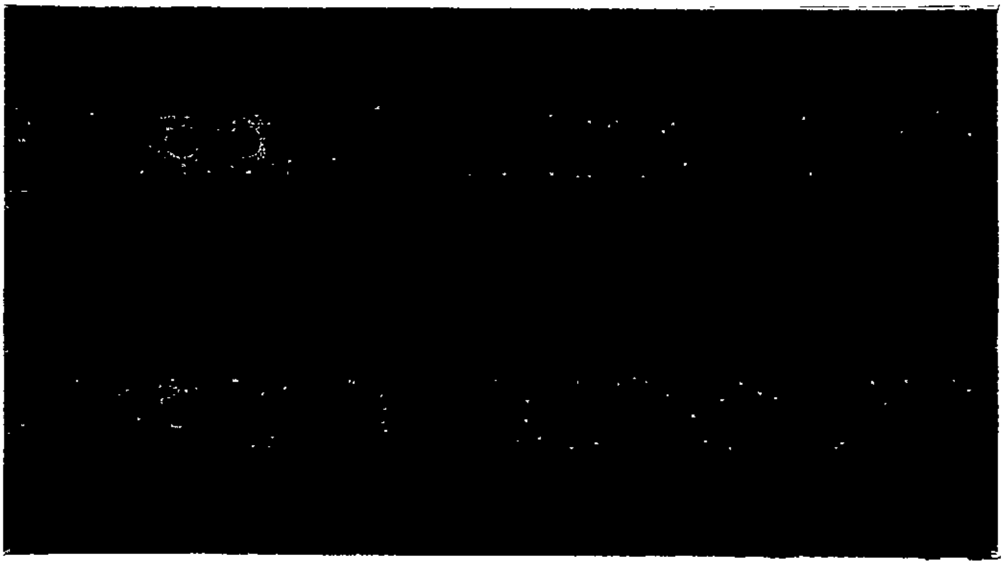
7.8-10013

NASA CR-151525



(E78-10013) PLANTING DATA AND WHEAT YIELD MODELS Final Report, 15 Feb. 1975 - 31 Mar. 1977 (Kansas State Univ.) 89 p
HC A05/MF A01 CSCL 02C

N78-10538

Unclas
G3/43 00013

KSU

Kansas State University

Manhattan, Kansas

PLANTING DATE AND WHEAT YIELD MODELS

CONTRACT NAS9-14533

FINAL REPORT FOR PERIOD 2/75 to 3/77

September, 1977

Prepared for: National Aeronautics and Space Administration
Johnson Space Center
Houston, Texas

Prepared by: Arlin M. Feyerherm
Principal Investigator
Department of Statistics
Kansas State University
Manhattan, Kansas 66506

Support Personnel:

Dr. Gary Paulsen
Ms. Jeanne Sebaugh
Mr. Dale Fjell
Mr. Michael Franzblau
Mr. John Olsowski
Mr. Michael Frerichs
Ms. Paulette Johnson

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
Final Report		

4. Title and Subtitle

Planting Date and Wheat Yield Models

5. Report Date

September 1, 1977

6. Performing Organization Code

7. Author(s)

Arlin M. Feyerherm

8. Performing Organization Report No.

9. Performing Organization Name and Address

Department of Statistics
Kansas State University, Calvin Hall
Manhattan, Kansas 66506

10. Work Unit No.

11. Contract or Grant No.

NAS 9-14533

12. Sponsoring Agency Name and Address

MASA/JSC

13. Type of Report and Period Covered

Final (2/15/75 to 3/31/77)

14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

New research products described in this report include: (1) a starter (time-of-planting) model for spring-planted wheat based on daily temperatures, (2) a model to estimate daily temperature extremes from a sample of eight readings taken at three-hour intervals, (3) a revised spring wheat yield model. Rust-induced losses were incorporated into our winter wheat yield model and the increase in precision for estimating Kansas yields was calculated. Amount of increased precision resulting from a six-fold increase in weather-station density was calculated for Kansas. Attempts to construct a precipitation-based variable-date starter model for fall-planted wheat, that would be an improvement on a fixed-date model, were unsuccessful.

17. Key Words (Suggested by Author(s))	18. Distribution Statement
Planting date, wheat yield models, daily extreme temperatures	

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price*

*For sale by the National Technical Information Service, Springfield, Virginia 22151

NASA — JSC

JSC Form 1424 (rev. July 74)

PREFACE

The continuing objective of our LACIE contract work is to develop universal wheat yield models for fall and spring-planted wheat which show response of grain yield to both weather and cultural practices and can be applied on a global basis.

This report covers work on development of starter (time-of-planting) models, estimation of daily temperature extremes from trihourly (every three hours) observations, development of a revised spring wheat yield model, investigation of the effect of increased density of weather stations on precision of yield estimates, estimation of increase in precision of yield estimates with knowledge of rust losses, and application of our winter wheat yield model to an oblast in the Ukraine area of the USSR.

The major conclusions arising from this effort were:

a. A variable-date starter model for spring wheat depending on temperature, is more precise than a fixed-date model. We could not reach the same conclusions for fall-planted wheat.

b. If the largest and smallest of eight temperatures are used to estimate daily maximum and minimum temperatures; respectively, a 1-4 °F bias will be introduced into these extremes. Some of the bias can be eliminated by using formulas developed for that purpose.

c. For our revised spring wheat yield model, regional yields should be related to WAC (a weather and cultural practices index) with a two-parameter equation; that is,

$$\hat{Y} = a + b*WAC$$

rather that the single-parameter (MAP) factor used for our winter wheat model.

d. For Kansas, a reduction of 0.5 bushels/acre in the RMSE (root-mean-square-error) between model and SRS yields was achieved by a six-fold increase (7 to 42) in the density of weather stations. An additional reduction of 0.3 b/A was achieved by incorporating losses due to rusts in the model.

TABLE OF CONTENTS

LIST OF TABLES

Table		Page

Table		Page

1.0 INTRODUCTION

This final report on Contract NAS 9-14533 covers the period February 15, 1975 to March 31, 1977. Some of the work performed under this contract was integrated into the final report on Contract NAS 9-14282, dated February, 1977, and titled "Response of Winter and Spring Wheat Grain Yields to Meteorological Variation."

Under both contracts, tasks were interrelated and pointed toward the common goal of development of universal wheat yield models for fall and spring-planted wheats. Specific goals included: (1) to quantify the effects of both meteorological variation and changing cultural practices on wheat yields and, (2) to be applicable on a global basis for both forecasting and estimating yields. Rather than repeat certain definitions and notations, it will be assumed that the final report on Contract NAS-14282 is available to the reader.

The principal products of our work, over and above those previously reported, were:

a. A starter (time-of-planting) model for spring-planted wheats based on daily temperatures.

b. An investigation of the relationship of planting dates for fall-planted wheats to temperatures and precipitation for which results were negative.

c. A model to estimate daily minimum and maximum temperatures from the eight daily readings taken at three-hourly intervals at many reporting stations on a world-wide basis.

d. A spring-planted wheat yield model (a revised form of that which appears in the final report for Contract NAS 9-14282) with application to North Dakota and two oblasts in the USSR.

e. Application of our winter wheat yield model to an oblast within the Ukraine area of the USSR.

f. An investigation of the effect of increased density of weather reporting stations on precision of yield estimates.

In sections to follow, we will give a summary of the results and use appendices for more detailed reporting on data sets, methodology and procedures.

2.0 SPRING WHEAT STARTER MODEL

Assuming historical data are available on percents of wheat planted by given dates, the simplest model to estimate the date (P50) when fifty percent of wheat acreage has been planted would be to use the mean value of P50's. Such a fixed-date model would not allow for season-to-season variation. A potential cause of yearly variation in planting is weather variation and we have chosen to measure this variation by accumulated values of what we call warming/planting (W/P) days.

A W/P day is defined as follows:

$$W/P = 0, \text{ if } TA < 32°F,$$
$$= 0.1\ (TA-32), \text{ if } 32°F \leq TA < 42°F,$$
$$= 1, \text{ if } TA \geq 42°$$

where

TA = average daily temperature.

To measure yearly variation, one can either, (1) accumulate W/P days between two fixed dates and record the accumulated sum or, (2) accumulate W/P days from a fixed date till the sum reaches a specified level (K) and record the date of occurrence. We used the latter method to estimate P50.

A summary of our findings follows with more details available in Appendix A.

2.1 Model development. Thirty-five data points were used to establish a value of K = 36. The Julian day EP50 (estimated P50) is the first day when

$$\sum_{J=19}^{EP50} (W/P)_J \geq 35.5 \quad .$$

Summation begins at J = 19 (January 19). Choice of January 19 for a beginning date was somewhat arbitrary but coincides roughly with the coldest time of the year in the northern hemisphere.

The value K = 36 warming/planting days was the average of

$$\sum_{J=19}^{P50} (W/P)_J$$

calculated over 35 location-years of data where (W/P) days were accumulated at one or two weather stations, and P50 values were interpolated from CRD data either for a single CRD or an average of two. The 35 data points were generated in the years 1967-73 over six regions in North and South Dakota.

2.2 Some test results. The variable-date starter model was applied to two sets of independent data. One set consisted of P50 dates from Montana for the period 1969-73, the other applied to P50 dates for CRD's in North Dakota for 1974 and 1975. The results are shown in Table 2.1 which compares use of the fixed and variable-date models. The variable-date is superior to the fixed-date model for these locations. Planting in North Dakota was late in both 1974 and 1975. In 1974, the model indicated a "normal year" but wet weather (not a contributing factor in our model) in May delayed planting. Both fixed and variable models underestimated the P50 dates by about 20 days in the northern CRD's and accounted for the large RMSE values. In 1975, use of the variable-date model was a considerable improvement over the fixed date model.

Table 2.1 Comparison of RMSE (root-mean-square-error) values (in days) for two models to estimate P50 values for n location-years.

	n	Fixed date†	Variable date
Montana	10	9.2	6.5
No. Dak.	12	17.1	13.1

†For Montana the fixed date was the mean P50 for the same years for which the RMSE was calculated. For No. Dak., the mean was for the years 1967-73 while the RMSE was based on 1974-75 data.

2.3 Application to USSR climates. The variable-date starter model was applied to daily temperature data at three different weather stations in the spring wheat area of the USSR. Results are shown in Table 2.2 for estimated P50 values. One can only judge that the results "look reasonable" since no ground truth data were available. At the least, the results indicate that April and May were cold enough in 1966, '69, and '72 to discourage early planting while 1974 was particularly warm in early spring.

Table 2.2 Estimates of P50 dates for three weather station locations in the USSR.

	Years										
Locations	'65	'66	'67	'68	'69	'70	'71	'72	'73	'74	'75
Kurgan	----	5/24	5/11	5/21	5/28	5/16	5/23	5/23	----	5/11	5/11
Atbasar	5/20	5/28	5/20	5/21	5/25	5/18	5/23	5/26	5/17	5/12	5/18
Tselinograd	5/2	5/23	5/18	5/19	5/22	5/19	5/20	5/24	5/13	5/13	5/16

2.4 Discussion. The variable-date starter model has the advantage that it can be applied in new areas without benefit of prior weather data to estimate parameters (e.g. mean dates for P50) assuming K = 36 is representative of all spring-planted wheat areas of the globe. This hypothesis should be tested in other areas but historical ground truth on P50 dates would be needed to determine how much K should vary from region to region.

Other models were tested along with the one described in Section 2.1. One of the alternates modified values of $(W/P)_J$ ($0 \leq (W/P)_J \leq 1$, J = day number) by amount of precipitation on day J; another by amount of precipitation on day J and the three preceding days. Neither of these alternate definitions reduced the RMSE value for Montana test data by an appreciable amount over that achieved with the temperature only. Another alternate used the line segment from TA = 32°F to 52°F to define $0 \leq W/P \leq 1$ but no advantage accrued.

Choice of the period 1967-73 for estimating K was based on a finding that the recorded data on planting dates for 1953-73 suggested that a shift toward later planting occurred over the two decades. A number of explanations are suggested in Appendix A. Whatever the cause(s), the time period 1967-73 seemed most appropriate for estimating K as 36 W/P days.

When applying the proposed variable-date model, the larger "misses" may be associated with either, (1) a year with very warm temperatures in late winter and early spring (central Montana, 1972) and farmers spend additional time in preparing the land to control weeds and/or wait for rain, or, (2) continued wet weather when planting must be postponed (North Dakota, 1974). In the latter situation, soil temperatures may be high enough but rains keep the farmers out of the field. More work is needed on the variable-date model to avoid "large misses" when they are weather dependent.

3.0 WINTER WHEAT STARTER MODEL

An intensive investigation was carried out to determine if time of planting winter wheat could be related to weather events prior to and during the period when planting occurred. Unlike spring planting, which is dependent on early spring temperatures, yearly variation in fall planting would seem more dependent on precipitation. Extremely wet or extremely dry weather in September/October could mean delays in planting. However, the problem is to find mathematical functions which enable one to quantify the amount of delay caused by these extreme conditions.

Variable-date starter models to estimate P15, P50, and P85 (15, 50, 85% planted) dates were developed using historical daily weather and percent planted data for crop reporting districts in Kansas. A workday ($0 \leq W_J \leq 1$, J = Julian day) was defined as a function of precipitation both on day J and prior to day J. The functions were all of the form of segmented lines.

Different variable-date models were created by changing parameter values which varied the slope and join-points of the segmented lines. However, none of the models reduced the RMSE (root mean square error) by an appreciable amount over that for a fixed-date model when testing results using the same data base from which the model was developed.

More details on this investigation are given in Appendix B. Prospects for improving on a fixed-date starter model to explain year-to-year variation due to weather look dim; especially if a universal model is needed.

4.0 ESTIMATING DAILY MINIMUM AND MAXIMUM TEMPERATURES FROM TRIHOURLY REPORTS

4.1 The need. Our yield models require daily minimum (TN) and maximum (TX) temperatures as inputs both directly and indirectly. Indirectly, they are used to generate daily increments of development in the crop calendar, and to estimate daily potential evapotranspiration. Directly, their mean values, between crop stages, are major factors in estimating yields.

The most accessible daily temperature data on a global basis are reports of temperatures taken on a trihourly (every three hours) schedule by the WMO network. The need for a computer algorithm to estimate TN and TX from the eight (or less) temperature readings available on a given day, was apparent both for model testing on historic data and for real-time yield forecasting and estimation.

4.2 <u>Definitions, assumptions, options.</u> Two options were considered to estimate TN and TX from trihourly readings.

a. Option 1: Determine the smallest (largest) of the trihourly readings and subtract (add) a predetermined number of degrees to estimate TN (TX).

b. Option 2: Select a specified trihourly temperature between midnight and 0800 and subtract a predetermined number of degrees to estimate TN. Likewise select a specific trihourly temperature between 0900 and 1700 hours and add a predetermined number of degrees to estimate TX.

<u>A combination of the two options can be used in practice and we recommend that, on a given day, option 1 be used if all eight trihourly readings are available and option 2 used if one or more temperatures are missing but at least one hourly temperature is present in each required time span to estimate TN and TX, respectively.</u>

Each predetermined value referred to in Options 1 and 2 is an estimate of one of the following daily differences (DD):

(4.1) $DDPTN(h_o) = PTN(h_o) - TN,\ h_o = 00,\ 01,\ 02;$

(4.2) $DDPTX(h_o) = TX - PTX(h_o),\ h_o = 00,\ 01,\ 02;$

(4.3) $DDTN(h) = T(h) - TN,\ h = 00,\ 01,\ \ldots,\ 08;$

(4.4) $DDTX(h) = TX - T(h),\ h = 09,\ 10,\ \ldots,\ 17;$

where

$PTN(h_o)$ = smallest of eight trihourly readings (a psuedo TN) taken at hours $h_o + 3i$ (h_o = 00, 01, 02; $i = 0, 1, \ldots, 7$),

$PTX(h_o)$ = largest of eight trihourly readings (a psuedo TX) taken at hours $h_o + 3i$ (h_o = 00, 01, 02; $i = 0, 1, \ldots, 7$),

$T(h)$ = temperature at clock hour h (h = 00, 01, ..., 17).

At a given location, h_o is unique and there is only one DD of interest in (4.1) and (4.2). There are three potential DD's of interest in (4.3) occurring at either (00, 03, 06) or (01, 04, 07) or (02, 05, 08) depending on h_o. There are also three DD's for (4.4).

To model estimates of TN [$\hat{TN} = PTN(h_o) - DDPTN$ or $\hat{TN} = T(h) - DDTN(h)$] and likewise for TX, we assumed that the only data available on a particular day would be the eight or less trihourly readings. Factors such as amount of cloud cover, passage of frontal systems, or any auxiliary meteorological data that might affect the size of a particular DD were assumed unknown. The best estimates of DD's for a particular day at a particular location would be long-term average daily difference (ADD's) for that Julian day and location. The modeling effort involved calculating ADD's for each day of the year at selected locations and relating the estimates to the following factors:

a. difference between sun (solar) time and clock time at the respective locations,

b. average daily temperature range,

c. daylength,

d. daily observation schedule (depending on h_o).

For DDTN(h) and DDTX(h), an additional factor is:

e. clock time (h) when temperature is recorded.

4.3 <u>Data set for model development</u>. Data used to estimate parameters in our model consisted of <u>hourly temperature</u> readings and <u>daily minimum and maximum temperatures</u>, over a 15-year period, at the following Kansas locations:

City	Longitude	Latitude	Elevations
Chanute	95° 29'	37° 40'	977
Russell	98° 51'	38° 54'	1834
Goodland	101° 42'	39° 22'	3651

For each location, daily values were computed for the entries in Equations 4.1 through 4.4 and then averaged for each Julian day over the 15-year period. As a further step toward removing "noise" from our estimates of ADD's, all DD's from a given month were averaged over both days within a month and years. Thus estimates of ADD's were based on approximately N = 450 (30 x 15) observations.

4.4 Statistical considerations. Means and standard deviations for DDPTN(00) and DDPTX(00) were computed over the 15 years of data for each Julian day at each of the three Kansas locations. A sample of the results appear in Table 4.1. A visual examination of Table 4.1 indicates:

a. That values of $\bar{X}$ tend to be smaller for TX then for TN and may vary systematically over locations and time of the year.

b. That standard deviations are approximately equal to the means.

Point (b) suggests that we approximate the distribution of DD's by a one-parameter probability distribution. The simplest continuous probability distribution, which has zero density for negative values and mean equal to standard deviation is the exponential (or negative exponential), a member of the Gamma distribution family. The density function is given by:

$$f(x) = \beta e^{-\beta x}, \; x \geq 0;$$
$$= 0, \; x < 0;$$

where

Table Sample means ($\bar{X}$) and standard deviations (s) of daily differences between psuedo maximum (minimum) and actual maximum (minimum) temperatures in °F (N = 15 years, h_o = 00).

	DDPTX(00) = [TX-PTX(00)]						DDPTN(00) = PTN(00) - TN					
	Chanute		Russell		Goodland		Chanute		Russell		Goodland	
Julian Day	$\bar{X}$	s	$\bar{X}$	s	$\bar{X}$	s	$\bar{X}$	s	$\bar{X}$	s	$\bar{X}$	s
1	1.4	1.6	1.2	0.8	1.3	1.0	1.8	1.6	3.1	2.3	2.9	1.8
15	0.5	0.6	1.3	1.4	2.0	1.7	4.7	3.7	3.9	3.8	3.5	2.2
30	0.7	0.7	1.0	1.0	1.7	1.8	2.2	2.4	3.0	2.3	2.3	1.4
45	0.6	0.8	1.5	1.2	1.2	1.1	2.5	2.6	3.3	2.2	3.5	3.9
60	0.8	0.7	1.1	1.0	1.8	1.3	1.8	2.3	2.8	1.7	2.6	2.2
75	1.1	1.0	1.2	1.3	1.3	1.3	1.6	1.6	2.3	2.6	2.2	1.6
90	1.4	0.9	1.1	1.2	1.6	1.1	1.7	1.7	3.0	4.0	3.4	2.5
105	1.0	0.8	1.3	1.1	1.8	1.3	3.6	4.5	2.8	3.4	4.3	4.4
120	1.7	1.2	1.2	1.0	2.2	1.9	2.1	2.6	3.0	4.1	3.4	3.9
135	1.5	1.2	1.7	1.2	2.1	1.4	1.7	1.0	2.5	2.8	2.8	2.0
150	0.9	0.9	1.4	1.1	2.5	1.6	1.5	0.7	2.4	1.7	1.9	1.2
165	1.7	1.4	1.3	1.1	1.8	1.4	2.3	1.7	2.1	1.4	3.0	2.8
180	1.3	0.8	1.5	1.2	1.2	0.8	1.9	1.0	2.1	1.7	2.7	2.1
195	1.5	0.9	1.6	1.0	1.8	1.2	1.3	1.0	1.8	1.5	3.2	2.6
210	1.5	1.0	0.9	1.6	2.4	1.6	1.6	1.0	1.4	1.2	2.8	1.4
225	0.9	1.3	1.0	1.8	2.2	1.2	1.9	1.3	1.4	1.4	2.3	1.4
240	1.1	1.0	1.6	1.2	1.6	0.9	1.7	1.7	1.2	0.9	2.7	1.5
255	1.4	1.2	1.0	0.8	1.4	1.0	2.8	4.0	2.3	1.7	3.9	3.3
270	1.3	1.0	1.3	1.0	1.8	1.1	1.6	1.4	2.0	1.9	2.0	1.7
285	1.2	0.9	1.5	2.3	1.4	1.2	1.3	1.0	1.7	2.5	2.2	1.2
300	1.0	1.3	1.1	1.4	1.4	1.2	2.1	2.8	2.3	2.4	3.1	2.5
315	0.7	0.8	1.2	1.3	1.9	1.5	2.8	2.2	3.1	3.2	3.4	2.9
330	0.7	0.6	1.2	1.4	1.3	0.9	2.8	3.3	4.0	3.2	3.9	3.3
345	1.0	1.1	1.0	1.0	1.7	1.6	4.9	5.9	4.2	3.2	4.3	2.8

$$\text{mean} = \mu = 1/\beta, \text{ variance} = \sigma^2 = 1/\beta^2.$$

For the exponential distribution, 63% of the observations are in the interval 0 to μ, 95% in the interval 0 to 3μ, and 99% in the interval 0 to 5μ. Here, μ = ADD for our work.

Another look at means of daily differences, involving psuedo maximum and minimums, is given by Table 4.2. Means are calculated over both years and days within months so that $N \cong 450$ and some of the "noise" has been removed. If we assume that day-to-day differences in the DD's are statistically independent (year-to-year certainly would be) then we are roughly 95% confident that the values in Table 4.2 are within $\pm$ 2 standard errors [s.e.($\bar{x}$)] of their respective ADD's. Assuming $\sigma_{DD} = \mu_{DD}$, and that changes in ADD's over a half-month are relatively small, we would estimate

$$(4.5) \qquad \pm 2 \text{ s.e.}(\bar{x}) = \pm 2\sigma_{DD}/\sqrt{N} \cong \pm 2\mu_{DD}/\sqrt{450} \cong \pm .09\ \bar{x}.$$

For a first estimate of ADD's at any location we might use some grand averages of entries in Table 4.2. However, the size of the standard errors for the entries in Table 4.2 suggest that these $\bar{x}$'s are estimating different values of ADD's. By modeling, we should account for part of the observed variation.

Next, consider DD's defined by Equations 4.3 and 4.4. Means and standard deviations over both years and days within months are shown in Tables C.1-C.3 of Appendix C. For $\bar{X}$'s less than about 3°, the assumption that $\sigma_{DD} = \mu_{DD}$ appears quite tenable but as the $\bar{X}$'s increase the standard deviations increase at a slower rate. This simply means that we have increasingly conservative estimates of standard errors if we apply Equation 4.5 to increasingly larger mean values. Estimates of ADD's shown in Table

Table 4.2 Sample means of daily differences between psuedo maximum (minimum) and actual maximum (minimum) temperatures (in °F) for three observational schedules (h_o = 00, 01, 02). N $\simeq$ 450

	$DDPTX(h_o) = TX - PTX(00)$									$DDPTN(h_o) = TN - PTN(h_o)$								
	Chanute			Russell			Goodland			Chanute			Russell			Goodland		
Month	00	01	02	00	01	02	00	01	02	00	01	02	00	01	02	00	01	02
Jan	0.8	1.2	1.3	1.1	1.3	1.8	1.6	1.9	1.7	3.0	1.7	1.4	3.2	1.9	1.6	3.3	2.3	2.5
Feb	0.9	1.1	1.5	1.4	1.2	1.6	1.6	1.9	1.8	2.5	1.4	1.3	3.0	1.8	1.7	3.2	2.3	2.3
Mar	1.3	1.2	1.5	1.4	1.3	1.6	1.7	2.0	2.1	2.3	1.4	1.3	2.6	1.7	1.8	2.9	2.5	2.2
Apr	1.2	1.2	1.4	1.4	1.3	1.4	1.8	2.0	2.1	2.2	1.8	1.3	2.4	2.2	1.7	3.1	2.4	2.0
May	1.2	1.4	1.5	1.4	1.4	1.4	2.0	2.0	2.1	2.0	1.6	0.8	2.2	2.0	1.4	2.9	2.0	1.8
Jun	1.5	1.5	1.5	1.4	1.4	1.4	1.8	1.9	1.9	1.7	1.5	1.0	2.0	2.0	1.4	2.8	2.0	2.1
Jul	1.5	1.5	1.6	1.3	1.5	1.4	1.9	2.0	2.0	1.6	1.6	1.1	1.7	2.0	1.3	2.7	2.0	2.0
Aug	1.5	1.4	1.6	1.4	1.4	1.5	1.8	1.9	2.0	2.0	1.7	1.3	1.6	2.1	1.6	2.8	2.5	1.9
Sep	1.2	1.2	1.4	1.2	1.2	1.4	1.7	1.8	1.9	2.5	2.0	1.5	2.1	2.1	1.8	3.2	2.9	2.1
Oct	1.0	1.2	1.2	1.5	1.2	1.3	1.6	1.8	1.7	2.6	1.6	1.6	2.6	1.9	2.0	3.2	2.9	2.4
Nov	0.9	1.3	1.1	0.9	1.3	1.5	1.4	1.8	1.5	3.2	1.4	1.5	2.9	1.6	1.8	3.3	2.5	2.5
Dec	0.9	1.1	1.2	1.0	1.4	1.6	1.6	1.8	1.6	3.0	1.4	1.3	3.0	1.9	1.7	3.6	2.6	2.7

C.1-C.3 suggest that in addition to variation with hour of the day ADD's vary among locations and within years. One of the tasks in modeling will be to quantify the effects of factors, mentioned in Section 4.2, on ADD's.

4.5 Formulas for Estimating ADD's. The modeling effort was directed toward deriving estimates of ADDPTN(h_o) and ADDPTX(h_o), (h_o = 00, 01, 02) for Option 1; and estimates of ADDTN(h) (h = 00, 01, ..., 08) and ADDTX(h) (h = 09, 10, ..., 17) for Option 2. Estimates for TN would be subtracted from the observed PTN(h_o) or observed T(h) to obtain $\hat{TN}$ and estimates for TX would be added to the observed PTX(h_o) or observed T(h) to obtain $\hat{TX}$.

4.5.1 Option 1: Use of PTN(h_o) and PTX(h_o). Formulas for deriving estimates of the respective ADD's were obtained by regressing the mean values shown in Table 4.2 on the following independent variables:

a. SUNCOR = sun (solar) time correction in hours

= $\pm$ 1/15 (longitude of weather station-longitude of standard meridian)

where the standard meridian is the longitude which defines the time zone for the weather station (0°, 15°, 30° etc.; east or west). The sign is (+) for stations east of the Greenwish meridian and (-) for stations west.

b. DL = daylength in hours,

c. TR = long-term average daily temperature range in °F,

The latter two variables vary daily but for model generation, DL and TR were considered "constant" within a month and values used for a given month were the value of DL on the 15th and the monthly mean value of TR, respectively. Actually SUNCOR is a mean value for the year for the difference between solar time and clock time at a given location.

Results of regressing ADD estimates from Table 4.2 on values for variables (a)-(c) above gave linear equations ($Y = B_0+B_1X_1+...+B_1X_r$) whose coefficients (B's) for the derived variables (X's) are given in Table 4.3. Separate equations were developed for the three possible observation schedules (h_o = 00, 01, 02). Each equation is based on N = 36 location-months of data.

4.5.2 <u>Option 2: Use of T(h)</u>. For estimating ADDTN(h) and ADDTX(h), a new variable was needed to estimate solar time. This was defined by

$$ST = h + SUNCOR = \text{solar (sun) time},$$

where

$$h = \text{clock time when temperature was taken.}$$

Results of regressing ADD estimates, shown in Tables C-1 to C-3 (Appendix C), on ST, DL, TR plus squared and cross-product forms of these basic independent variables are shown in Table 4.4. The entries in Table 4.4 are coefficients of the derived variables shown in the right-hand column. Thus ADDTX(h) would be estimated by:

$$\begin{aligned}\widehat{ADDTX(h)} = {} & 77.23 - 9.984 * (h + SUNCOR) - 1.533 * DL \\ & + 0.426 * TR + 0.368906\ (h + SUNCOR)^2 \\ & + 0.053822 * DL^2 - 0.021126(h + SUNCOR) * DL.\end{aligned}$$

In application, we recommend that Option 2 be exercised only if at least one trihourly observation is missing. Table 4.5 shows priorities for choice of a specified hour among those for which readings are available. One could use more than one reading but there is little gain in information due to the large correlations among multiple readings on the same day.

Table 4.3 Coefficients (B's) for derived variables (X's) in equations to estimate average daily differences between psuedo temperature extremes and the respective extreme values.

h_o = 00		h_o = 01		h_o = 02		Derived Variables† (X's)
ADDPTN	ADDPTX	ADDPTN	ADDPTX	ADDPTN	ADDPTX	
+11.552863	+0.491811	+0.458731	+1.181802	-2.453513	+1.177449	1 (Intercept)
+0.503515	+0.409910		+0.691554	+0.368277	+0.398488	SC
-1.472859				-0.457507		DL
				+0.500000		TR
+0.045956				+0.014372		DL * DL
		+0.002285		-0.007771		TR * TR
+0.006668	+0.003199		+0.001536		+0.001586	DL * TR

†SC = SUNCOR (See Section 4.5.1)

DL = daylength (calculated by subroutine) in hours

TR = daily temperature range (see Section 4.5.1) in °F

Table 4.4 Coefficients (B's) for derived variables (X's) in equations to estimate average daily differences between specified hourly temperatures and the respective extreme values.

ADDTN	ADDTX	Variables† (X's)
7.712	77.232	1 (Intercept)
6.648138	-9.984249	ST
	-1.533265	DL
	+0.426257	TR
-0.425058	+0.368906	ST * ST
-0.018818	+0.053822	DL * DL
-0.791654		ST * DL
-0.187218	-0.021126	ST * TR
+0.008460		DL * TR
+0.055679		DL * ST * ST
+0.020133		ST * DL * DL
+0.004224		ST * TR * TR

†ST = h + SUNCOR = sun (solar) time (see Section 4.5.2)

DL = daylength (calculated by subroutine) in hours

TR = daily temperature range (monthly average) in °F

Table 4.5 Priorities for choice of hours (h) to use when estimating ADDTN(h) and ADDTX(h). Entries are hours of the day.

Estimation of:	Time zone with:	1st choice	2nd choice	3rd choice
ADDTN	$h_o = 00$	03	06	00
	$h_o = 01$	04	01	07
	$h_o = 02$	05	02	08
ADDTX	$h_o = 00$	15	12	09
	$h_o = 01$	13	16	10
	$h_o = 02$	14	11	17

The priority scheme was determined by examining standard deviations in Tables C-1 to C-3. Hours for which standard deviations tended to be smallest were given highest priority. High priority hours tended to be those closest to the time when minimums and maximums "normally" occur.

4.6 Test of model

4.6.1 Sious Falls, SD. Temperature data from Sioux Falls, SD was used for an initial test of the model. Sioux Falls is located at longitude 96° 44' W, latitude 43° 34'N, and elevation 435 feet. Hourly temperatures, together with daily minimum and maximum temperatures were available for the period 1949 to 1964. Four months (March, June, September, December) were selected for test purposes.

Model-generated and measured mean daily minimum and maximum temperatures for monthly periods are shown in Table C-4 of Appendix C. Missing data in the table indicates sufficient missing temperature data to give misleading test results.

Summary statistics in Table C-4 include a root-mean-square error (RMSE), calculated over years, for each month, and the range of differences between model-generated and measured values. The RMSE for Option 1 (use of psuedo maximums and minimums to estimate TX and TN) was less than that for Option 2 (use of specified hourly readings to estimate TX and TN) for all four months and is the basis for recommending use of Option 2 only if there were missing data among the eight readings.

Careful examination of Table C-4 leads to the following conclusions:

a. Bias - Some bias for particular months is almost sure to exist because a quadratic function of the variables could not be expected to give exact values for ADD's for every month at every location.

For Option 1, the bias is nondetectable and for Option 2, the bias appears to be at most about 1°F in some months.

b. Range of differences. For Option 1, the largest difference between a mean of model-generated and measured temperatures was 2°F. For Option 2, it was 3°F. Roughly, this translates into about a one bushel difference in yield estimates for our spring wheat yield model.

4.6.2 Some USSR Locations. Results of further testing of the model for three locations in the USSR are shown in Table C-5. Geographical characteristics are as follows:

City	Latitude	Longitude	Elevation
Chernovtsy	48° 18'N	25° 56' E	787 feet
Sverdlovsk	56° 52'N	60° 35'E	2549 feet
Kurgan	55° 30'N	65° 20'E	< 100 feet

Some of the location-months show very close agreement between model and measured values; others show discrepancies up to 5.4°F. Possible causes for large discrepancies include:

a. Whole days of missing data in the tape containing trihourly readings (computer program would use estimates based on "valid" estimates before and after the day of interest),

b. Erroneous entries in the trihourly tape, secured from NWC at Asheville and/or erroneous values for measured minimum and maximum values secured from the Gramex library in Washington, D.C.,

c. Bias in our model for certain location-months.

An indirect check on the veracity of model-generated minimums and maximums is contained in the application of our spring wheat (Section 5.2) and winter wheat (Section 6.2) yield models to USSR environments over an 8-year period. Yields were computed using the trihourly data base. Precision of results was as good as application in the USGP where measured minimums and maximums were used.

4.7 Implementation. Formulas and algorithms were programmed at CCEA-NOAA, Columbia, Missouri for computer application. A description of the programs is contained in a document prepared by B. Juen and K. Williams under the direction of Dr. Gerald Barger, LEC, for the EOD-JSC dated April, 1977. The document is titled "As built design specification for historical daily data bases for testing advanced models."

Computer runs were made at CCEA-NOAA to test the model (Table C-5) and to read daily precipitation and generate estimated daily minimum and maximum temperatures for 19 stations in the USSR over the period 1965-75. Model-generated temperatures were used to test our spring and winter wheat yield models in USSR environments.

5.0 SPRING WHEAT YIELD MODEL - REVISED

An initial spring wheat yield model was proposed in the final report of Contract NAS 9-14282. Our data set for model development showed a rather strong correlation between the long-term average daily temperature in January (ADTJ) and plot yields. Locations with lower values of ADTJ tended to have higher yields. We included ADTJ as a multiplier of some soil moisture stress terms in the initial model but the results were hard to explain agronomically.

A second problem with our initial spring wheat yield model seemed to involve use of a seven-inch capacity VSMB (Versatile Soil Moisture Budget) combined with the thresholds we used for winter wheat (based on a 10-inch capacity) to measure soil moisture stress (SM and SSM variables). Under the given conditions, the moisture stress variables did not adequately differentiate between dry and wet years.

The major changes made to develop a new model were:

a. Use of a 10-inch capacity budget for all locations.

b. To allow the variables AE (actual evapotranspiration - simulated) and RE = AE/PE (actual divided by potential evapotranspiration - simulated), defined over the different phases of the crop calendar, to compete with thresholded SM variables for entry into the model.

c. To allow the variable CNTS (sum of contents of zone 4 and 5 in the VSMB) to come into the model in competition with the thresholded SSM variable.

d. To introduce thresholded precipitation variables to compete with other moisture-related terms for entry into the model.

e. To allow use of both intercept and slope parameters (rather than a single slope = MAP value) when relating regional yields to model values, developed from plot data.

5.1 Form and substance of the model. The weather related component of the spring wheat yield model is shown as equation 5.3. Equation 5.3 is not the final form of the model. In the final form, we relate yields $[\hat{Y}(R, S)]$ for a region (R) to weather at a station (S) by

$$(5.1) \qquad \hat{Y}(R, S) = a(R, S) + b(R, S) * WAC(R, S)$$

where

a(R, S) and b(R, S) are the intercept and slope, respectively, of a linear function relating recorded regional yields to model-generated WAC values;

and

(5.2) WAC(R, S) = a Weather And Cultural practice component

$$= VYA(R) * \sum_{j=1}^{3} p_j(R)\ [W_j(S) + NI_j(R) * W_o(S)]$$

where

VYA(R) = a varietal yielding ability factor determined by the VYA of varieties grown in region R,

j = 1, 2, 3 represent continuous, fallow, and irrigated cropping practices, respectively,

$p_j(R)$ = proportion of wheat under cropping practice j in region R,

$NI_j(R)$ = amount of nitrogen applied (pounds/acre) under cropping practice j in region R,

$W_j(S)$ = a weather related component for cropping practice j generated by daily weather and climatological factors at station S,

$W_o(S)$ = a weather related component, generated at station S, which is the coefficient of NI in what we formerly called the plot-based part of our model [$W_o(S)$ is a constant for spring wheat but not for winter wheat].

Equation 5.3 gives values for $W_j(S)$ (j=1,2,3). Variables whose values may change with cropping practice are subscripted with a j. For j=1, values would be used from the VSMB for cropping on a continuous basis while for j=2 the values would come from a VSMB budget carried forward in time with every other year fallowed. Equation 5.3 may also be used to estimate irrigated (j=3) yields by setting the following variables at the specified values:

Variable = Value
TP_3_PJ = 3.0
TP_5_PM = 5.0
TP_9_PH = 5.0
AE_HM = 2.5
CNT_M = 4.0

Finally $W_o(S) = 0.0905$ for all cropping practices.

(5.3)
$$
\begin{aligned}
W_j(S) = {} & 154.45 \\
& + 3.660 * (TP\_3\_PJ)_j \\
& + 3.175 * (TP\_5\_PM)_j \\
& - 2.446 * (TP\_9\_PH)_j \\
& - 9.162 * (RE\_TJ) \\
& + 3.861 * (AE\_HM)_j \\
& + 1.887 * (CNT\_M)_j \\
& \\
& - 0.473 * ATX\_JF \\
& - 0.370 * ATX\_FH \\
& - 0.342 * ATX\_HM \\
& - 0.036 * ATX\_HM * PR\_HM \\
& - 0.590 * ATX\_MD \\
& - 0.294 * ATX78\_DR, \qquad (j=1,2,3)
\end{aligned}
$$

where

a) $(TP\_\beta\_rs)_j$ = PR_rs if PR_rs $\leq \beta$ (inches)

$= \beta$ if PR_rs $> \beta$,

and PR_rs = cumulative precipitation from simulated (BMTS) crop calendar stage r to stage s, (j=1 or 2). For j=3, TP_β_rs = a constant.

b) $(AE\_HM)_j$ = total simulated (VSMB) actual evapotranspiration (in inches) from simulated stage H(BMTS=3) to stage M(BMTS=3.5) (BMTS values on the Robertson scale), (j=1 or 2). For j=3, AE_HM = 2.5.

c) RE_TJ = (AE_TJ)/(PE_TJ) = ratio of simulated actual evapotranspiration to simulated potential evapotranspiration during the period from stage T(BMTS=1.5) to J(BMTS=2.0).

d) $(CNT\_M)_j$ = contents (in inches) of zone 4 plus zone 5 in the VSMB at crop stage M(BMTS=3.5), (j=1 or 2). For j=3, CNT_M = 4.0.

e) ATX_rs = average of daily maximum temperatures (°F) from crop stage r to stage s.

f) ATX78_DR = 0, if ATX_DR $\leq$ 78,

= (ATX_DR − 78), if ATX_DR > 78.

All values involving Baier and Robertson's VSMB are based on a 10-inch capacity budget.

5.2 Application

5.2.1 Summary of Results. The revised yield model was applied to North Dakota and two oblasts in the USSR (Kurgan and Tselinograd) for test and evaluation.

For North Dakota, the root-mean-square-error (RMSE) when comparing model and USDA-SRS yields was 2.1 bushels/acre for the ten-year period 1967-76. The largest difference between the model and SRS estimates was 3.4 b/A which represented a considerable improvement over results shown in the final report for Contract NAS 9-14282, for the original model. For the Tselinograd oblast, the RMSE was 2.8 b/A for the period 1965-73 when using the jackknife method as a testing technique. Only seven years of data were available for Kurgan and the jackknife method did not show yields responding to weather [$b(R, S) \cong 0$]. However, this provided an opportunity to demonstrate how to build a weather-related model for Kurgan using results from both Tselinograd and Kurgan.

5.2.2 General Conditions for Test Runs. The model-generated yields were produced under the following conditions:

a. The spring wheat variable-date starter model was used for all locations for all years. The fifty percent planted (P50) date was estimated as the date when the sum of the warming/planting (W/P) days reached 35.5.

b. North Dakota estimates were made using the bootstrap technique and Kurgan and Tselinograd by the jackknife technique. For North Dakota, values of $a(R,S)$ and $b(R,S)$ in Equation (5.1) were determined from regional (CRD) yields and model-generated WAC-values for the ten-year period preceding the year for which an estimate was given. For the USSR oblasts, the year for which an estimate was made was not included in calculating a(R,S) and b(R,S).

5.2.3 Input Parameters. In application one has to assign seasonal values to:

a. parameters associated with cultural (agronomic) practices [applied nitrogen (NI), proportion of continuous (p_1), fallow (p_2), and irrigated (p_3) wheat, and varietal yielding ability (VYA)] for a given region (R),

b. parameters used to relate regional recorded (published) yields to model-generated WAC (weather and cultural practice) values; namely, a(R,S) and b(R,S).

For North Dakota, values of VYA, NI, and p_1 ($p_1 + p_2 = 1.0$) are given in Table 5.1. The following region-station combinations were used:

Region (CRD)	Station
NW-NC	Minot (MNT)
NE	Grand Forks (GRD)
WC-SW	Dickinson (DCN)
C	Jamestown (JAM)
SC	Bismarck (BIK)
EC-SE	Fargo (FGO)

For the Kurgan and Tselinograd oblasts we set VYA = 1.0, NI = 0, and p_2 = 1.00 (all fallow).

We have moved away from the single MAP parameter to use of both intercept and slope parameters since use of the single MAP parameter was unnecessarily restrictive in fitting recorded yields to model-generated quantities. Table

Table 5.1 Values of input parameters† for crop reporting districts in North Dakota.

CRD's	NW-NC			NE			WC-SW			C			SC			EC-SE		
Yr	VYA	NI	p_1	VYA	NI	p_1	VYA	NI	p_1	VYA	NI	p_1	VYA	NI	p_1	VYA	NI	p_1
57	1.12	0	0.28	1.10	5	0.46	1.12	0	0.36	1.10	0	0.54	1.13	0	0.73	1.11	5	0.63
58	1.12	0	0.16	1.10	5	0.37	1.12	0	0.38	1.10	0	0.45	1.13	0	0.68	1.11	5	0.58
59	1.12	0	0.10	1.10	5	0.30	1.12	0	0.41	1.10	0	0.42	1.13	0	0.69	1.11	5	0.60
60	1.12	0	0.17	1.10	5	0.29	1.12	0	0.38	1.10	0	0.42	1.13	0	0.74	1.11	5	0.59
61	1.12	0	0.12	1.10	5	0.23	1.12	0	0.32	1.10	0	0.36	1.13	0	0.59	1.11	5	0.56
62	1.12	0	0.11	1.10	5	0.25	1.12	0	0.28	1.10	0	0.28	1.13	0	0.55	1.11	5	0.46
63	1.12	0	0.08	1.10	5	0.10	1.12	0	0.16	1.10	0	0.20	1.13	0	0.45	1.11	5	0.38
64	1.11	2	0.08	1.10	7	0.12	1.12	1	0.19	1.10	3	0.25	1.13	3	0.49	1.10	7	0.48
65	1.11	2	0.08	1.11	7	0.18	1.12	1	0.16	1.10	3	0.22	1.12	3	0.49	1.11	7	0.42
66	1.12	2	0.10	1.12	8	0.21	1.12	2	0.16	1.11	3	0.25	1.12	3	0.50	1.11	7	0.42
67	1.12	3	0.13	1.13	9	0.25	1.12	2	0.18	1.12	4	0.30	1.11	4	0.54	1.12	9	0.48
68	1.14	3	0.14	1.14	11	0.28	1.12	2	0.19	1.13	4	0.33	1.11	4	0.55	1.13	11	0.50
69	1.14	5	0.08	1.14	17	0.20	1.12	3	0.11	1.13	7	0.20	1.10	7	0.39	1.14	17	0.39
70	1.15	6	0.05	1.15	19	0.17	1.12	4	0.06	1.15	8	0.14	1.12	8	0.29	1.15	19	0.30
71	1.16	6	0.10	1.15	21	0.22	1.14	4	0.08	1.15	8	0.23	1.14	8	0.32	1.16	21	0.38
72	1.14	7	0.08	1.14	23	0.26	1.14	5	0.08	1.15	9	0.25	1.15	9	0.26	1.15	23	0.44
73	1.14	7	0.07	1.17	31	0.21	1.14	6	0.07	1.15	12	0.21	1.15	12	0.23	1.15	31	0.45
74	1.14	8	0.19	1.17	25	0.38	1.15	5	0.16	1.16	10	0.42	1.15	10	0.42	1.16	25	0.68
75	1.15	8	0.19	1.17	26	0.43	1.15	5	0.16	1.16	11	0.49	1.16	11	0.54	0.16	25	0.68
76	1.15	9	0.19	1.17	26	0.43	1.15	5	0.16	1.16	12	0.49	1.16	11	0.54	1.16	25	0.68

†VYA = varietal yielding ability, NI = amount of nitrogen (lbs.), p_1 = proportion of continuous cropping.

5.2 shows values of a(R,S) and b(R,S) used in Equation (5.1) to calculate estimated yields for the separate regions of North Dakota. The a's and b's were determined by least squares (linear regression) methods with CRD-SRS yields regressed on WAC values for the ten-year period preceding the year for which an estimate was being made. The estimation process was, in essence, the same as would be used on a real time basis.

Table 5.3 shows values of a(R,S) and b(R,S) for USSR data where the jackknife technique was employed.

The R^2 values for North Dakota in Table 5.4 are of considerable interest. For all CRD's, one finds increasingly better fit of recorded CRD yields to model-generated WAC values as you move from the 1955-64 period to the 1967-76 period. Clearly, the results indicate that we should not use more than 10 to 12 years to calculate a(R,S) and b(R,S) values for the current crop year. We recommend use of the 1967-76 period for the 1977 crop year.

The R^2 values for the two USSR oblasts (Table 5.5) also bear further investigation. The R^2 values for Atbasar show unusually good fit of recorded to model-generated data; those for Tselinograd are of about the same magnitude as the later years for North Dakota data, while those for Kurgan show essentially zero correlation.

A partial explanation could lie in the fact that the published yields for the Tselinograd oblast range from 3.0 to 24.3 bushels/acre while those for Kurgan had a smaller range (18.9-30.3). However, the range of SRS yields for regions in North Dakota for 1967-76 (see Table 5.6) are no larger but fit of recorded to model-generated yields is quite good. The R^2 values for Kurgan look much like those for the North Dakota data in the 1955-64 ten-year period. It may be that in both cases the CRD (oblast) data were sufficiently in error in two or three years out of ten (six for Kurgan) to give practically zero correlation when coupled with random model errors.

Table 5.2 Values of a(R,S) and b(R,S) when regressing North Dakota CRD yields on WAC values over ten-year periods.

	Regions (CRD's) and Stations											
	SC BIK		WC-SW DCN		EC-SE FGO		NE GRD		C JAM		NW-NC MNT	
	a	b	a	b	a	b	a	b	a	b	a	b
55-64	11.73	0.14	8.32	0.43	17.36	0.11	15.81	0.28	14.58	0.21	10.37	0.32
56-65	8.75	0.33	8.79	0.46	22.22	0.03	15.56	0.32	13.28	0.30	8.43	0.43
57-66	9.20	0.34	7.60	0.56	21.62	0.01	13.85	0.40	14.30	0.27	4.96	0.60
58-67	11.42	0.24	6.42	0.61	20.58	0.07	16.33	0.35	15.20	0.27	6.97	0.54
59-68	7.70	0.49	7.83	0.55	8.68	0.54	15.65	0.37	12.72	0.41	10.28	0.42
60-69	8.57	0.49	9.42	0.52	8.84	0.56	13.01	0.49	13.61	0.42	10.97	0.43
61-70	9.10	0.46	9.08	0.53	10.56	0.49	13.03	0.50	13.84	0.41	11.48	0.42
62-71	14.77	0.26	17.97	0.22	7.80	0.61	18.33	0.36	18.17	0.30	20.02	0.19
63-72	11.61	0.35	15.72	0.29	5.35	0.67	19.17	0.33	13.93	0.44	17.62	0.24
64-73	12.97	0.30	15.76	0.30	8.05	0.60	20.33	0.30	15.48	0.41	15.63	0.30
65-74	11.99	0.35	16.99	0.26	10.13	0.54	10.82	0.55	13.92	0.45	13.98	0.34
66-75	11.99	0.35	16.80	0.27	13.79	0.44	10.64	0.56	13.57	0.49	14.08	0.33
67-76	11.98	0.33	17.20	0.26	14.75	0.42	14.40	0.47	13.76	0.47	14.14	0.33

Table 5.3 Values of a(R,S) and b(R,S) when regressing USSR oblast yields on WAC values for the jackknife test. (All years except omitted year used in regression.)

	Regions (Oblasts) and Stations					
	Tselinograd				Kurgan	
	Atbasar		Tselinograd		Kurgan	
Omitted Year	a	b	a	b	a	b
1965	4.28	0.48	4.73	0.39	Missing WX Data	
1966	2.62	0.53	0.95	0.49	25.95	-0.03
1967	3.41	0.52	1.69	0.47	25.26	-0.01
1968	2.63	0.54	2.91	0.43	20.84	0.08
1969	2.63	0.53	1.60	0.46	23.97	0.02
1970	3.11	0.53	2.86	0.44	22.14	0.08
1971	2.58	0.53	2.55	0.44	23.41	0.06
1972	-0.84	0.79	5.21	0.29	25.05	-0.04
1973	2.80	0.53	2.08	0.48	Missing WX Data	

Table 5.4 Values of R^2 when regressing North Dakota CRD yields on WAC values over ten-year intervals.

	Regions (CRD) and Stations					
Years	SC BIK	WC-SW DCN	EC-SE FGO	NE GRD	C JAM	NW-NC MNT
55-64	.02	.20	.02	.09	.05	.15
56-65	.10	.20	.00	.19	.13	.25
57-66	.10	.35	.00	.31	.11	.40
58-67	.07	.39	.01	.26	.13	.37
59-68	.22	.40	.33	.27	.24	.36
60-69	.36	.50	.46	.45	.34	.47
61-70	.30	.45	.39	.53	.34	.47
62-71	.15	.20	.70	.64	.24	.19
63-72	.56	.61	.80	.64	.48	.42
64-73	.51	.63	.78	.57	.56	.47
65-74	.73	.66	.60	.80	.66	.55
66-75	.73	.66	.54	.79	.76	.56
67-76	.59	.68	.55	.68	.68	.54

Table 5.5 Values of R^2 when regressing USSR oblast yields on WAC values for the jackknife test. (All years except omitted year used in regression.)

Omitted Year	Regions (Oblasts) and Stations		
	Tselinograd (1965-73)		Kurgan (1966-72)
	Atbasar	Tselinograd	Kurgan
1965	.89	.72	--
1966	.91	.71	.00
1967	.91	.58	.00
1968	.89	.61	.05
1969	.91	.67	.00
1970	.90	.64	.01
1971	.91	.61	.03
1972	.89	.27	.01
1973	.90	.63	--

Table 5.6 Comparison of Model (KSU_F) and USDA (SRS) Yields for North Dakota. Entries are Bushels Per Acre.

Year		Regions (CRD's), Stations, and Percent Acreage						
		NW-NC Minot 29%	NE Grand Forks 19%	WC-SW Dickinson 18%	SC Bismarck 6%	C Jamestown 10%	EC-SE Fargo 18%	State
1967	KSU	15.9	24.7	20.2	12.6	17.7	21.9	19.4
	SRS	19.0	28.9	22.5	17.2	20.2	26.1	22.6
1968	KSU	32.8	29.0	28.8	18.3	24.8	23.1	27.9
	SRS	23.8	31.4	23.6	22.9	29.0	30.7	26.8
1969	KSU	28.1	29.6	24.7	22.9	27.0	27.2	27.2
	SRS	31.1	33.8	25.8	23.1	30.8	29.6	29.8
1970	KSU	23.6	28.2	19.8	20.0	23.1	27.5	24.2
	SRS	23.3	28.1	21.0	16.5	23.7	24.3	23.6
1971	KSU	27.4	35.1	24.6	22.8	25.2	30.2	28.4
	SRS	30.1	35.8	27.8	26.5	33.4	34.4	31.8
1972	KSU	28.4	32.9	27.7	25.1	27.6	30.4	29.2
	SRS	29.2	31.3	29.2	23.8	27.0	28.5	28.9
1973	KSU	27.1	31.3	26.2	19.3	20.6	25.5	26.3
	SRS	29.7	30.3	27.6	19.7	22.2	27.2	27.5
1974	KSU	23.1	28.0	20.0	14.9	20.5	23.5	22.8
	SRS	20.7	22.4	20.9	12.8	17.7	27.3	20.1
1975	KSU	23.9	29.6	22.9	21.5	24.2	22.6	24.5
	SRS	24.7	31.0	23.9	21.1	25.5	26.1	25.9
1976	KSU	26.8	25.4	26.0	22.0	25.9	25.1	25.7
	SRS	25.4	29.6	25.6	16.5	21.8	23.8	24.7

5.2.4 Estimated Yields. In this section we show estimated yields, for the three test areas, together with recorded yields. Table 5.6 contains results for North Dakota and Table 5.7 results for the Kurgan and Tselinograd oblasts in the USSR.

While various statistics must be calculated to properly test the results, it is clear that the KSU_F revised spring wheat model is an improvement over the earlier version. In the initial version, the largest discrepancy between model and SRS data for North Dakota state average yields was 6.1 bushels. In this version it is 3.4 bushels.

In the Tselinograd area, where one expects weather to be very influential and large year-to-year variation in yields, the model properly reflects that variation.

In Kurgan, higher precipitation and cooler temperatures would be expected to give higher yields than in the Tselinograd area and both recorded data and WAC-values (see Table 5.8) reflect this. However, the published data does not reflect the year-to-year variation due to weather that published data for North Dakota (during the past 10-15 years) and Tselinograd show.

There remains the problem of estimating yields for Kurgan that reflect weather variation. One solution is as follows:

a) Let $\bar{R}$ and $\overline{WAC}$ be the mean recorded and WAC yields for Kurgan.

b) The average yields should satisfy the relation

$$\bar{R} = a + b\,\overline{WAC}$$

or, from Table 5.8,

$$24.5 = a + b\,(42.8) \tag{5.4}$$

Table 5.7 Comparison of Model (KSU_F) and Recorded Yields for Two Oblasts in the USSR. Entries are Bushels Per Acre.

		Regions (Oblasts) and Stations			
		Tselinograd			Kurgan
Year		ATBASAR	BOTH +	TSELINOGRAD	Kurgan
1965	KSU	6.7	9.0	11.3	
	OBLAST		3.0		
1966	KSU	12.2	10.3	8.5	25.2
	OBLAST		14.2		22.6
1967	KSU	9.9	7.7	5.5	25.0
	OBLAST		7.5		22.2
1968	KSU	9.6	10.4	11.3	23.1
	OBLAST		9.8		30.3
1969	KSU	12.1	11.0	9.8	24.5
	OBLAST		14.2		24.3
1970	KSU	12.3	12.8	13.3	25.9
	OBLAST		10.5		24.3
1971	KSU	12.0	12.7	13.4	25.5
	OBLAST		14.4		18.9
1972	KSU	33.5	25.7	17.9	23.6
	OBLAST		24.3		28.8
1973	KSU	14.0	16.2	18.4	
	OBLAST		15.3		

+Both indicates that the mean of KSU_F model for Atbasar and Tselinograd was used.

c) From Table 5.3, note that the values of b for Atbasar and Tselinograd are approximately 0.50.

d) Assume b = 0.50, substitute into Equation 5.4 and calculate a = 3.1.

e) For Kurgan, estimate yields with the formula:

$$\hat{Y} = 3.1 + 0.5 * WAC \qquad (5.5)$$

Note that the value of a = 3.1 is close to the values for Atbasar and Tselinograd. If Equation (5.5) is a reliable indicator of yield then the major contributing difference in yields between the Kurgan and Tselinograd oblasts is weather since cultural practice terms were assumed constant when computing WAC values.

Finally, if Equation (5.5) is applied with the historic Kurgan WAC values, the resulting yields in bushels per acre are shown in Table 5.9. The RMSE (root mean square error) for the model using Equation (5.5) is 3.87 while that using a and b values from Table 5.3, and comparing yields given in Table 5.7, gives RMSE = 4.47.

If one wants to trust the model, then the results in Table 5.9 suggest that oblast yields were overestimated in 1968 and 1969 and underestimated in 1970 and 1971 which would have contributed to small R^2 over all the different six-year period used for jackknife testing.

Use of Equation (5.5) for 1974 and 1975 (we had weather data but no yield data) would have produced estimates of 19.0 and 15.4 bushels per acre, respectively, for the Kurgan oblast. These values, which would indicate relatively poor yields, can be checked against recorded data to further test such a model for Kurgan.

Table 5.8 Comparison of Kurgan and Tselinograd oblasts for (a) recorded yields and (b) WAC values.

	Recorded yields		WAC values	
Year	Kurgan	Tselinograd	Kurgan	Tselinograd†
1966	22.6	14.2	39.0	16.6
1967	22.2	7.5	37.5	10.3
1968	30.3	9.8	43.5	15.5
1969	24.3	14.2	32.2	17.9
1970	24.3	10.5	55.1	20.5
1971	18.9	14.4	39.3	21.2
1972	28.8	24.3	53.0	43.5
Means	24.5	13.6	42.8	20.8
Difference of means	10.9		22.0	

†Mean WAC value for Atbasar and Tselinograd

Table 5.9 Comparison of recorded and model (Equation 5.5) yields for Kurgan.

	Years						
	'66	'67	'68	'69	'70	'71	'72
Equation (5.5)	22.6	21.8	24.8	19.2	30.6	22.8	29.6
Recorded	22.6	22.2	30.3	24.3	24.3	18.9	28.8

6.0 APPLICATION OF WINTER WHEAT MODEL

6.1 Kansas

The final report for Contract NAS 9-14282 contained results for application of the KSU_F winter wheat model using nine cooperative weather stations (one per CRD). Yields were estimated for each CRD using the "bootstrap" technique. The CRD estimates were weighted, by acreage, and combined into state yields for the 10-year period 1967-76.

Subsequently, a more exhaustive analysis was undertaken for the 22-year period 1955-76 using a network of 42 weather stations in Kansas. A subset of seven (four first-order Weather Bureau and three FAA) of the forty-two were selected for real-time testing of our model.

In addition, data were secured, from A. P. Roelfs of the USDA Cereal Rust Laboratory at St. Paul, Minnesota, on yearly percent loss of wheat due to stem and leaf rust in Kansas for the 1955-76 period. This provided the opportunity to include losses due to rusts in our yield model.

The results in Table 6.1 show a comparison of model-generated yields and RMSE (root mean square error) values, where

$$\text{RMSE} = [\sum_{t=55}^{76} (\text{model} - \text{SRS})^2/N]^{\frac{1}{2}},$$

for the four combinations obtained by varying:

(a) density of stations (7 and 42),

(b) amount of information (info) on leaf and stem rust loss (none and full).

<u>Results in Table 6.1 show about a 0.5 bushel/acre reduction in the RMSE when station density is increased from 7 to 42 and an additional reduction of</u>

<u>0.3 bushel/acre when losses due to rusts are included in the model</u>. A more detailed discussion of these results follow:

6.1.1 <u>Estimation procedure</u>. Steps used to calculate model-generated yields shown in Table 6.1 were as follows:

(a) WAC values (see Eq. 5.2) were calculated for each year for each of the 42 stations,

(b) a statewide weighted average of WAC ($\overline{WAC}$) values was computed for each year using 7 and then 42 stations where weights for each station were based on acreage (long-term averages),

(c) For each year and station density, the quantity

$$(1 - p) * \overline{WAC},$$

where p = proportional loss due to stem and leaf rust, was computed,

(d) MAP values were calculated for the four cases as

$$MAP = \sum_{t=55}^{76} (\text{SRS yield})_t / \sum_{t=55}^{76} (1-p_t)\overline{WAC}_t$$

where p_t = 0 for the case of no information on rust losses.

(e) The quantity

$$MAP * (1-p_t)\ \overline{WAC}_t \qquad (t=55,\ 56,\ldots,76)$$

for each of the four cases gave model-generated yield estimates.

The results for the revised spring wheat model suggest that both an intercept a(R, S) and slope b(R, S) term should be used to relate a regional yield estimate to WAC (or $\overline{WAC}$) values. A comparison of RMSE values using both intercept and slope versus slope (MAP value) only showed no advantage to including the intercept for the winter wheat model applied in Kansas.

6.1.2 <u>Density of weather stations</u>. The results in Table 6.1 show the gain in precision (reduction in RMSE) achieved by increasing density from

Table 6.1 Comparison of model-generated yields (two station densities with and without rust information) and USDA-SRS estimates in Kansas. (Yields in bushels per acre)

	Model-Generated Yields				
	7-Stations		42-Stations		USDA-SRS
Year	No Info	Rust Info	No Info	Rust Info	Estimates
1955	17.9	18.2	17.7	17.8	15.0
1956	16.8	17.6	14.6	15.2	15.5
1957	22.5	20.4	19.7	17.8	19.0
1958	25.3	26.2	27.4	28.2	28.5
1959	24.1	22.7	24.0	22.4	20.5
1960	26.3	26.6	26.6	26.8	28.5
1961	26.0	24.9	27.1	25.9	26.5
1962	22.7	22.7	23.2	23.1	23.5
1963	25.1	26.1	22.0	22.8	21.5
1964	21.7	22.5	20.9	21.5	22.0
1965	22.8	22.8	22.8	22.7	23.5
1966	26.6	26.6	24.4	25.3	19.5
1967	19.0	19.8	20.8	21.6	20.0
1968	27.9	26.8	26.6	25.3	26.0
1969	31.2	32.5	31.4	32.6	31.0
1970	29.6	30.5	29.8	30.5	33.0
1971	29.8	30.4	30.4	30.9	34.5
1972	29.4	30.3	28.6	29.3	33.5
1973	32.4	31.0	34.2	32.6	37.0
1974	31.8	27.9	33.4	29.2	27.5
1975	28.8	27.6	31.0	29.5	29.0
1976	29.7	30.9	30.0	31.1	29.5
RMSE	3.1	2.9	2.6	2.3	

seven to forty-two stations in Kansas. The reduction of 0.5 bushel/acre in the RMSE represents a respectable gain in information. The most noticeable gain is in 1963 when the larger sample of stations gives an estimate 3.3 bushels closer to the SRS estimate than the smaller sample for the "rust info" case.

The gain in precision does not begin to approach the amount one obtains from independent random samples where the standard error of a mean is reduced by the factor $1/\sqrt{N}$. There are three reasons why a sizeable reduction in RMSE with a very dense network of stations cannot be expected:

(a) Weather-related variables are highly correlated from station-to-station which transforms into highly correlated model-generated yields (see Table 6.2).

(b) The SRS estimates are based on sample data and have a variance associated with them. Therefore, if model-generated values were "exact", the RMSE still would not be zero.

(c) The model is an incomplete expression and factors which influence yields are not included. Therefore, there tends to be a lower limit to the RMSE values.

An example of (c), is the lack of terms in the model to express the effect of severity of diseases. This model deficiency has been partially corrected as discussed in the following section.

6.1.3 Effect of stem and leaf rust. As indicated in 6.1.1, the effect of stem and leaf rust was incorporated into the model by multiplying our state-wide $\overline{WAC}$ values by (1 - p) where p = percent loss/100. Percent loss by year is shown in Table 6.3.

Use of the information on rust gave a further reduction in the RMSE of 0.2 bushel/acre for 7-station density and 0.3 bushel/acre for 42 stations. The

Table 6.2 Correlations between model-estimated yields at selected weather stations in Kansas (n = 21 seasons).

	HLC	DGD	CON	SAL	WIC	TOP	CHA
HLC	1.000	0.840	0.865	0.791	0.808	0.801	0.640
DGD		1.000	0.760	0.847	0.867	0.826	0.494
CON			1.000	0.877	0.831	0.937	0.722
SAL				1.000	0.919	0.887	0.637
WIC					1.000	0.838	0.637
TOP						1.000	0.717
CHA							1.000

HLC = Hill City, DGD = Dodge City, CON = Concordia, SAL = Salina, WIC = Wichita, TOP = Topeka, CHA = Chanute

Table 6.3 Percent loss due to leaf and stem rust in Kansas[†].

Year	%loss	Year	%loss
1955	3.5	1966	0.0
1956	0.0	1967	0.0
1957	13.0	1968	8.0
1958	0.5	1969	0.0
1959	10.0	1970	1.0
1960	3.0	1971	2.0
1961	8.0	1972	1.0
1962	4.0	1973	8.0
1963	0.0	1974	15.5
1964	0.6	1975	8.0
1965	4.0	1976	0.0

[†] Data made available by A. P. Roelfs, USDA Cereal Rust Laboratory, St. Paul, MN.

reduction takes on added significance when you consider that the RMSE is getting down to a value that only major contributory factor(s) could cause further reduction.

From Table 6.1, it is clear that leaf and stem rust can play a major role in reducing yields over Kansas. $\overline{\text{WAC}}$ values for 1973 and 1974 were close in size but there was a 9.5 bushel/acre difference in SRS estimates. Use of rust information helped to bring our 1974 estimates more in line with SRS. The same was true for 1959.

6.1.4 Freezes at heading. One of the factors for which we were not able to find a good weather-related variable to represent its effect was freezing temperatures near heading. The effect of this missing factor was most evident in 1966, when the four different model-generated yields overestimated SRS by 4.9 to 7.1 bushels/acre. Much of western Kansas was subjected to two hard freezes in the first two weeks of May and yield losses were apparent at the time. New modeling efforts could possibly correct this deficiency.

6.2 Khmel'Nitskiy, USSR

Results of applying our winter wheat yield model to the Khmel'Nitskiy oblast in the Ukraine area of the USSR are shown in Table 6.4. Measures of cultural practices (see footnote of Table 6.3) were assumed constant over the 8-year period. By taking VYA = 1.0, the MAP factor absorbs the comparison of yielding ability of varieties, used over that time period, with the "standard" (Pawnee/Commanche) used to develop our winter wheat model.

If a single station (Khmel'Nitskiy) is used to estimate yields, the RMSE is 3.4, compared with 2.3 for a seven-station average. The year 1968 showed the lowest recorded yield and the model indicates that it was a poor year weather-wise relative to the other years.

Table 6.4 KSU model (winter wheat) estimates compared with yields for Khmelnitskiy oblast in USSR. (yields in bushels/acre)

Year	Recorded yields for Oblast	Yields from Khmelnitskiy WX†	Average of Yields from 7 stations††
1967	45.3	40.5	40.9
1968	33.2	30.5	34.5
1969	40.8	45.7	42.9
1970	37.2	38.6	39.3
1971	44.1	42.0	41.4
1972	43.0	43.6	42.4
1973	41.2	45.5	43.1
1974	40.6	37.4	39.4

†MAP = .94, VYA = 1.0, Nitrogen = 40#/A

††MAP = .91, VYA = 1.0, Nitrogen = 40#/A

The seven stations had WMO numbers in block 33 as follows:

301, 317, 415, 429, 562, 658, 663.

ACKNOWLEDGMENTS

Excellent cooperation and support of the many individuals listed in our final report on Contract NAS9-14282 has continued through the present contract period and we gratefully acknowledge their continuing contributions.

Specific contributions to the present effort included the following. Jeanne Sebaugh constructed our wheat yield model computer program (WHYMOD) with assistance from Mike Franzblau and John Olsowski. Dale Fjell was in charge of processing agronomic data to incorporate it into the model. Mike Frerichs did the necessary programming to develop a model to estimate daily temperature extremes. Kathy Elliott and Betty Skidmore provided excellent typing and clerical services.

APPENDIX A

DEVELOPMENT OF SPRING WHEAT STARTER MODEL

1.0 Definitions

We approached the problem of modeling seasonal and regional variation in time-of-planting by considering that a cool/wet early spring could mean delays in planting while warm/dry conditions permit earlier planting. If weather conditions are not a significant determinant of time-of-planting than a fixed-date starter model may be the best one can do in fixing when p% ($0 \leq p \leq 100$) of a spring wheat crop has been planted in a given region in a given year.

To measure weather variation among seasons/regions, we defined a warming/planting (W/P) day which assigns a number from zero to one to each calendar day beginning January 19 (somewhat arbitrarily chosen but coinciding roughly with the coldest time of the year in the northern hemisphere) and continuing through the planting season. Accumulated W/P days were then related to percent of wheat planted in a given region.

All definitions for a W/P day were special cases of the general form

$$
\begin{aligned}
W/P &= 0, && \text{if } TA \leq 32, \\
&= \alpha(TA-32)(PRE), && \text{if } 32 < TA \leq 32 + 1/\alpha \\
&= 1, && \text{if } 32 + 1/\alpha < TA,
\end{aligned}
$$

where,

TA = average daily temperature (°F)

α = a selected threshold value,

PRE = 1, for Julian day J, if all the following conditions were met:

(i) Precipitation (PR) on day J < $\beta_1(TA-32)_J$,

(ii) Accumulated PR on days J, J-1 < $\beta_2(TA-32)_J$,

(iii) Accumulated PR on days J, J-1, J-2 < $\beta_3(TA-32)_J$,

(iv) Accumulated PR on days J, J-1, J-2, J-3 < $\beta_4(TA-32)_J$.

If one or more of the above conditions were not met, then PRE = 0.

Table A-1 shows values of α and β's assigned to generate a set of definitions of W/P days.

Table A-1. Definitions of a warming/planting day.

Definition	α	β_1	β_2	β_3	β_4
I	0.10	.005	.015	.025	.035
II	0.10	.001	99.9	99.9	99.9
III	0.10	99.9	99.9	99.9	99.9
IV	0.05	99.9	99.9	99.9	99.9

Definition I makes a W/P day dependent on precipitation thresholds, over a four-day period, that increase with temperatures on day J. In essence, Definition II considers precipitation on day J only since conditions (ii) - (iv) above are obviously met. Definitions III and IV depend on temperature only, with Definition III specifying an upper threshold at 42°F (W/P = 1 if TA > 42°F) and Definition IV puts the upper threshold at 52°F. A lower threshold for temperature is 32°F for all definitions.

2.0 Data sets for model development and testing.

Data on percent planted by a given Julian day by CRD were made available through the Crop Reporting Services of North Dakota, South Dakota, and Montana. The data were collected at seven-day intervals and it was necessary to interpolate such statistics as P15 (day when 15% planted was reached), P50 (day when 50% planted was reached) and other percentiles of interest. To relate cumulative W/P days to percent planted it was necessary to set up correspondences between weather stations and crop reporting districts. The following correspondences were made:

Weather Stations	Regions (CRD(s))	Region No.
Minot, ND	(NW) ND	1
Minot-Langdon, ND	(NC) ND	2
Langdon-Fargo, ND	(NE-NC) ND	3
Dickinson, ND	(WC) ND	4
Dickinson, ND - Bison, SD	(SW)ND-(NW)SD	5
Fargo, ND - Eureka, SD	(SE) ND	6
Williston, ND	(NE) MT	7
Moccasin - Bozeman, MT	(C) MT	8

Regions 1-6 were used for model development for the years 1953-73 and for testing in 1974-75. Regions 7 and 8 were used for testing only for the years 1969-73. Data was not available for all years in all regions because either weather data or planting-date data may have been missing. If two CRD's comprise a region, then P15 or P50 days were averaged to give a regional value. If two weather stations were used, then values of

$$C(W/P) = \sum_{J=19}^{P50} (W/P)_J$$

from the two stations were averaged to give an accumulated W/P value to reach P50 for the given region.

3.0 A shift toward later and more variable planting dates.

A significant shift toward later planting occurred as measured by P15 (and also P50) values during the time period used for model development. This is shown in Table A-2. This was accompanied by larger year-to-year variation in the dates by which 15% of the crop was planted (see Table A-3) in the 1967-73 period than in previous periods.

To check whether weather accounts for some or all of the shifts in means and variances, we used Definition I to compute the means and standard deviations

Table A-2. Means (over years) of dates (Julian day) when 15% of crop was planted in specified crop reporting districts.

State	CRD	Time Periods (Years) 53-59	60-66	67-73
ND	NW	118	117	129
	NC	116	118	127
	NE	117	122	126
	WC	111	111	120
	C	108	114	119
	EC	110	117	121
	SW	109	112	114
	SC	107	109	111
	SE	108	114	115
SD	NW		103	106
	NC		103	105
	NE		103	109

Table A-3. Standard deviations (over years) of dates when 15% of the crop was planted in specified crop reporting districts.

State	CRD	Time Periods (Years) 53-59	60-66	67-73
ND	NW	4	5	12
	NC	7	4	10
	NE	9	6	12
	WC	4	6	13
	E	6	5	10
	EC	7	6	12
	SW	4	11	14
	SC	3	5	10
	SE	2	4	10
SD	NW		8	9
	NC		8	9
	NE		10	12

for C(W/P) from J=19 to J=120 (May 1) calculated over years. The results in Table A-4 show no significant shifts in means or standard deviations over time. Thus, weather is eliminated as a major factor.

Conversations with agronomists have produced the following explanation. Two factors are important in the early life of the plant (a) soil temperature and (b) control of weeds. The advent of large equipment meant that farmers could cover more ground in a shorter time and be more "timely" in their field operations. Later plantings mean a greater opportunity to destroy weeds that have emerged. More year-to-year variation suggests that farmers are timing their plantings more closely to ideal soil temperatures.

Of further significance is the fact that we analyzed mean values of number of calendar days from P15 to P50 and did not find a shift over time. Thus, it was the time when planting began rather than or in addition to the rate of planting that showed changes over time.

It should be noted that a shift in P15 dates was translated into a shift in P50 dates. It is P50 dates that we will concentrate on in model development.

4.0 Choice of K

An algorithm (rule) to determine an estimated date when 50% of the crop was planted (EP50) was established by finding the mean value of

$$C(W/P) = \sum_{J=19}^{P50} (W/P)_J$$

when averaged over a specified set of region-years. Then in real-time application EP50 will be the first day when the following inequality holds

$$\sum_{J=19}^{EP50} (W/P)_J \geq K .$$

From the discussion in Section 3.0, it is clear that K should be based on the later time period (1967-73) for real-time operation.

Table A-4. Means and standard deviations of C(W/P) (cumulative warming/planting days) to May 1 for specified location

		North Dakota Weather Stations			
Statistic	Time Period	Dickinson	Minot	Langdon	Fargo
Means	1953-59	28	24	20	27
	1960-66	27	23	18	24
	1967-73	29	25	20	27
Standard Deviations	1953-59	6	6	9	8
	1960-66	5	5	5	5
	1967-73	8	8	5	7

Table A-5 shows values of K for different regions over two different time periods for three of the four definitions of a (W/P) day. If C(W/P) days explained all the year-to-year and region-to-region variation in recorded P50, then all averages in Table A-5 would have been the same. Other factors cause variability but the means appear homogeneous enough that for simplicity one could choose K=36 (Definition III) for all regions for real-time operation.

5.0 Choice of definition of (W/P) day.

We recommend Definition III on the basis of results shown in Table A-6 and its simplicity. For the period 1953-66 a variable-date starter model showed no advantage over a fixed-date model. However, during the 1967-73 period when planting became more dependent on the weather, as discussed in section 3.0, then all variable-date models had smaller RMSE values than the fixed-date for Regions 1-6.

Evidence to favor use of either Definition III or IV over I came from Regions 7-8 which provided 10 region-years of test data in Montana (Regions 1-6 provided model-development data). Finally, we recommended Definition III simply because most of our testing work used this definition and we tried definition IV as an afterthought. The gain in use of Definition IV was not sufficient to warrant changing our recommendation.

Results for Definition II were not included in Table A-6. Tests in Regions 7-8 gave RMSE values very close to those for Definition I.

6.0 Some test results.

In addition to the test results shown for Regions 7-8 in Table A-6, we also generated EP50 dates for six regions in North Dakota for 1974-75. Results are shown in Table A-7 and represent testing with data independent of that from which the model was developed.

Table A-5. Average number of C(W/P) days to reach 50% planted.

Definitions	Periods	(NW) ND	(NC) ND	(NE-EC) ND	(WC) ND	(SW)ND-(NW)SD	(SE) ND	All
		Regions						
I	1953-59	22	20	16	23			20
	1960-66	22	20	21	20		17	20
	1967-73	30	27	25	27	25	22	27
II	1953-59	25	21	16	24			22
	1960-66	21	20	22	21		20	21
	1967-73	32	28	27	27	26	23	28
III	1953-59	31	27	21	29			28
	1960-66	28	26	28	27		27	27
	1967-73	41	36	37	37	36	30	36

Table A-6. Comparison of RMSE[+] for different definitions of a (W/P) day.

Definition	Period	Regions 1-6	Regions 7-8
Fixed-date[++]	1953-66	6.2	
	1967-73	10.7	9.2
I	1953-66	6.8	
	1967-73	6.2	9.1
III	1953-66	5.8	
	1967-73	6.4	6.5
IV	1953-66	6.4	
	1967-73	6.0	6.1

[+]RMSE = $[\Sigma(\text{Recorded} - \text{Model})^2/N]^{\frac{1}{2}}$ where N=58 (53-66) and 35 (67-73) respectively, for Regions 1-6 and N=10 for Regions 7-8.

[++]Fixed-date model used EP50 = $\overline{\text{P50}}$ = mean of P50 values over region-years.

Table A-7. Comparison of EP50 dates for fixed and variable date starter models with USDA-SRS estimates of P50 for regions in North Dakota.

Year	Regions		Fixed-Date	Variable-Date	SRS
1974	NW-NC		136	131	156
	NE		135	139	161
	WC-SW		128	120	135
	C		128	132	151
	SC		124	128	129
	EC-SE		125	134	140
		RMSE	17.8	17.0	
1975	NW-NC		136	138	149
	NE		135	137	143
	WC-SW		128	140	148
	C		128	140	146
	SC		124	137	143
	EC-SE		125	136	142
		RMSE	16.4	7.4	

In 1974, planting was delayed due to a very wet weather in May. Since wetness in not a part of Definition III, the result was that both the variable-date and fixed-date models missed by a considerable amount. In 1975, planting was again late but the variable-date model detected the situation and gave much closer estimates than the fixed-date model.

To avoid "misses", as occurred in 1974, our spring wheat starter model needs to have some precipitation conditions. Going back to Definition I or II does not seem to be the answer because of the test results in Montana. Possibly higher thresholds for precipitation in, say Definition II, would give a more sensitive (W/P) day measure.

APPENDIX B

STARTER MODELS: A STUDY OF DATES OF PLANTING WINTER WHEAT

by

Paulette M. Johnson and Arlin M. Feyerherm

1.0 SUMMARY AND RECOMMENDATIONS

An intensive investigation was carried out to determine if the dates at which certain percentages of winter wheat were planted could be related to weather events prior to or during the period when planting occurs. Results were negative in that no mathematical function of daily precipitation amounts was found which explained any significant portion of the yearly variation in specified per cent planting dates (e.g. the fifty percent planting date (P50) is the date at which 50% of the crop is planted). The investigation was carried out using data collected by the USDA-SRS, along with daily meteorological data (3 to 5 stations per CRD), to estimate the specified per cent planting dates for Kansas.

In light of the results of this study it is recommended that a given date, fixed over years but variable over locations, be used to start up winter wheat crop calendars (e.g. Robertson's biometeorological time scale).

The remainder of this report gives details of the investigation that led to the above recommendation.

2.0 YEARLY VARIATION IN PER CENT PLANTED

For the past twenty-six years, the USDA-SRS and Kansas Crop and Livestock Reporting Service have collected data every seventh day, during the planting season, to estimate the per cent of winter wheat planted to that date. For each year, from 1951 through 1975, percentages were linearly interpolated for the six days between data collection points. Simple arithmetic means of the dates when 15, 50, and 85% of the crop, respectively, were planted, together with their standard deviations were determined for each CRD in Kansas and are shown in Table B-1.

The state of Kansas provided an excellent environment for studying

Table B-1. Means and Standard Deviations (S.D.) for Julian Days when 15, 50, and 85% of Crop was Planted by CRD. (Data Base: 1951-1975)

Per Cent Planted	Mean	S.D.	Mean	S.D.	Mean	S.D.
	Northwest		North Central		Northeast	
15	258*	3.1	264	6.0	265	4.7
50	266*	6.0	273	7.8	275	6.7
85	274	8.2	283	9.3	287	7.0
	West Central		Central		East Central	
15	252	7.1	268	6.8	269	6.1
50	263	9.0	277	7.1	281	7.6
85	276	8.7	287	8.0	294	7.2
	Southwest		South Central		Southeast	
15	254	5.3	265	6.5	271	6.3
50	266	8.8	275	7.1	285	7.0
85	279	9.3	286	10.1	300	7.2

*Mean values based on 1963-75 data only. Shift to later planting after 1962 due in part to outbreaks of wheat streak mosaic and Hessian Fly.

starter models. As seen in Table 1, normal (mean) dates when 50% of the crop was planted (NP50) varied from Julian day 263 in west central Kansas to 285 in the southeast. Annual precipitation amounts vary from 16 inches in the west to 40 inches in the southeast. Delays in planting are due, in part, to dry weather in the west and wet weather in the east.

Standard deviations of P50 dates varied from 6.0 to 9.0. Any model used to explain variation in P50's will have to generate estimated P50's with standard deviations less than those shown in Table 1, if it is to replace use of a mean date (NP50).

3.0 ACCUMULATED "WORKDAY" MODEL

Assume there exists a definition of a "workday" such that the number of workdays from some predetermined Julian date [(NP15)-19] to Julian date x=P15 is a constant (γ_{P15}) over years and locations. Assume similar constants γ_x for x=P50, P85 exist. Let a workday be defined by the function

$$(3.1) \quad W_i = (1-\alpha_1 PR_i)^+ (1-\alpha_2 PR_{i-1})^+ (1-\alpha_3 PR_{i-2})^+ [1-(1-\beta * CPR_i)^+]$$

where

W_i = a workday measured as a proportion ($0 \le W_i \le 1$) of the ith Julian day,

α_1, α_2, α_3 = parameters (constants) for a three-day wetness factor,

β = parameter (a constant) for a dryness factor

PR_j = precipitation on day j (j=i, i-1, i-2),

CPR_i = cumulative precipitation from day i-19 through day i,

$(\quad)^+$ = zero, if the quantity in parentheses is negative, and equal its value otherwise.

If our assumptions were correct, then there would exist constants α_1, α_2, α_3, β and γ_x such that

$$(3.2) \qquad \gamma_x = \sum_{i=(NP15)-19}^{x} W_i, \qquad x=P15,\ P50,\ P85$$

for all years and locations. To predict x=P50 for a given location-year, accumulate W_i values until the sum γ_{P50} is reached; the Julian date on which this occurs is the predicted P50.

Due to the fact that other sources of variation are involved besides precipitation or the lack thereof, in determining the number of calendar days to get to 15 or 50 or 85% planted, a statistical approach becomes necessary.

According to one statistical criterion the problem involves estimating parameters α_1, α_2, α_3, β, and γ_x so that the mean square error.

$$(3.3) \qquad MSE = \sum_{j=1}^{n} (\hat{x}_j - x_j)^2/n$$

is a minimum, where n= number of location-years, and for the jth location-year,

x_j = Julian date when a given per cent of the crop is planted (x = P15, P50, P85)

$\hat{x}_j$ = estimated Julian date when a given per cent of crop is planted where $\hat{x}_j$ is the first Julian date when the W_i's sum to a constant $\hat{\gamma}_x$ (called a "cutoff values").

Various combinations of parameter values were tried in (3.1). The right hand side of (3.2) was evaluated and the average value over a set of location-years used as $\hat{\gamma}_x$. Evaluation of (3.3) for each set of parameters indicated which set gave the smallest RMSE. Comparison with results in Table B-1 indicated how well a variable planting date model performed relative to using a fixed date model for P15, P50, or P85.

Results for one set of parameters are given in Table B-2. The lower limit in (3.2) was equal to the NP15 for the associated CRD less 19 days. Cutoff values ($\hat{\gamma}_x$) of 11, 16, and 23 days were used for x = P15, P50 and P85; respectively.

Table B-2. Means and root-mean-square errors (RMSE) for estimated Julian dates when given percent of crop planted. (α_1 = 10, α_2 = 5, α_3 = 3.33, β = 2).

Per Cent	West Central		Central		Southeast	
Planted	Mean	RMSE	Mean	RMS	Mean	RMSE
15	249	9.8	270	7.6	273	7.3
50	259	12.7	278	7.5	281	10.3
85	272	10.4	292	10.8	294	12.7

Cutoff values were calculated for each district by accumulating workdays (W_i) up to actual P15, P50, and P85 dates in a given location-year, and averaging the results. Values of $\hat{\gamma}_x$ varied from 9-12 workdays for P15, 14-19 workdays for P50, and 19-26 workdays for P85 for the definition used in Table B-2. Five weather stations were used in each of the three western districts, five in each of the central districts and three in the three eastern districts. Hence, not finding a significant statistical relationship between per cent planting dates and weather was not due to sparseness of weather data.

A variety of other sets of parameter values were tried in an effort to improve on results in Table 2. While improvements were obtained in particular districts with particular parameter sets, the search for a more universal model was fruitless.

An effort was also directed toward an analytical solution to the problem of determining parameter values α_1, α_2, α_3, β, and γ_x to minimize

$$(\gamma_x - \sum_{i=(NP15)-19}^{x} W_i)^2$$

summed over location-years. A simplex algorithm was used but the estimated parameters often depended on the values used in the initial iteration. Convergence to a unique vector was rare even with $\alpha_2=\alpha_3=0$. When α_2 was allowed to vary in addition to α_1, β, and γ_x, solutions showed a very erratic pattern. As a positive contribution, this approach strongly indicated that a given minimum could be closely approximated by a wide range of vectors of parameter values.

4.0 DRYNESS-INDUCED PLANTING DELAYS

A linear regression model was constructed in a final attempt to isolate the dryness factor, a factor which should account for almost all of the long delays in planting in western Kansas. Three definitions for a dryness factor were tested along with two wetness factors. The models used were of the form:

$$P15 = NP15 + \beta_1(DRY) + \beta_2\ (WET) + \varepsilon$$

where NP 15 = mean 15% planting date in each district,

$$DRY = 1-(1-\beta\ CPR)^+ \text{ where } CPR = \sum_{i=NP15-19}^{NP15} PR_i,$$

$$WET = \sum_{i=NP15-19}^{NP15} (1-\ \alpha PR_i)^+$$

Beta values of 2.0 and 5.0 were tried in the definition of "DRY" along with a simple 0-1 dryness factor (DRY = 1 if CPR $\leq$ 0.2 inches and zero otherwise). Alpha values of 2.0 and 5.0 were tested in "WET". Similar models for P50 and P85 were constructed except that the precipitation was accumulated for only ten days before the mean planting dates, NP50 and NP85.

The best models for the three western districts, along with their standard errors of estimate (S.E.E.) were:

(4.1) $\widehat{P15} = NP15 + 3.83(DRY)$; S.E.E. = 5.66 where $DRY = 1-(1-2.0\ CPR)^{+}$.

Models using the other two "DRY" definitions gave slightly larger values of S.E.E.

(4.2) $\widehat{P50} = NP50 + 6.02(DRY)$; S.E.E. = 7.56

where $DRY = 1-(1-5.0\ CPR)^{+}$. Again, the other two "DRY" definitions gave slightly larger values of S.E.E.

(4.3) $\widehat{P85} = NP85 + 20.56 + 4.25\ (DRY)-2.31\ (\underline{WET})$; S.E.E. = 8.55

where $DRY = 1-(1-2.0\ CPR)^{+}$

$$WET = \sum_{i-NPLD_{.85}-9}^{NPLD_{.85}} (1-5.0\ PR_i)^{+}.$$

None of these models performed better than the "fixed date" model on the basis of a comparison of S.E.E. with the S.D. (standard deviation) values in Table B-1.

Our conclusion from this investigation is that the many factors involved which dictate when farmers, individually and collectively, plant wheat in a particular year overshadow the effects of precipitation on this decision. At least such is the case for the methods by which we have attempted to measure precipitation effects.

APPENDIX C

TABLES RELATED TO ESTIMATING DAILY MINIMUM
AND MAXIMUM TEMPERATURES FROM EIGHT
TRIHOURLY OBSERVATIONS

Table C-1a Sample means ($\bar{x}$) and standard deviations (s) of daily differences between model-generated specified hourly and measured minimum temperatures (in °F) at Chanute, KS. N ≅ 450.

Month		Clock Hour (h) 00	01	02	03	04	05	06	07	08
Jan	$\bar{x}$	7.6	7.0	6.5	6.1	5.8	5.3	5.0	4.7	5.2
	s	7.1	6.9	6.7	6.6	6.5	6.3	6.0	5.8	5.8
Feb	$\bar{x}$	7.3	6.6	5.9	5.4	4.7	4.1	3.8	3.5	4.5
	s	6.0	5.9	5.7	5.4	5.1	4.8	4.6	4.6	4.4
Mar	$\bar{x}$	7.1	6.2	5.5	4.8	4.2	3.7	3.4	3.8	6.1
	s	5.8	5.5	5.3	5.2	5.0	4.7	4.5	4.5	4.7
Apr	$\bar{x}$	7.0	6.1	5.3	4.4	3.9	3.4	3.3	5.0	8.1
	s	5.1	4.9	4.7	4.5	4.4	4.2	4.2	4.1	5.0
May	$\bar{x}$	6.0	5.0	3.9	3.5	2.8	2.2	3.1	5.7	8.5
	s	3.7	3.6	3.4	3.2	3.1	2.9	2.7	3.2	4.3
Jun	$\bar{x}$	5.5	4.5	3.7	2.9	2.3	1.8	3.0	5.7	8.6
	s	3.3	3.0	2.8	2.6	2.3	2.2	2.2	3.0	4.2
Jul	$\bar{x}$	5.6	4.6	3.8	3.0	2.2	1.6	2.4	5.2	8.3
	s	2.9	2.7	2.5	2.2	2.0	1.9	1.9	2.7	3.8
Aug	$\bar{x}$	6.2	5.2	4.2	3.3	2.5	1.9	1.9	4.5	8.3
	s	3.0	2.8	2.5	2.2	2.0	1.9	1.7	2.4	3.7
Sep	$\bar{x}$	7.0	6.0	5.1	4.3	3.5	2.7	2.2	3.8	7.7
	s	4.4	4.0	3.8	3.6	3.5	3.3	3.0	2.9	4.3
Oct	$\bar{x}$	7.3	6.4	5.5	4.5	3.8	3.2	2.8	3.3	6.9
	s	4.6	4.4	4.1	3.7	3.5	3.4	3.4	3.2	3.7
Nov	$\bar{x}$	7.5	6.7	6.2	5.5	4.9	4.4	4.1	3.8	5.6
	s	6.0	6.0	5.8	5.8	5.5	5.4	5.4	5.4	5.3
Dec	$\bar{x}$	7.2	6.5	6.0	5.5	5.2	4.8	4.6	4.2	4.7
	s	6.2	6.0	5.9	5.9	5.8	5.7	5.6	5.5	5.4

Table C-1b Sample means $\bar{x}$ and standard deviations (s) of daily differences between model-generated specified hourly and measured maximum temperatures (in °F) at Chanute, KS. N $\cong$ 450

Month		Clock hour (h) 09	10	11	12	13	14	15	16	17
Jan	$\bar{x}$	13.4	10.4	7.8	5.6	4.0	2.9	2.6	3.3	5.2
	s	7.3	6.0	5.1	4.7	4.6	4.9	5.7	5.4	5.6
Feb	$\bar{x}$	14.0	10.9	8.1	5.9	4.0	2.9	2.3	2.7	4.2
	s	7.4	5.5	4.3	3.6	3.3	3.3	3.8	4.2	4.6
Mar	$\bar{x}$	13.4	10.4	7.9	5.9	4.2	3.1	2.6	2.8	3.8
	s	6.4	5.1	4.2	3.6	3.3	3.4	3.7	4.1	4.6
Apr	$\bar{x}$	11.8	9.2	7.0	5.3	3.8	2.8	2.3	2.5	3.4
	s	4.9	4.2	3.5	3.1	2.7	2.7	2.9	3.4	4.0
May	$\bar{x}$	10.5	8.2	6.3	4.6	3.3	2.4	2.1	2.3	3.2
	s	4.5	3.5	3.0	2.6	2.3	2.2	2.5	2.8	3.2
Jun	$\bar{x}$	10.6	8.3	6.3	4.7	3.6	2.6	2.3	2.5	3.2
	s	3.5	3.2	2.8	2.6	2.9	2.8	3.1	3.5	3.6
Jul	$\bar{x}$	11.4	8.9	6.7	4.9	3.6	2.4	2.4	2.6	3.2
	s	3.4	3.2	2.9	2.6	2.6	2.1	3.6	3.6	3.8
Aug	$\bar{x}$	12.1	9.0	6.7	4.8	3.4	2.4	2.1	2.4	3.4
	s	3.3	2.8	2.4	2.3	2.3	2.2	2.3	2.8	3.5
Sep	$\bar{x}$	12.5	9.2	6.6	4.6	3.1	2.2	1.8	2.2	3.8
	s	3.7	3.1	2.8	2.4	2.4	2.6	2.7	3.1	3.7
Oct	$\bar{x}$	13.0	9.3	6.4	4.4	2.8	1.9	1.8	2.4	4.7
	s	5.1	3.6	2.9	3.2	2.2	2.8	2.7	3.1	3.2
Nov	$\bar{x}$	13.0	9.4	6.7	4.6	3.0	2.2	2.2	3.4	6.1
	s	6.0	4.4	3.5	3.3	3.4	4.4	4.1	4.3	4.5
Dec	$\bar{x}$	12.5	9.4	6.6	4.5	2.8	2.1	2.2	3.1	5.5
	s	6.4	4.8	3.8	3.2	3.1	3.4	3.6	4.2	4.6

Table C-2a Sample means ($\bar{x}$) and standard deviations (s) of daily differences between model-generated specified hourly and measured minimum temperatures (in °F) at Russell, KS. N ≅ 450

Month		Clock hour (h) 00	01	02	03	04	05	06	07	08
Jan	$\bar{x}$	8.5	7.4	6.7	6.4	5.6	5.2	4.9	4.3	4.5
	s	6.2	5.5	5.4	5.4	4.9	4.8	4.7	4.5	4.6
Feb	$\bar{x}$	8.3	7.5	6.7	5.9	5.3	4.9	4.5	4.1	5.0
	s	6.4	5.9	5.7	5.5	5.1	5.0	4.8	4.9	4.9
Mar	$\bar{x}$	8.1	7.0	6.2	5.6	4.8	4.2	3.7	3.8	6.2
	s	5.8	5.2	5.0	4.9	4.4	4.1	4.0	3.8	4.2
Apr	$\bar{x}$	8.3	7.0	6.0	5.3	4.3	3.6	3.2	4.8	8.3
	s	5.3	4.8	4.5	4.4	4.0	3.8	3.8	3.8	4.7
May	$\bar{x}$	7.4	5.9	5.0	4.4	3.3	2.7	3.1	5.7	8.8
	s	4.7	4.2	4.0	3.7	3.3	3.1	3.0	3.3	4.5
Jun	$\bar{x}$	6.8	5.7	4.5	3.7	2.9	2.3	2.8	5.4	8.8
	s	3.7	3.4	3.0	2.9	2.6	2.5	2.5	3.1	4.1
Jul	$\bar{x}$	7.1	5.9	4.8	3.6	2.7	1.9	2.2	5.0	8.5
	s	3.5	3.3	3.0	2.5	2.3	2.1	2.0	2.8	3.9
Aug	$\bar{x}$	7.5	6.1	5.0	4.0	3.1	2.3	2.0	4.0	7.7
	s	3.6	3.3	3.0	2.5	2.4	2.1	2.4	2.6	3.6
Sep	$\bar{x}$	7.9	6.7	5.7	4.9	4.0	3.1	2.6	3.5	7.3
	s	4.8	4.5	4.1	3.8	3.4	3.1	3.0	3.0	4.1
Oct	$\bar{x}$	8.8	7.4	6.4	5.6	4.7	3.9	3.3	3.4	6.7
	s	5.0	4.7	4.5	4.2	4.0	3.9	3.7	3.5	4.1
Nov	$\bar{x}$	8.4	7.2	6.5	5.9	5.2	4.6	4.1	3.7	5.2
	s	5.8	5.6	5.4	5.0	4.8	4.5	4.2	4.2	4.4
Dec	$\bar{x}$	8.0	7.1	6.5	6.2	5.6	5.2	4.9	4.5	4.8
	x	5.8	5.4	5.2	4.9	4.8	4.7	4.5	4.6	4.7

Table C-2b Sample means ($\bar{x}$) and standard deviations (s) of daily differences between model-generated specified hourly and measured maximum temperatures (in °F) at Russell, KS. N $\approx$ 450

Month		Clock hour (h) 09	10	11	12	13	14	15	16	17
Jan	$\bar{x}$	17.6	13.8	10.2	7.3	4.8	3.2	2.5	2.7	5.3
	s	9.2	7.4	5.8	4.8	3.8	3.7	4.0	4.0	4.5
Feb	$\bar{x}$	16.7	12.6	9.3	7.0	4.6	3.3	2.5	2.4	3.8
	s	8.9	7.1	5.6	4.6	3.8	3.6	4.1	4.2	4.5
Mar	$\bar{x}$	16.0	11.8	8.8	6.7	4.5	3.2	2.5	2.4	3.4
	s	8.4	6.4	5.0	4.0	3.2	3.0	3.0	3.4	3.7
Apr	$\bar{x}$	14.4	10.9	8.3	6.3	4.3	2.9	2.3	2.3	3.1
	s	6.0	4.8	4.0	3.3	2.8	2.3	2.6	3.2	4.1
May	$\bar{x}$	12.9	9.7	7.5	5.7	4.0	2.7	2.3	2.2	3.1
	s	5.2	4.2	3.5	3.0	2.6	2.5	2.4	2.7	3.6
Jun	$\bar{x}$	13.3	10.4	7.9	5.8	4.0	2.8	2.2	2.1	2.8
	s	3.9	3.3	2.8	2.6	2.1	2.4	2.2	2.9	3.6
Jul	$\bar{x}$	13.3	10.3	7.8	5.8	4.0	2.8	2.0	2.1	2.9
	s	3.8	3.3	3.0	2.8	2.5	2.3	2.0	2.9	4.0
Aug	$\bar{x}$	14.4	11.0	8.1	6.0	4.0	2.7	2.0	2.1	3.0
	s	4.1	3.7	3.3	2.7	2.3	2.4	2.5	3.1	3.9
Sep	$\bar{x}$	14.8	11.1	8.1	5.8	3.8	2.5	2.0	1.9	3.0
	s	5.0	3.9	3.4	3.2	2.6	2.4	2.5	2.7	3.0
Oct	$\bar{x}$	16.5	11.9	8.4	5.8	3.5	2.1	1.8	2.1	4.2
	s	6.4	4.6	3.5	3.0	2.5	2.3	2.7	2.8	3.4
Nov	$\bar{x}$	16.6	12.1	8.4	5.7	3.4	2.2	1.8	2.6	6.0
	s	8.3	6.3	5.0	3.7	3.4	3.3	3.2	3.8	4.3
Dec	$\bar{x}$	16.3	12.4	8.9	6.3	3.9	2.5	2.1	3.0	6.4
	s	8.9	6.7	4.9	3.8	3.1	3.0	3.0	3.5	4.1

Table C-3a Sample means ($\bar{x}$) and standard deviations (s) of daily differences between model-generated specified hourly and measured minimum temperatures (in °F) at Goodland, KS. N $\approx$ 450

		Clock hour (h)								
Month		00	01	02	03	04	05	06	07	08
Jan	$\bar{x}$	8.0	7.5	6.9	6.7	6.4	6.1	5.8	5.9	8.4
	s	6.5	6.2	5.9	5.7	5.5	5.4	5.5	5.3	5.3
Feb	$\bar{x}$	7.7	7.1	6.3	5.9	5.7	5.1	4.9	5.6	9.4
	s	5.8	5.5	5.4	5.2	5.0	4.7	4.9	4.6	5.3
Mar	$\bar{x}$	7.7	6.5	5.7	5.4	4.8	4.3	4.2	7.4	11.3
	s	5.7	5.0	4.6	4.3	4.0	3.9	3.9	4.7	6.6
Apr	$\bar{x}$	7.8	6.6	5.6	4.9	4.0	3.6	5.1	10.0	13.8
	s	5.3	4.6	4.3	3.9	3.5	3.5	4.0	5.3	7.1
May	$\bar{x}$	7.0	5.6	4.7	4.0	3.0	3.0	5.6	10.3	13.7
	s	4.7	4.1	3.5	3.4	2.8	2.8	4.1	5.8	7.1
Jun	$\bar{x}$	6.8	5.5	4.4	3.6	2.7	3.0	5.7	11.0	14.5
	s	4.1	3.5	3.0	2.9	2.6	2.6	4.1	5.2	6.0
Jul	$\bar{x}$	6.6	5.4	4.3	3.4	2.5	2.5	5.0	10.2	13.8
	s	3.6	3.2	2.7	2.3	2.2	2.2	3.4	4.4	5.5
Aug	$\bar{x}$	7.1	6.0	4.8	4.0	3.2	2.4	4.3	9.6	13.7
	s	3.8	3.4	2.9	2.7	2.4	2.2	3.0	4.2	5.5
Sep	$\bar{x}$	7.8	6.9	5.8	4.9	4.2	3.3	4.2	9.2	13.2
	s	4.9	4.3	4.1	3.6	3.3	3.2	3.5	5.0	7.1
Oct	$\bar{x}$	8.0	7.1	6.1	5.5	4.9	4.2	4.2	8.3	13.4
	s	5.0	4.7	4.5	4.2	4.1	3.9	3.7	4.5	6.7
Nov	$\bar{x}$	7.8	7.2	6.4	6.1	5.7	5.0	5.1	6.7	10.5
	s	5.6	5.2	5.0	4.5	4.4	4.1	4.1	4.4	5.5
Dec	$\bar{x}$	7.9	7.4	6.8	6.8	6.5	6.1	6.0	6.4	8.8
	s	5.8	5.4	5.0	5.1	4.8	4.7	4.8	4.9	5.2

Table C-3b Sample means ($\bar{x}$) and standard deviations (s) of daily differences between model-generated specified hourly and measured maximum temperatures (in °F) at Goodland, KS. N ≈ 450

Month		09	10	11	12	13	14	15	16	17
		Clock hour (h)								
Jan	$\bar{x}$	15.3	9.5	6.9	5.2	3.2	2.7	3.3	6.2	10.4
	s	7.5	5.2	4.4	4.4	4.1	4.2	4.6	5.4	5.7
Feb	$\bar{x}$	14.3	9.4	7.0	5.2	3.4	2.6	2.9	4.5	7.4
	s	7.1	4.6	3.9	3.8	3.6	3.5	3.9	4.7	5.0
Mar	$\bar{x}$	13.5	9.0	7.0	5.5	3.7	2.9	2.9	3.8	5.7
	s	6.8	4.5	3.8	3.5	3.0	3.1	3.6	4.1	4.5
Apr	$\bar{x}$	12.9	9.1	7.1	5.6	3.8	3.0	3.0	3.7	5.1
	s	5.8	4.3	3.5	3.3	2.6	2.9	3.3	3.8	3.9
May	$\bar{x}$	12.2	8.7	6.5	5.2	3.4	2.9	3.2	4.2	5.5
	s	5.3	4.0	3.2	2.9	2.3	2.6	3.1	4.2	4.6
Jun	$\bar{x}$	12.7	9.0	6.8	5.3	3.3	2.5	2.6	3.5	5.0
	s	4.6	3.7	3.1	2.9	2.3	2.2	3.1	4.1	4.5
Jul	$\bar{x}$	12.8	8.8	6.7	5.2	3.2	2.7	2.8	3.9	5.5
	s	4.0	2.9	2.6	2.5	2.2	2.6	3.6	4.6	5.3
Aug	$\bar{x}$	12.9	8.9	6.4	5.0	3.1	2.6	2.8	3.8	5.6
	s	4.3	3.1	2.6	2.5	2.4	2.5	3.3	4.0	4.6
Sep	$\bar{x}$	13.8	9.2	6.6	5.1	2.9	2.2	2.4	3.6	5.8
	s	5.2	3.6	2.9	3.1	2.2	2.1	2.6	3.3	3.7
Oct	$\bar{x}$	14.1	8.8	6.3	4.4	2.6	2.1	2.3	4.3	9.0
	s	6.2	3.7	3.1	2.9	2.5	2.6	2.6	3.4	4.5
Nov	$\bar{x}$	13.7	8.3	5.6	4.0	2.3	2.1	2.8	6.8	11.4
	s	7.0	4.4	3.3	3.0	2.5	2.9	3.1	4.2	5.3
Dec	$\bar{x}$	14.1	8.7	5.7	4.1	2.5	2.1	3.1	7.3	11.4
	s	7.2	4.5	3.5	3.0	2.5	2.8	3.3	4.2	5.4

Table C-4. Comparison of monthly means of model-generated estimates and actual daily maximum and minimum temperatures for Sioux Falls, SD. Entries are in °F. (h_o = 00)

		Maximum			Minimum		
		Model-Generated			Model-Generated		
Month	Year	Hour specified	Psuedo	Measured	Hour specified	Psuedo	Measured
March	49	37.5	38.9	38.3	23.0	22.7	22.2
	50	----	----	----	----	----	----
	51	28.1	29.0	29.6	12.8	11.9	11.0
	52	33.5	33.7	33.5	18.3	17.9	17.5
	53	42.7	43.0	42.8	24.2	23.9	22.8
	54	35.4	36.1	36.4	20.0	19.4	20.0
	55	39.8	40.3	40.7	19.1	18.6	17.4
	56	36.5	38.2	37.8	18.6	17.5	17.6
	57	41.4	41.6	41.7	21.0	21.9	21.0
	58	39.5	39.4	39.2	23.5	24.1	23.4
	59	43.9	44.8	44.8	25.6	25.8	25.5
	60	26.5	27.1	27.5	9.5	8.5	7.8
	61	43.5	43.9	44.0	27.3	27.6	27.4
	62	32.1	32.5	32.6	17.1	17.0	17.4
	63	49.3	50.2	49.7	29.4	28.5	28.1
	64	36.4	37.9	38.0	18.2	16.1	15.2
RMSE†		0.9	0.4		1.3	0.7	
Range of Differences††		-0.3 to +1.6	-0.6 to +0.6		-3.0 to +0.3	-1.2 to +0.6	

†RMSE = root-mean-square error

††Difference = measured - model

Table C-4. (continued)

		Maximum			Minimum		
Month	Year	Hour specified	Psuedo	Measured	Hour specified	Psuedo	Measured
June	49	----	----	----	----	----	----
	50	----	----	----	----	----	----
	51	72.1	72.6	73.0	50.4	50.9	51.4
	52	----	----	----	----	----	----
	53	82.1	82.9	82.9	58.5	58.6	58.8
	54	80.1	80.3	80.4	58.8	58.0	58.3
	55	76.6	76.6	77.2	54.2	54.7	55.0
	56	88.2	88.1	88.6	61.3	61.9	61.6
	57	77.5	78.5	78.2	54.7	54.9	55.2
	58	74.8	76.5	76.8	52.1	51.9	52.0
	59	83.8	84.3	84.1	60.7	60.9	60.7
	60	75.7	76.6	76.4	54.5	54.9	54.8
	61	80.0	80.8	80.9	56.4	56.5	56.7
	62	77.0	77.7	77.5	56.2	57.2	57.7
	63	84.8	85.2	85.0	60.9	61.7	62.0
	64	82.3	83.1	83.0	57.1	56.9	57.1
RMSE		0.8	0.3		0.7	0.3	
Range of Differences		+0.2 to +2.0	-0.3 to +0.6		-0.5 to +1.5	-0.3 to +0.5	

Table C-4. (continued)

		Maximum			Minimum		
Month	Year	Hour specified	Psuedo	Measured	Hour specified	Psuedo	Measured
Sept.	49	71.8	71.4	70.9	44.1	44.1	44.4
	50	----	----	----	----	----	----
	51	----	----	----	----	----	----
	52	79.5	78.6	78.9	49.6	50.1	50.0
	53	76.2	75.3	76.0	47.1	46.6	47.2
	54	74.2	73.3	74.0	50.6	50.5	50.9
	55	77.1	76.5	77.0	49.6	49.7	49.8
	56	76.6	76.5	77.0	46.3	45.8	45.3
	57	70.5	70.0	70.4	47.1	48.1	47.8
	58	77.5	77.7	77.7	51.8	50.5	50.1
	59	72.2	72.3	72.4	50.0	49.9	49.9
	60	74.5	73.9	74.0	51.5	51.7	51.6
	61	70.4	70.2	70.5	48.5	48.1	48.0
	62	71.4	71.4	70.3	47.2	47.4	47.4
	63	76.1	75.6	76.4	51.8	52.4	52.1
	64	70.6	71.0	70.5	47.9	47.6	48.0
RMSE		0.5	0.5		0.6	0.3	
Range of Differences		-1.1 to +0.4	-1.1 to +0.8		-1.7 to +0.4	-0.5 to +0.6	

Table C-4. (continued)

		Maximum			Minimum		
Month	Year	Hour specified	Psuedo	Measured	Hour specified	Psuedo	Measured
Dec.	49	----	----	----	----	----	----
	50	27.0	27.1	26.9	8.6	7.7	7.3
	51	22.8	23.5	23.8	6.8	5.0	4.2
	52	33.5	32.1	31.7	13.2	13.9	13.1
	53	29.5	31.0	31.6	13.1	10.5	10.7
	54	32.7	31.7	31.4	14.5	14.2	14.3
	55	21.2	21.4	22.7	4.6	2.0	2.8
	56	35.8	35.1	34.6	15.3	15.1	14.2
	57	41.0	40.3	38.3	20.5	19.8	18.9
	58	28.2	28.4	29.6	8.5	7.1	8.0
	59	37.9	37.1	36.4	22.2	22.3	22.2
	60	30.8	30.4	30.4	11.8	11.8	11.4
	61	----	----	----	----	----	----
	62	----	----	----	----	----	----
	63	----	----	----	----	----	----
	64	----	----	----	----	----	----
RMSE		1.5	0.9		1.4	0.7	
Range of Differences		-2.7 to +2.1	-2.0 to +1.3		-2.6 to +0.0	-0.9 to +0.6	

Table C-5. Comparison of mean values (°F) for reported and estimated daily minimum and maximum temperatures at three USSR locations in 1975.

	Location	WMO-number	Month	Reported	Model: Using psuedo min. and max.	Model: Using specified hour
MAX	Chernovtsy	33658	July	76.1	75.1	76.2
			Oct.	55.2	56.1	55.8
			Jan.	37.5	37.7	37.8
	Sverdlovsk	28440	Apr.	58.0	55.7	55.7
			July	78.0	74.5	75.0
			Oct.	37.5	37.2	36.8
	Kurgan	28661	Oct.	42.4	41.0	40.0
MIN	Chernovtsy	33658	July	58.9	54.4	58.9
			Oct.	41.4	37.3	38.3
			Jan.	27.6	23.8	22.2
	Sverdlovsk	28440	Apr.	38.0	35.9	37.2
			July	56.2	56.6	58.5
			Oct.	26.5	24.5	22.2
	Kurgan	28661	Oct.	24.8	22.8	20.7