1973 for Y_t = dry weight (gms.) per ear at t days after silks begin to dry, we have the following parameter estimates for data sets available after various field visits.

	Estimates of model parameters based upon all data available after various field visits				
Parameter	IV	VI	VII	VIII	X
α	.0059597	.0061557	.0062934	.0063149	.0063487
β	.12777	.12930	.14616	.14869	.15380
ρ	.88271	.88108	.87514	.87428	.87267

Thus, the specific model based upon data obtained on field visits I-VII is

$$\hat{Y}_t = \frac{1}{.0062934 + .14616(.87514)^t} ,$$

where $\hat{Y}_t$ is the estimated dry weight (gms.) per ear at t days after silks began to dry. For $\hat{Y}_{t'}$ = estimated dry weight (gms.) per kernel at t′ days after silks emerged based upon data from visits I-VII the model is

$$\hat{Y}_{t'} = \frac{1}{3.8654 + 333.95(.87113)^{t'}} .$$

Numerical values of the dependent variables for various values of the two time variables are shown below for these two models.

Time after silks started to dry (t)	Estimated dry weight per ear ($\hat{Y}_t$)	Time after silks emerged (t′)	Estimated dry weight per kernel ($\hat{Y}_{t'}$)
(Days)	(Grams.)	(Days)	(Grams.)
0	6.56	0	.0030
10	22.32	10	.0114
20	60.82	20	.0400
30	111.52	30	.1088
40	142.90	40	.1921
50	154.34	50	.2380
60	157.67	60	.2531
70	158.57	70	.2573
80	158.81	80	.2583
90	158.87	90	.2586
100	158.90	100	.2587
110	158.90	110	.2587
120	158.90	120	.2587
∞	158.90	∞	.2587

Evaluation of Models

Two methods of evaluating the performance of the logistic growth model for various time variables and as data becomes available for later stages of growth are:

1. The magnitude and sign of the departure of the forecast from actual mean dry weight at maturity.

2. The magnitude of the relative standard deviation of the "primary" parameter, α.

Mean dry weight at maturity was estimated from a large sample of plants with mature ears. The mean was for the population of plants sampled from during the entire growth period for which the time variable in the model being evaluated was defined. That is, the model forecast and estimated mean weight make valid inferences about the same sub-population. The relative standard deviation is the estimated standard deviation divided by the estimate of the primary parameter $(\frac{\hat{\sigma}_{\hat{\alpha}}}{\hat{\alpha}})$.

For the two examples previously discussed, departures of the forecast from the actual mean dry weight and the relative standard deviations are shown below.

Dependent variable	Independent time variable	Data from visits	Departure of forecast from actual mean dry weight	Relative standard deviation of the primary parameter
			(Percent)	(Percent)
Dry weight of grain per ear (Y_t)	Days after silk starting to dry (t)	I(only)	(No convergence to model)	
		I & II	+22.0	35.46
		I - III	+0.7	6.96
		I - IV	+7.8	4.32
		I - V	+8.5	2.74
		I - VI	+4.4	1.74
		I - VII	+2.1	1.29
		I - VIII	+1.7	1.16
		I - IX	+2.2	1.02
		I - X	+1.2	0.92
Dry weight of grain per kernel ($Y_{t'}$)	Days after silk emerged (t')	I(only)	-89.6	16.07
		I & II	-71.4	12.09
		I - III	-37.3	10.10
		I - IV	+4.6	6.61
		I - V	-6.0	2.24
		I - VI	-1.8	1.59
		I - VII	+0.4	1.37
		I - VIII	+0.5	1.15
		I - IX	+0.9	0.98
		I - X	+1.2	0.86

Preliminary Conclusions

There is strong evidence that biological laws do relate corn growth, in terms of dry matter deposition, to time scales that are closely associated with the timing of kernel fertilization. Conclusions based upon this preliminary study are limited by two primary factors:

1. The population sampled was small relative to the population of potential application.

2. Only the type of model discussed in this report was used to evaluate the various time scales.

Time after silk emergence and time after the three stages of silk drying show a strong relationship with the amount of dry weight in kernels. Among the other time variables considered, time after tassel emergence shows some relationship, but appears to have a less consistent relationship than do the time scales based upon silk events. Perhaps the elapsed time after various silk events are more closely associated with the development of the individual plant's primary ear. The so-called "time" variable, dry matter percentage of grain when sampled, did not demonstrate a consistent relationship to either dry weight of grain per ear or per kernel in this study. The fact that the time variable and dry weight are observed as a result of the same laboratory procedures may distort any real relationship.

Many models used in making objective forecasts of yield components rely on the invariance of model parameters over years. That is, parameters developed based upon the relationship of early season observations to variables observed at maturity during a base period of years are assumed to apply to the current year. This is equivalent to an assumption that, so far as the parameters are concerned, the current year is a random sample from the same "super population" from which each of the base period years are sampled. This assumption is quite useful for years that are fairly typical, but can cause difficulties in atypical years. Within year time related growth models have the advantage of providing the forecast, a measure of its precision and estimates of the model parameters which are independent of data from other years. The model discussed in this report has the highly desired statistical property of producing a point estimate and an estimate the error associated with that estimate from the same set of data.