ADJUSTING SURVEY ESTIMATES OF GRAIN YIELD WITH WEATHER VARIABLES

Arlin M. Feyerherm, Department of Statistics, Kansas State University, Manhattan, KS 66506

KEY WORDS: Crop estimates, weather-yield models

ABSTRACT

The objective of this research was to increase precision of preharvest yield estimates of winter wheat by using weather variables as covariates. Yield per unit area is a product of number of heads and average weight per head. Present methods to make preharvest estimates include counting heads and/or estimating number of heads from stalk counts for the density component and using five-year historical averages and/or information on number of fertile spikelets/head or grains/head for the weight component. A model was developed for wheat in Kansas that related average weight/head to a function of weather variables, time (as a surrogate for technology), and the estimated number of heads. In an eight-year test of preharvest forecasts of weight/head, the model showed reductions in the root-mean-square-error of 30% on May 1 and 19% on June 1 over present methods.

1. INTRODUCTION

The National Agricultural Statistics Service (NASS) of the U. S. Department of Agriculture uses sample survey methods to select small plots within fields for its objective yield indicator of mean yield per unit area for certain crops (U. S. Department of Agriculture, 1983). For winter wheat (planted in the fall), randomly selected plots are laid out in April to take subsequent counts (number of stalks, number of heads) and measurements (weight per head, weight of grain per head) as the season progresses. After the wheat has matured, a net yield per acre (Y_i) for the ith plot is calculated as

$$Y_i = H_i W_i C_i - L_i; \quad i = 1,2,\dots,n, \tag{1}$$

where H_i = number of heads, W_i = grain weight/head, C_i = an appropriate constant to transform the weight per plot to a bushel per acre basis, and
L_i = harvest loss (Thiessen, 1989).

It is the practice of NASS to issue preharvest state-by-state forecasts of production on a monthly basis. For winter wheat, these begin in May. At that time, a count of the number of stalks is possible, and heads may be visible in some plots, so H_i in (1) may be estimated. However, the weight of grain/head (W_i) is unavailable. A wheat objective yield (WOY) forecast is prepared by substituting the five-year historical average of statewide mean weights/head for W_i in (1). The main contributors to yearly variation in statewide means are technology, weather, and pests. In this paper, a model is developed to measure the influence of technology and weather on the postharvest, statewide, mean weight/head in Kansas. The modeled estimate then is compared with preharvest forecasts made using present methods.

2. DATA AND METHODS

To study the influence of weather on wheat yields, one must use available observational data. For this study, annual wheat yields were available for the nine agricultural statistics divisions (ASD's) in Kansas for the period 1950 to 1988, together with daily weather data (temperatures and precipitation) at two weather stations per ASD, to represent weather effects on ASD yields. Thus, 351 (9 ASD's times 39 years) vectors of variables were available to fit a model of the form

$$Y_{st} = \sum_{i=1}^{k} \beta_i X_{ist} + \sum_{r=1}^{9} \gamma_r Z_r + \sum_{r=1}^{9} \delta_r Z_r T + \epsilon_{st};$$
$$s = 1,2,\dots,9;\; t = 1950,1951,\dots,1988;$$
$$T = (t - 1950); \tag{2}$$

Y_{st} = NASS estimate of yield for the sth ASD in year (t),
X_{ist} = value of ith weather variable for the sth ASD in year (t),
$Z_r = 1$ if $r = s$; $Z_r = 0$ if $r \neq s$,
ϵ_{st} = random error for the sth ASD in year (t),
β_i $(i=1,2,\dots,k)$, γ_r and δ_r $(r = 1,2,\dots,9)$
were parameters to be estimated. In (2), the first summation is called the weather-yield function (WYF), where for each i, s, and t an X_{ist} is a departure from a 39-year mean, the second summation is associated with different average yield levels in each ASD in 1950 (T = 0), and the third summation measures linear trends in yield since 1950 associated with technology.

A smaller data set was available to estimate parameters to relate statewide average weight/head ($\overline{W}_t$) in year (t) to weather, technology, and head counts. Here the model was

$$\overline{W}_t = \alpha_0 + \alpha_1 T + \alpha_2(\text{WYF}_t) + \alpha_3(\overline{H}_t - \hat{\overline{H}}_t) + \alpha_4(\text{FREEZE}) + \epsilon_t, \tag{3}$$

$t = 1964,1965,\dots,1988$; $T = (t - 1964)$;
$\overline{W}_t$ = NASS estimate of statewide mean weight/head (in grams) at harvest calculated over plots for year (t),
WYF_t = statewide weighted average of WYF values over ASD's in year (t) using harvested acres per ASD (1971-1975) to develop weights,
$\overline{H}_t - \hat{\overline{H}}_t$ = departure of statewide average head count per plot at harvest in year (t) from a trend value for average head count in year (t),
FREEZE = 0; but if t = 1966 or t = 1981 then FREEZE = 1,
ϵ_t = random error for year (t),
and the parameters α_0 = average statewide weight/head in 1964 (if all other factors had zero influence), α_1 = yearly increase in head weights from technology, α_2 = impact of change of one unit of WYF on $\overline{W}$, α_3 = change in weight/head for unit change in head count, and α_4 = impact of very late freezes (May 13 in 1966 and May 11 in 1981). The model assumed not only influences of technology (improved varieties, increased nitrogen, better management) and weather but also the ability of wheat to produce heavier weight of grain per head when fewer heads were present and vice versa.

Table 1. Estimates of α's, RMSE's, and R^2's as Additional Terms Were Added to Model in (3).

Growth Stages Included In WYF (see (4))	Estimates of α's						
	Intercept	T	WYF	$\bar{H}_t - \hat{H}_t$	FREEZE	RMSE	R^2
None	0.44	0.007				0.055	0.47
E	0.45	0.005	0.010			0.048	0.61
E,D,V	0.45	0.005	0.009	-0.0010		0.041	0.72
E,D,V,R	0.45	0.005	0.013	-0.0011		0.037	0.77
E,D,V,R	0.45	0.005	0.011	-0.0010	-0.063	0.034	0.82
E,D,V,R,F	0.45	0.005	0.011	-0.0011	-0.065	0.032	0.84

3. RESULTS

3.1 A Weather-Yield Function

Based on the model in (2), values of the weather-yield function for a given year and ASD were estimated by:

$$\begin{aligned} WYF = & \{ -0.413\ TX(E) + 2.95\ P(E) - 0.171\ [P(E)]^2 \} + \\ & \{ -0.274\ TN(D) + 1.24\ P(D) \} + \\ & \{ -0.229\ TN(V) + 0.045\ [AP-30][TX(V) - TN(V)] \\ & -0.082\ [CP(V)][P(V)] + 0.161\ [P(V)]^2 \} + \\ & \{ -0.396\ TN(R) + [3.18 - 0.258\ CP(R)]P(R) \\ & - [0.304 - 0.0228\ CP(R)][P(R)]^2 \} + \\ & \{ -0.261\ TN(F) - 0.003\ [CP(F)][P(F)]^2 \} \end{aligned} \quad (4)$$

where the letters in parentheses ( ) denote the approximate stage of development, i.e., E = a 60-day period in the fall when the crop is being established, D = time span of dormancy (variable by ASD), V = 40-day period of vegetative growth, R = 40-day period of reproductive activity from jointing to heading, and F = 20-day period of grain-fill. The weather variables for the different growth states were TX = avg. daily maximum temperatures in °F, TN = avg. daily minimum temperature °F, P = total precipitation in inches, CP = accumulated precipitation to the beginning of a stage, and AP = long-term avg. annual precipitation for the sth ASD (s = 1,2,...,9). All terms in (4) were departures from 39-year means, and the braces were used to delineate terms associated with different growth stages. All coefficients were significantly different from zero at a $p < .01$ level, with one exception when $p = .04$. Insight into development of a WYF was gained from previous work (Feyerherm and Paulsen, 1981, 1986).

The unit for WYF was bushels/acre and (4) represents a transformation of heat and moisture units into yield per unit area measured in bushels per acre. Agronomically, the terms in (4) indicated that throughout the season, winter wheat grain yields were increased when temperatures decreased. During the vegetative stage, yields were larger in arid areas, (AP-30) < 0, in years when the average daily range in temperature [TX(V) - TN(V)] was small. An opposite effect occurred in humid regions, (AP-30) > 0. The impact of an inch of precipitation varied during the season and was positive or negative depending on growth stage and prior cumulative precipitation; the coefficients of precipitation were functions of CP values in the vegetative, reproductive, and grain-fill stages.

3.2 A Model for Weight/Head

Estimation of the parameters (α_i's) in (3) using n = 25 years of Kansas data produced the following relation:

$$\hat{W}_t = 0.452 + 0.0052\ T + 0.011\ WYF_t - 0.0011\ (\bar{H}_t - \hat{H}_t) - 0.065\ FREEZE . \quad (5)$$

All coefficients were significantly different from zero at the $p < .01$ level. In addition, data on statewide average head counts when regressed against time and some weather variables produced $\hat{H}_t = 361 + 1.0\ T$ for substitution in (5). Results in Table 1 demonstrate the increase in precision of estimates of weight/head that can be expected as the season progresses from planting to harvest. Increased precision was obtained by including more terms from (4), by adding estimated head counts into the model, and by introducing the indicator variable (FREEZE) when a late-season freeze occurred.

3.3 Precision of Model vs. Present Methods

Eq. (5) represents an alternative to present methods of making preharvest (May 1 and June 1) estimates of W_i in (1). To compare use of the model in (3) with present methods, years 1981 through 1988 were used for testing, and data from 1964 up to a test year (e.g., 1964 through 1980 when 1981 was the test year) were used to re-estimate the coefficients in (3). All terms in (4) and head count at harvest were used for model derivation. However, in a test year, terms in (4) through the vegetative (V) stage only and through the reproductive (R) stage only were evaluated for WYF for May 1 and June 1 estimates, respectively. Similarly, May 1 and June 1 NASS-estimates of head counts were substituted for H_t. The root-mean-square errors of differences between model and at-harvest NASS values of $\bar{W}_t$ and between preharvest (May 1 and June 1) NASS estimates and at-harvest NASS values are shown in Table 2.

Table 2. Root-Mean-Square-Errors for Modeled and NASS Estimates of Weight/Head over Eight Test Years.

Estimates for:	Model	NASS
May 1	0.049	0.070
June 1	0.043	0.053

The results suggest that use of the model in (3), updated yearly, would give more precision than present methods to May 1 and June 1 preharvest estimates for the statewide mean weight/head ($\bar{W}_t$).

REFERENCES

Feyerherm, A. M. and Paulsen, G. M. (1981), "Development of a Wheat Yield Prediction Model," Agronomy Journal, 73, 277-282.

____________ (1986), "Development of a Weather-Yield Function for Winter Wheat," Agronomy Journal, 78, 1012-1017.

Thiessen, E. J. (1989), "Kansas Wheat Objective Survey," Proceedings of Conference on Applied Statistics in Agriculture, May 1-2, 1989, Manhattan, KS.

U. S. Department of Agriculture (1983), "Scope and Methods of the Statistical Reporting Service," USDA Statistical Reporting Service, Misc. Publication Number 1308. U.S. Government Printing Office, Washington, D. C.

ACKNOWLEDGEMENT

The author appreciates the review of this paper by Kenneth E. Kemp, Department of Statistics, Kansas State University and the cooperation of the Kansas Agricultural Statistics Office in providing data for this study.