25X1

13 MAY 1983

MEMORANDUM FOR: (See Addressees List)

FROM: [redacted] 25X1
Chief, Strategic Resources Division, OGI

SUBJECT: PRED: A Predictive Weather Model for Crop Forecasting [redacted] 25X1

1. The attached memorandum describes a new weather forecasting model which will be used primarily in future assessments of the Soviet grain crop. [redacted] 25X1

2. This paper was prepared by [redacted], Agricultural Assessments Branch, Strategic Resources Division, Office of Global Issues. [redacted] 25X1

25X1

3. Comments and queries are welcome and may be addressed to the Chief, Agricultural Assessments Branch, on [redacted] 25X1

[redacted] 25X1

Attachment:
PRED: A Predictive Weather Model for Crop Forecasting, May 1983, GI M 83-10129 [redacted] 25X1

[redacted] 25X1

[redacted] 25X1

25X1

SUBJECT: PRED: A Predictive Weather Model for Crop Forecasting [redacted]

25X1

Addressees List:

Mr. Elmer Klumpp
Special Assistant to the Under Secretary
International Affairs and Commodity Programs
Department of Agriculture

Dr. Terry N. Barr
Acting Chairman, World Agriculture Outlook Board
Department of Agriculture

Mr. Jimmy Murphy
Acting Assistant Administrator
International Agriculture Statistics
Department of Agriculture

Mr. Donald J. Novotny
Director, Grain and Feed Division
Foreign Agriculture Service
Department of Agriculture

Mr. David Schoonover
Director, Centrally Planned Economics Division
Foreign Agriculture Service
Department of Agriculture

Mr. Gerald A. Bange
Director, Foreign Production Estimates Division
Department of Agriculture

Mr. Anton F. Malish
Chief, USSR/Eastern Europe Branch
Economic Research Service
Department of Agriculture

Mr. Keith Severin
Foreign Production Estimates Division
USSR/Eastern Europe
Department of Agriculture

Mr. Frank Gomme
Foreign Agriculture Service
Department of Agriculture

Mr. John Danylyk
Bureau of Intelligence and Research
Office of Economic Analysis
Department of State

Mr. David Peterson
Director, Office of Intelligence Liaison
Department of Commerce

[redacted]

25X1

[redacted] 25X1

SUBJECT: PRED: A Predictive Weather Model for Crop Forecasting [redacted] 25X1

OGI/SRD/AA [redacted] (13 May 1983) 25X1

Distribution List: (Attachment with Each Copy)
Orig - Each Addressee
1 - D/GI
1 - DD/E
1 - DD/SOVA
1 - C/SOVA/SE
1 - SOVA/SE/R [redacted] 25X1
1 - C/ECD/CM
1 - C/SRD
4 - C/SRD/AA
8 - OGI/PS

[redacted] 25X1

25X1

Central Intelligence Agency

Washington, D.C. 20505

DIRECTORATE OF INTELLIGENCE

23 MAY 1983

Pred: A Predictive Weather Model For Crop Forecasting [redacted] 25X1

Summary

The estimation of grain output in the USSR will be significantly improved this year by the introduction and use of the PRED model. In the past the estimation of crop yield has been handicapped by the lack of a reliable predictive weather input to UPSTREET's agronomically based, weather driven, crop-growth simulation model. In the absence of this essential data, analysts could only forecast the maximum yield. This required an evaluation of the impact of elapsed weather on crop yield and the assumption that ideal growing weather would prevail for the remainder of the season. With the completion of PRED, which has the capacity to produce representative weather data for an entire season, the acceptance of the ideal weather assumption is no longer necessary. Furthermore, the accuracy and utility of the assessments produced during the early season are enhanced. In this report the use of the PRED model, as an adjunct to the UPSTREET analytical effort, is examined and evaluated. [redacted] 25X1

This paper was prepared by the [redacted] *Agricultural Assessments Branch, Strategic Resources Division, Office of Global Issues (AAB/SRD/OGI). Comments and queries may be addressed to the Chief, AAB/SRD/OGI* [redacted]

25X1

25X1

[redacted] 25X1

[redacted]

GI M 83-10129

25X1

PRED: A Predictive Weather Model for Crop Forecasting

25X1

PRED was developed to provide predictive weather inputs to the crop model used in the UPSTREET[1] analysis of grain production in the USSR. By its calculated employment of historical weather data to determine the likely character of future weather, beyond the date on which the crop model is run, PRED provides information that has previously been unavailable. By generating a variety of representative weather scenarios, an average, (i.e. best estimate) crop yield can be determined as well as information regarding the uncertainty of this estimate. This approach enhances the applicability of the UPSTREET methodology and lends credence to assessments of crop production made early in the growing season.

25X1

The following sections describe the PRED model. Section I describes how the historical data is processed and details the creation of 200 independent, season long weather scenarios. Section II describes how a sample of weather scenarios (typically 25) are selected from this collection of 200 to reflect specified long term weather projections; Section III then describes the calculation and analysis of summary weather and crop-yield statistics.

25X1

Section I: Data Collection and Assembly

The first section of PRED compiles historical weather data[2] to provide a base for generating parameters to describe weather conditions and develop weather scenarios. The historical weather data base includes at least eight years of information on daily maximum and minimum temperatures, precipitation, and evapotranspiration (ETP)[3] for up to 3000 locations.

25X1

To create future weather scenarios, the historical data are processed according to a series of statistical routines which establish probabilities, expressed as curves[4], for the following weather parameters:

- o occurrence of precipitation

---

1 UPSTREET - The UPSTREET methodology is an interdisciplinary approach to grain production forecasting developed by CIA. It was first employed to forecast Soviet grain production in 1975. The methodology is dynamic and improvements are made nearly every year.

25X1

2 The historic weather data is obtained from ETAC, the Environmental Technical Applications Center, of the US Air Force.

25X1

3 ETP - The potential amount of moisture lost because of the combined effects of evaporation from the earth's surface and transpiration from leaves and other parts of plants.

25X1

4 These are actually cumulative probability curves with, for example, maximum temperature values on the x-axis and cumulative probability values ranging from 0 to 1.0 on the y-axis.

25X1

25X1

25X1

- o amount of precipitation
- o maximum temperature level
- o minimum temperature level
- o ETP

There are two sets of probability curves for the temperatures and ETP: one for days with precipitation and one for days without. For maximum and minimum temperature and ETP, probability curves are derived for each half-month period in the year. For precipitation amount, a probability curve is derived for each month.

25X1

From these probability curves, 200 weather scenarios are created. The occurrence of precipitation, temperature, precipitation and ETP values for each day are selected from the probability distributions using a random number generator. It is assumed that these 200 scenarios are a representative sample of all possible weather situations, given past history.

25X1

Section II: Selection of Scenario Subsets

In the second section of PRED, a subset of scenarios, usually 25, is selected to represent the weather at each location and to serve as input to the crop model. Each scenario created in Section I of PRED is characterized as above normal, normal, or below normal in terms of both precipitation and temperature.

25X1

Scenario selection is initiated by the user who provides forecasts for each location. These are usually monthly forecasts but can be for any interval of time. The user can derive forecasts independently or use forecasts that are available from several universities and the National Weather Service Long Range Predictive Group. Each forecast consists of four declarations:

- o temperature - below, near, or above normal
- o confidence - a subjective probability that the temperature will actually fall into the above categories
- o precipitation - below, near, or above normal
- o confidence - a subjective probability that the precipitation will actually fall into the above categories.

Since weather is never totally predictable, this type of declaration lets some of the 25 scenarios chosen fall outside of the forecasted precipitation and temperature categories.

25X1

Table 1 presents some examples of scenario selection for different forecasts. For example, if the forecast is for near normal temperatures and near normal precipitation, with 90% confidence levels for both forecasts, then the vast majority of the scenarios selected will be normal in both temperature and precipitation. With the same forecast but only a 60% confidence level on both forecasts, only five of the scenarios selected will be normal in both temperature and precipitation. As the confidence level decreases, the dispersion in the scenarios selected increases.

25X1

25X1

Table 1

Examples of 25 Scenarios Selected for 3 Different Forecasts

Example	Forecast	Number of Scenarios in Each Temperature/Precipitation Class				
				Temp		
				B	N	A
Example 1	Near Normal Temperature (N)		B	0	2	0
	Near Normal Precipitation (N)	Precip	N	1	19	1
	with 90% Confidence on both		A	0	2	0
				Temp		
				B	N	A
Example 2	Near Normal Temperature (N)		B	2	3	2
	Near Normal Precipitation (N)	Precip	N	3	5	3
	with 60% Confidence on both		A	2	3	2
				Temp		
				B	N	A
Example 3	Above Normal Temperature (A)		B	2	1	3
	Above Normal Precipitation (A)	Precip	N	2	2	3
	with 60% Confidence on both		A	3	4	5

B = Below Normal
N = Near Normal
A = Above Normal

25X1

Section III: Utilization of PRED Output

Once the sample of weather scenarios has been selected for each location, it is simply a matter of appending them, one at a time to observed weather to date, and running the crop model. The results are 25 (or whatever number was selected) end-of-season percent of maximum-yield-remaining estimates for each location. PRED averages crop yields over all scenarios to give a best estimate of the final yield at that location. The average yield is given along with one standard deviation above and below the average (68% confidence intervals).[5]

25X1

Applications and Limitations

PRED significantly enhances the capacity of UPSTREET-type analysis by developing in a rational and objective way, predictive weather which can be used to drive the UPSTREET crop model. PRED is not restricted in use to the USSR; it can be used in any country for which historical weather data is available. In addition, it also has potential for use with CROPCAST, a generic version of the UPSTREET crop model, also used primarily for grain estimates; with regression-type models designed for the assessment of grains and non-grains; and with the Ritchie generic crop growth model which can be employed to forecast yields of a wide variety of crops.

25X1

PRED also can be used to assess the effect of weather on crops in any given season. In this operational mode the user inputs a forecast of average temperature and precipitation for the entire growing season. The resulting yield values can be compared with yield values obtained by running the crop model on actual weather for the full growing season.

25X1

Furthermore, PRED has the potential to project the range of grain yields within established time frames by forecasting extremes in weather, i.e., in seasons that are excessively dry or very cool. It does not have the capacity to forecast short-term weather conditions, however, and it can neither anticipate such phenomena nor accept forecasts of the likelihood of their occurrence.

25X1

Finally, PRED is structurally limited to 40 stations or locations per run, but multiple runs can be made. In an area as large as the USSR, it will therefore be most useful when applied to key indicative areas.

25X1

---

[5] See the appendix for examples of output.

25X1

25X1

25X1

Appendix 1

This appendix describes the functions and interrelationships of files and programs that collectively comprise the PRED model. It will be most valuable when used in conjunction with the diagram of the entire PRED model (see figure 1). Sections I, II and III are divided by dotted lines on the figure. The crop model and random number generator module shown on the figure are exterior routines. Appendix 2 describes the files and their required storage space.

25X1

Section I of PRED compiles historical weather data and arranges it for use in developing weather scenarios. It contains a total of six files (the Historic Data Base file; the Historic Monthly file; the Monthly Check file; the Station Statistics file; the Scenario file; and the Scenario Attributes file) and three programs (MONFILE 2, CLIMSTAT, and SCENARIO).

25X1

Section II of PRED selects predictive weather for input to the crop model. Within this Section there are three files (the Forecast Input file; the PREDMET file; and the PREDMET Output file) and one program (PREDMET).

25X1

Section III identifies the locations in the crop model that correspond to the PRED stations, and provides PRED weather to the crop model so that it can compute final yields. This Section contains a total of six files (the PRED Cell List; the Yield Strip file; the Last Day Model file; the Last Day Strip file; the Last Day PRED file; and the LOSS file) and the PREDSTRPR (PRED Stripper) and LOSS programs.

25X1

Historical Data Base

This file is comprised of daily weather records of maximum and minimum temperatures (in degrees C), precipitation amounts (in millimeters) and evapotranspiration (ETP) rates (in millimeters) for a maximum of 3,000 locations. Stations for which less than eight years of data are available are not included. The data, arranged chronologically, are derived from reports generated by stations participating in the World Meteorological Organization (WMO) reporting program; they are quality controlled and reformatted for agricultural applications by the US Air Force Environmental Technical Applications Center. The data base includes information on the USSR dating from 1968.

25X1

Historical Data Base Variable Description

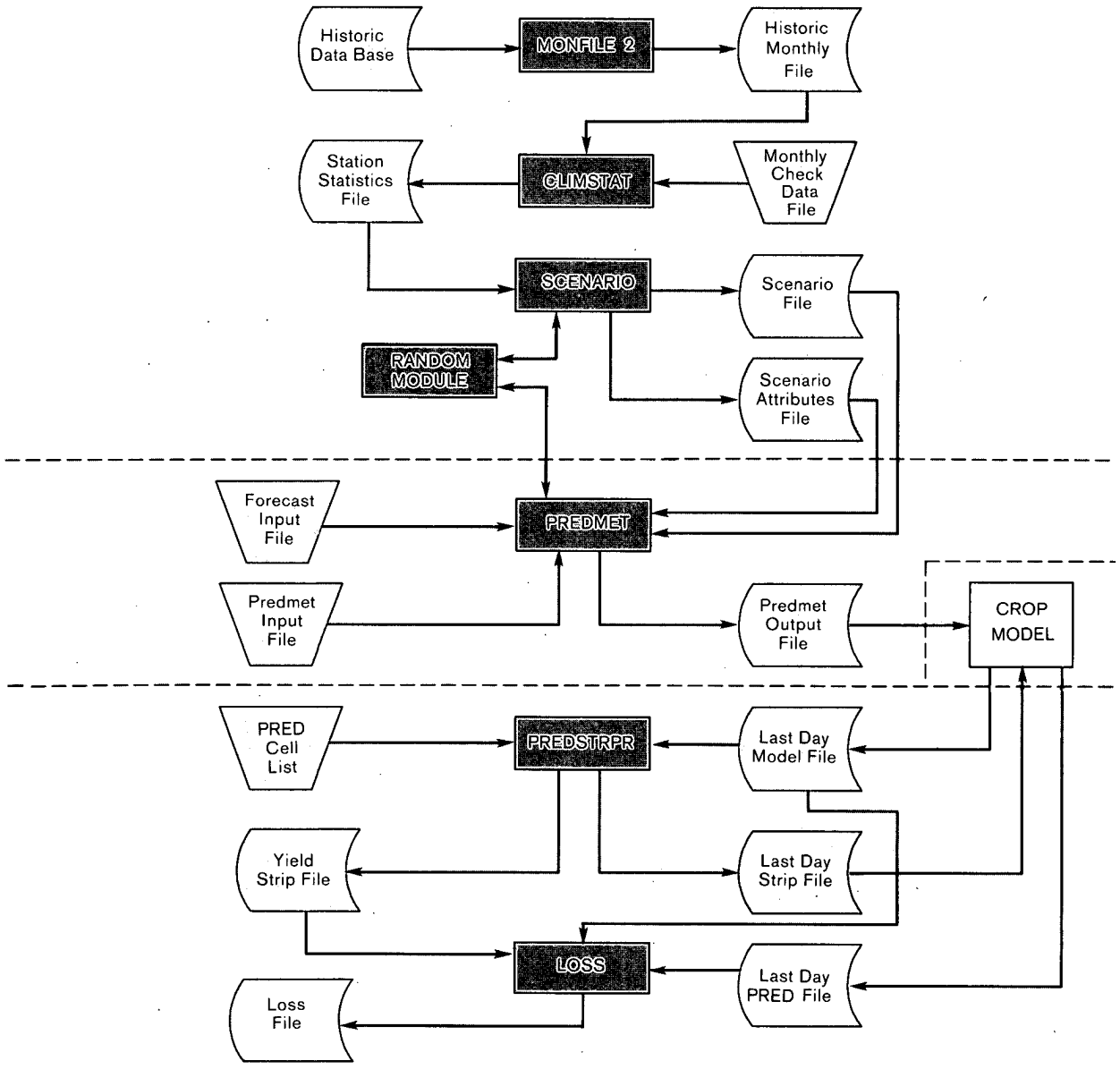
Station Number
Year
Julian Day
ETP (MM)
Precipitation (mm)
Maximum Temperature (C)
Minimum Temperature (C)

Monfile 2 Program

The MONFILE 2 program reads the Historical Data Base file, groups and

25X1

Figure 1
PRED Flow Diagram



589449 4-83

reformats the information into monthly increments, and stores the results in the Historic Monthly file.

Historic Monthly File

The Historic Monthly file is the repository of data reformatted by the MONFILE 2 program. In this file the data for each location are grouped into monthly records, which become input to the CLIMSTAT program.

Monthly Check Data File

The Monthly Check Data file is built by the user and contains information used to control operation of the CLIMSTAT program. The information includes: the station number, the total number of months of weather data in the station's data base, and a processing flag which indicates if the station is to be processed.

CLIMSTAT Program

The CLIMSTAT program reads the Monthly Check Data file to establish, for each half-month in the year, precipitation, temperature, and evapotranspiration parameters for each station. The probability of precipitation for a certain day is defined by a first order Markov chain which expresses the probability of precipitation as a function of whether or not precipitation occurred on the previous day. The values of temperature and ETP are assumed to follow normal distributions that are defined by a mean and standard deviation computed from historical data. The amount of precipitation is assumed to follow an Incomplete Gamma distribution defined by two parameters--gamma and beta. For maximum and minimum temperatures and ETP, distributions are defined for each half-month; for precipitation amount, the resolution is at the whole month level.

Station Statistics File

This file receives and stores the output of the CLIMSTAT program, in which one set of statistics per station is produced. Each set of statistics contains climatological parameters for each station, and these parameters (a total of 16) are organized by half-month increments.

Station Statistics File Parameters

Precipitation Probabilities for the Next Day:

1. Probability when previous day had precipitation.
2. Probability when previous day had no precipitation.

For a Precipitation Day:

3. Maximum temperature - mean
4. Maximum temperature - standard deviation
5. Minimum temperature - mean
6. Minimum temperature - standard deviation
7. ETP - mean

 25X1

8. ETP - standard deviation
9. Gamma -- (Used to define the Incomplete Gamma Function
10. Beta -- that describes the amount of precipitation).

For a Dry Day:

11. Maximum temperature - mean
12. Maximum temperature - standard deviation
13. Minimum temperature - mean
14. Minimum temperature - standard deviation
15. ETP - mean
16. ETP - standard deviation

Scenario Program

The Scenario program uses data in the Station Statistics file, in conjunction with a random number generator, to build 200 season-long weather scenarios for each station. Random numbers are used to translate the parameters from CLIMSTAT into weather outcomes for each day. The program also builds the Scenario Attributes file, which contains the day-to-day accumulations of temperature and precipitation for each scenario. [redacted] 25X1

Scenario File

The Scenario file is created by the SCENARIO program and is the repository of the 200 season-long weather scenarios produced by the latter. This very large file is used by the PREDMET program. To save space the records contain no information on stations or date; hence PRED software must be provided with accurate information on the sequence of stations as well as the start and end dates of the scenarios. The user must keep separate records on the order of data in this file. [redacted] 25X1

Scenario Attributes File

The Scenario Attributes file contains accumulations of temperature and precipitation for each scenario in the Scenario file. The records in this file are used by the PREDMET program to evaluate relative warmth and wetness of the developed scenarios in the scenario selection process. In addition, the Scenario Attributes file allows the user to make comparisons of any portion of a scenario with a like portion of any other scenario. [redacted] 25X1

Forecast Input File

The Forecast Input file is created by the user and contains weather forecasts for each station for specified time periods. The forecasts indicate whether temperature and precipitation will be below average, average, or above average. Forecasts can be given for a maximum of 11 time intervals. A confidence level is also established for each forecast. This file contains a temperature forecast code, a temperature confidence level, a precipitation forecast code and a precipitation confidence level for each of the 11 forecast intervals. [redacted] 25X1

[redacted] 25X1

Forecast Input File Description
3-Month Forecast - 3 Locations

Location	Month	Temperature Forecast-Confidence	Precipitation Forecast-Confidence
1	1	A 90%	N 80%
1	2	N 70%	N 70%
1	3	N 60%	N 60%
2	1	B 80%	A 80%
2	2	B 70%	N 70%
2	3	N 60%	N 60%
3	1	N 90%	A 90%
3	2	N 70%	N 70%
3	3	N 70%	A 70%

A = Above Normal
N = Normal
B = Below Normal

PREDMET Input File

The PREDMET Input file is created to control the operation of the PREDMET run. These controls establish starting and stopping dates for each run, state the number of iterations to be included in the run, define the forecast intervals to be employed, and list the stations that are to be assessed. The file is updated by the user as needed.

Random Module

The Random Module uses random numbers from 0.0 to 1.0 as probabilities to determine whether or not precipitation occurs, the amount of that precipitation, maximum and minimum temperatures, and ETP for each day in building a year-long scenario. It is also used in selecting the subset of scenarios by the PREDMET program.

PREDMET Program

The PREDMET program uses the Forecast Input file, PREMDET Input file, Scenario file, Scenario Attributes file, and the Random Module to control the selection of scenarios covering the period that extends from a selected date to the end of the growing season or other selected stopping date. The 25 scenarios thus selected for each location will be biased toward the forecast but with a statistically meaningful sample spanning a broader range of conditions. Program output is stored in the PREDMET Output file and serves as the meteorological input for the crop model.

PREDMET Output File

The PREDMET Output file is structured exactly like the meteorological data input file (MET) used by the crop assessment model. Information in it can thus be used by the crop model without change. This file contains the daily meteorological data of the 25 scenarios selected for each location.

 25X1

Each record in this file contains a day's worth of meteorological information for a single location and one of the 25 iterations. 25X1

Last Day Model File

This file is created by the crop model, using actual weather up to the day PRED is run. PRED uses this file to generate the Last Day Strip file and, via the crop model, the Last Day PRED file. 25X1

PRED Cell List

The PRED Cell List is a control file that matches the selected locations in PRED with locations in the crop model.[6] 25X1

PREDSTRPR Program

The PREDSTRPR program (PRED Stripper) reads the output of the Last Day Model file and selects records that pertain to the locations specified in the PRED Cell List. It creates the Last Day Strip file, which inputs to the crop model, and the Yield Strip file, which inputs directly to the LOSS program. 25X1

Last Day Strip File

The Last Day Strip file, generated by PREDSTRPR, is identical to the Last Day Model file except that it also contains an iteration number that corresponds to the weather scenario being processed. It is input to the crop model. 25X1

Yield Strip File

This file contains the estimated percent of yield remaining, as of the beginning day on which the crop model is run on PRED data, at each crop model location. 25X1

Last Day PRED File

The Last Day PRED file is the output Last Day file derived from running the crop model on PRED data. It contains end-of-season values, expressed in yield percent, for each location and each weather scenario. This is input to the LOSS program. 25X1

LOSS Program

The LOSS program computes the mean and standard deviation of the results of the Last Day PRED file. When weather data are not available for all locations in a region, the results of the Last Day PRED file are used in conjunction with the Yield Strip file to determine the mean incremental change in yield and its standard deviation. These results are assigned to locations using a nearest neighbor assignment scheme. This output forms the LOSS

---

6 The crop model runs on a gridded cell format. 25X1

25X1

file. [redacted]

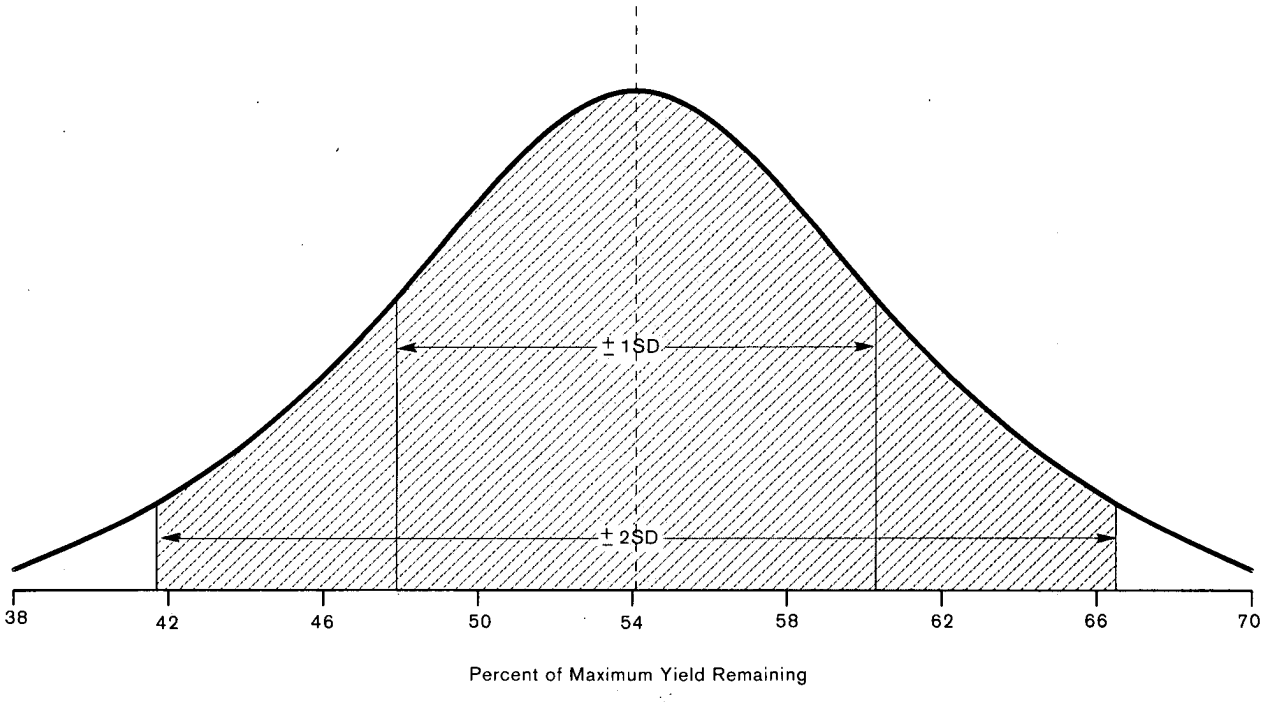
LOSS File

The LOSS file is the output mean and standard deviation for all locations in the LOSS program. Figure 2 demonstrates these results for one location. To illustrate, this average end-of-season crop yield for spring wheat is given to be 54.1% of the historical maximum yield for location 1. Some weather scenarios determined crop yields that were less than this, some more, as indicated by the spread in the distribution. The actual standard deviation (SD) was calculated to be 6.2, so that a 95% confidence interval about 54.1% would be 54.1% $\pm$ 12.4% (i.e., $\pm$ 2 standard deviations). The interval 41.7% to 66.5% will contain the true or actual yield 95% of the time. Table 2 contains a sample output for 30 locations. [redacted]

[redacted]

25X1

Figure 2
Distribution of Crop Yield as a Percent of Maximum for Location 1



25X1

589450 4-83

25X1

25X1

Table 2

Estimated Crop Yield As A Percentage of Total
With One Standard Deviation Limit

25X1

Location	Spring Wheat		
	-1SD	AVG	+1SD
1	47.9	54.1	60.3
2	59.8	66.3	72.8
3	54.4	58.7	63.0
4	43.6	50.0	56.4
5	52.6	56.3	60.0
6	66.8	70.7	74.6
7	63.3	68.8	74.3
8	45.3	48.8	52.3
9	57.0	61.8	66.6
10	55.2	60.4	65.6
11	64.0	67.5	71.0
12	53.5	58.7	63.9
13	66.9	72.3	77.7
14	44.7	47.2	49.7
15	69.8	73.0	76.2
16	68.9	73.5	78.1
17	72.1	76.1	80.1
18	69.1	73.8	78.5
19	63.9	67.3	70.7
20	62.4	66.2	70.0
21	81.1	84.5	87.9
22	50.6	54.0	57.4
23	55.6	57.9	60.2
24	61.0	65.0	69.0
25	81.8	85.3	88.8
26	73.1	76.9	80.7
27	78.3	81.5	84.7
28	73.4	78.3	83.2
29	68.5	74.1	79.7
30	69.3	73.4	77.5

25X1

25X1

25X1

Appendix 2

PRED SYSTEM FILES

Description Files	Created By	Used By	Medium	Storage (Bytes)
Historic Data Base	External	MONFILE2	Tape	8000/location/year
Historic Monthly	MONFILE2	CLIMSTAT	Tape	6000/location/year
Monthly Check Data	User	CLIMSTAT	Disk	14/location
Station Statistics	CLIMSTAT	SCENARIO	Tape or Disk	1536/location
Scenario	SCENARIO	PREDMET	Tape	600,000/location
Scenario Attributes	SCENARIO	PREDMET	Tape	300,000/location
Predmet Input	User	PREDMET	Disk	1000
Forecast Input	User	PREDMET	Disk	200/location
Predmet Output	PREDMET	Crop Model	Tape or Disk	28/location/iteration/day
Pred Cell List	user	PREDSTRPR	Disk	7/location
Yield Strip	PREDSTRPR	LOSS	Disk	32/location
Last Day Strip	PREDSTRPR	Crop Model	Disk	264/location/iteration
Last Day Pred	PREDMOD	LOSS	Disk or Tape	264/location/iteration
Last Day Model	Crop Model	PREDSTRPR/LOSS	Tape	264/crop model locations
Loss	LOSS	External	Disk or Tape	92/crop model locations

Notes: number of locations : 20-40
number of years : 10-15
number of iterations: 25-50
number of days : 30-300

25X1