United States
Department of
Agriculture

Statistical
Reporting
Service

Statistical
Research
Division

February 1983

Large Area USSR Barley-Yield Models: Development and Evaluation

IN COOPERATION WITH THE ATMOSPHERIC SCIENCE DEPARTMENT OF THE UNIVERSITY OF MISSOURI

Felix N. Kogan

LARGE AREA USSR BARLEY-YIELD MODELS:

DEVELOPMENT AND EVALUATION

by

Felix N. Kogan
Agricultural Meteorologist

This research was conducted as part of the AgRISTARS* project. It is part of task 2 in major project element number 2 as identified in the 1983 Yield Model Development Project Implementation Plan. As an internal project document, this report is identified as shown below.

---

*AgRISTARS is an acronym for Agriculture and Resources Inventory Surveys Through Aerospace Remote Sensing. It is a multi-agency research program to meet some current and new information needs for the U. S. Department of Agriculture.

AgRISTARS
Yield Model Development
Project

YMD-2-2-3(83-02.1)

LARGE AREA USSR BARLEY-YIELD MODELS: DEVELOPMENT AND EVALUATION. By Felix N. Kogan, Statistical Research Division, Statistical Reporting Service, U. S. Department of Agriculture*, Columbia, Missouri 65201. February 1983. SRS Staff Report No. AGES830121.

ABSTRACT

Barley yield models were developed and tested for five economic regions of the USSR. The aim of model development was early prediction of average barley yields for large areas. A regression approach was used for model development. This approach was modified in an attempt to more accurately represent trend in the historic yield series and incorporate fully the impact of weather variables on large area yield fluctuations. Weather was represented in the form of index variables which were designed to more accurately reflect the weather-related variability of yield. Application of the modified approach was shown to improve the accuracy of yield forecasts by both dependent and independent tests of the performance of the developed models.

Key words: Barley yield, yield model development, model evaluation, index variable, trend, weather.

ACKNOWLEDGEMENTS

Wendell Wilson (USDA) has given generously of his time to review this manuscript and has made many valuable suggestions for its improvement. For his work he deserves and has my profound appreciation. I wish also to thank my colleagues, Professor Wayne Decker (UMC), Jeanne Sebaugh, Merritt Padgitt, Fred Warren (USDA), Clarence Sakamoto and Sharon LeDuc (NOAA) for their very valuable comments. Special recognition should also be given to Galen Hart (USDA) for fruitful discussions on the problems touched on in this report and to Margaret Weidenhamer (USDA) who conducted a very beneficial review of the report. My profound appreciation is given to Jean Sparks for typing the manuscript and to Jerry Wright for preparation of the figures.

*Research on this task was conducted cooperatively by the Statistical Reporting Service (SRS) and the Atmospheric Science Department, University of Missouri-Columbia (UMC) under Research Agreement No. 58-319T-0207X.

TABLE OF CONTENTS

List of Figures

List of Figures

List of Tables

List of Tables

LARGE AREA USSR BARLEY-YIELD MODELS: DEVELOPMENT AND EVALUATION

Felix N. Kogan

SUMMARY

Large area barley yield models were developed for five economic regions of the USSR; these are South-West, Belorussia, Volgo-Vyatka, North Caucasus, and Kazakh regions. These regions were selected both because they are important in the production of barley and because they represent a wide range of growing conditions. The models calculate predicted average yield of barley for each region from the estimated quantitative impact of technological improvements and of fluctuations in weather on yield. The only current data required to operate these models would be the daily weather reported through the World Meteorological Organization network. Modified bootstrap tests of the models over an eight-year period (1971-78) resulted in relative mean squared errors (MSE) from 6.4 percent (South-West) to 20.4 percent (Volgo-Vyatka). (The relative MSE's for the other regions were between 11 and 12 percent.)

The models developed allow Soviet regional barley yields to be predicted with a lead time of two or three months in advance of the barley harvest. Independent testing of the models showed good correspondence between predicted and actual barley yields for the 1971-1978 period.

The impact of technological improvements was approximated by a non-linear function of year number as a surrogate for technological changes in the barley yield series. The impact of weather was also approximated by a non-linear function of actual total monthly precipitation and average monthly air temperatures in each region.

Since the analogue modeling approach has some drawbacks, a special technique to overcome these drawbacks was developed. This technique involved a partially empirical approach to more accurately represent technology and weather in the last years of the yield series. This approach was also utilized to reduce the number of weather variables in order to eliminate some discrepancies between a shortage of historical data for modeling and the existence of collinearity in the crop weather system.

INTRODUCTION

The objectives of this study were first to determine if existing models could be used for estimating the yield of barley in the USSR, and if not, to develop suitable models. This subject is of importance because barley is the most important grain crop in the USSR, and the size of the crop in any one year has great impact on the USSR requirements for imported feed grains.

The USSR is the largest producer of barley in the world. A quarter of the world's barley production is normally grown by Soviet farmers. Thus, the global barley supply depends to a great degree on the amount of barley produced in the USSR.

In the past 10 years, Soviet spring barley production ranged from 33 to 67 million metric tons. Much of this variation in production levels can be attributed to increased weather fluctuations [13]. The most recent striking example is the reduction in barley output during the past three years (1979-1981).

To overcome deficiencies between domestic grain supply and demand, the USSR has substantially increased purchases of grain, particularly feed grain, on the international market since the early 1970's. Overall these purchases are gradually increasing, but they fluctuate from year to year, depending on production of grain in the USSR. Therefore, timely forecasts of the Soviet barley production could be very useful for assessing global barley production, world trade in barley and other grains, and some perspectives for the future. These forecasts will be useful also for the assessment of Soviet domestic resources for the development of agriculture, particularly livestock production.

The interest in forecasting USSR barley production stems from the two priority areas of the Agriculture and Resources Inventory Surveys Through Aerospace Remote Sensing (AgRISTARS) Joint Agency program. These priority areas are "Foreign Commodity Production Forecasting" and "Early Warning/Crop Condition Assessment." These areas have been given emphasis because of the immediate need for better and more timely information on crop condition and expected production in the major grain producing countries of the world. Yield models which will provide early, mid and late season (harvest) yield estimates would help USDA to assess the potential impact of weather events throughout the crop season [38, 39].

A thorough literature search was conducted for available models that could be used for forecasting USSR barley yields. (These models are described in Appendix I.) Since the search was not successful in terms of the possibility of using these models for forecasting USSR barley yields, a program of developing barley yield models was initiated. The goals were to develop simple and low cost models for large area (at least economic region) use, which could provide timely (2-3 months in advance of harvest) assessment of barley yields. The models were required to reflect the influences of environmental conditions and improvement of technology on barley. They were also required to be scientifically and statistically reliable.

This report consists of five sections and three appendices. The introductory part describes the formulation of the problem in terms of AgRISTARS' tasks. Descriptions of historical data and methodology for model development are shown in the section on Crop Yield Modeling and the Available Data Base. Particular attention is given to the problem of defining an optimal way of reflecting the effect of technological and environmental factors on yield fluctuation and to problems connected with the application of regression analysis to the limited samples from historical data. Specific features of

Soviet soil, climatic and weather conditions are presented in the section on Peculiarities of the USSR's Natural Resources. Dynamics of the yield series in terms of reflecting technological and climatic variability and numerical assessment of weather-yield interaction are described in the section on Development of Barley Yield Models. The results of model testing and evaluation are presented in the last section. The three appendices contain (I) a review of Russian literature on barley yield model development, (II) an overview of model development and use and (III) a discussion of some considerations in the application of indices for yield modeling.

CROP YIELD MODELING AND THE AVAILABLE DATA BASE

The analogue approach has been chosen to apply in this study for development of barley yield models. This approach was the most appropriate for the tasks set in this study in terms of optimal combination of the advantages which are discussed in Appendix II. In developing these models, consideration was given to the physiological requirements of barley and to specific features of the geographic and climatic conditions of the studied regions. Special attention was given to overcoming some of the limitations of the analogue modeling approach.

Limitations of Analogue Modeling Approach

Normally, in the analogue approach for modeling crop yields, we encounter some problems (in addition to those mentioned in Appendix II)which can be overcome to some extent. These additional problems result from the lack of historical data, limitations of the statistical tool, lack of knowledge about crop-environment interactions and lack of accuracy in measuring crop and environmental variables. Some of these problems make model development difficult and do not allow us to fully rely on the results of modeling. Three of these problems are very important. They are multidimensionality (large number of influence factors) of crop-environment systems, multicollinearity (natural correlation between various environmental factors), and the approximation of trend. An impression of the multidimensionality problem in crop-environment systems can be obtained from Figure 1, which illustrates the large number of factors affecting crop productivity.

In dealing with multidimensionality, scientists normally try to increase the size of a sample and/or decrease the number of variables entering the statistical analysis. In the present case, increasing the size of the sample is not feasible, since only one barley yield observation is available annually and the series of years with reliable data is limited. Another way of artificially increasing the sample size is through the application of the cross-sectional method, which involves taking several sets of weather variables from the same (usually geographically homogeneous) region. It also has shortcomings. If several sets of weather data are matched with the same yield value, the resulting increase in degrees of freedom is fictitious, and there is no real improvement in precision. The elements of the sample formed by combining several subsamples are highly interdependent due to strong correlation between weather variables measured at neighboring locations. This

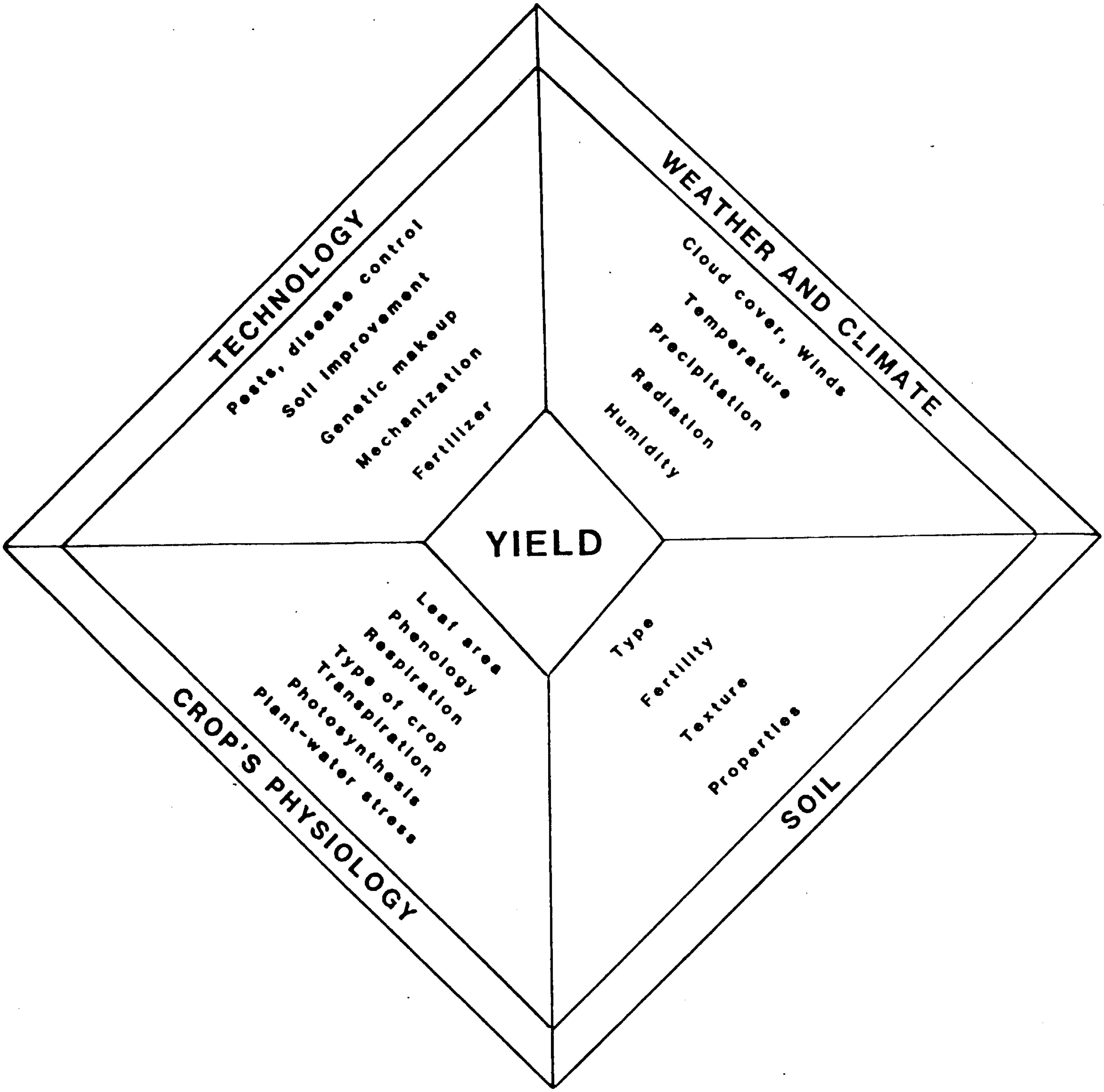

Figure 1. Factors affecting crop yields.

dependence contradicts the assumption of statistically independent elements which is made in regression analysis. Besides that, the statistical value of information contained in such a sample does not increase proportionally to the increase in the number of observations in the sample. Accordingly, the combined sample has sometimes only 30 percent of statistically significant additional information [30]. The estimated model coefficients based on this method will always have inflated values. Therefore, there is no efficient way of increasing the size of a sample for regression type of crop modeling.

The other approach lies in decreasing the number of variables. This may be accomplished in several ways. One of them is to neglect some of the factors, assuming that their impact on yield is constant over years. Such an assumption is normally applied to various soil factors. This assumption may be appropriate for a particular geographic region, especially if shifts in the type of soil utilized for the crop have been small. Even where some shifts over time have occurred, the variance of soil properties may be small when compared to the variance of weather conditions for a region. Another way is to express the quantitative impact of the technology-group and of some of the physiology-group factors in terms of trend over time. This approach is appropriate when our knowledge concerning the quantitative interpretation of technology-crop interaction is insufficient.

However, there are difficulties in the modeling of trend for the purposes of employing it in forecasting and the further interpretation of yield variation around the trend as a weather-dependent phenomenon. The first question which arises is how the trend should be computed. As is well known, the range of errors of the trend line, as defined by the least squares procedure, increases in the direction of the earliest and the latest years in the time-series. Accordingly, the magnitude of errors depends mostly on information for the earliest and latest years. In terms of predicting future yields, information in the latest years is very important. For example, if we assume that the yield series in Figure 2 is limited to 1945-1972 (Point A), then the trend in the latest part of the period would go up (trend line 2) since most of the years after 1965 had better than average cereal crop yields. If we assume that the yield series is limited to 1945-1965 (Point B), then the trend at the end of this period would go down (trend line 3). In each of these cases, the trends and fluctuations around the trends do not reflect the same response of yield to technology and weather changes as when the trend is calculated for the entire period between 1945 and 1980 (trend line 1). To overcome this problem a partially empirical approach is suggested later in this Section.

Another problem is related to the way the trend is represented in the model. A commonly used approach consists of approximating trend as an independent variable along with independent weather variables in the technology-weather-crop model [17, 33]. This approach has some shortcomings. Over the past 30 years, when technology factors greatly contributed to yield increases, the trend in a technology-weather-crop model has generally explained a much larger proportion of variation in yield (up to 90%) than has the weather [13, 17]. In such a case, the trend will dim, often significantly, the weather signal in yield variance, especially for those parts of the growing season when influence of weather on crops is not clearly defined. Thus, it is possible to lose useful weather information for yield prediction. Besides that, employing a weather variable for yield modeling without removing technological trend

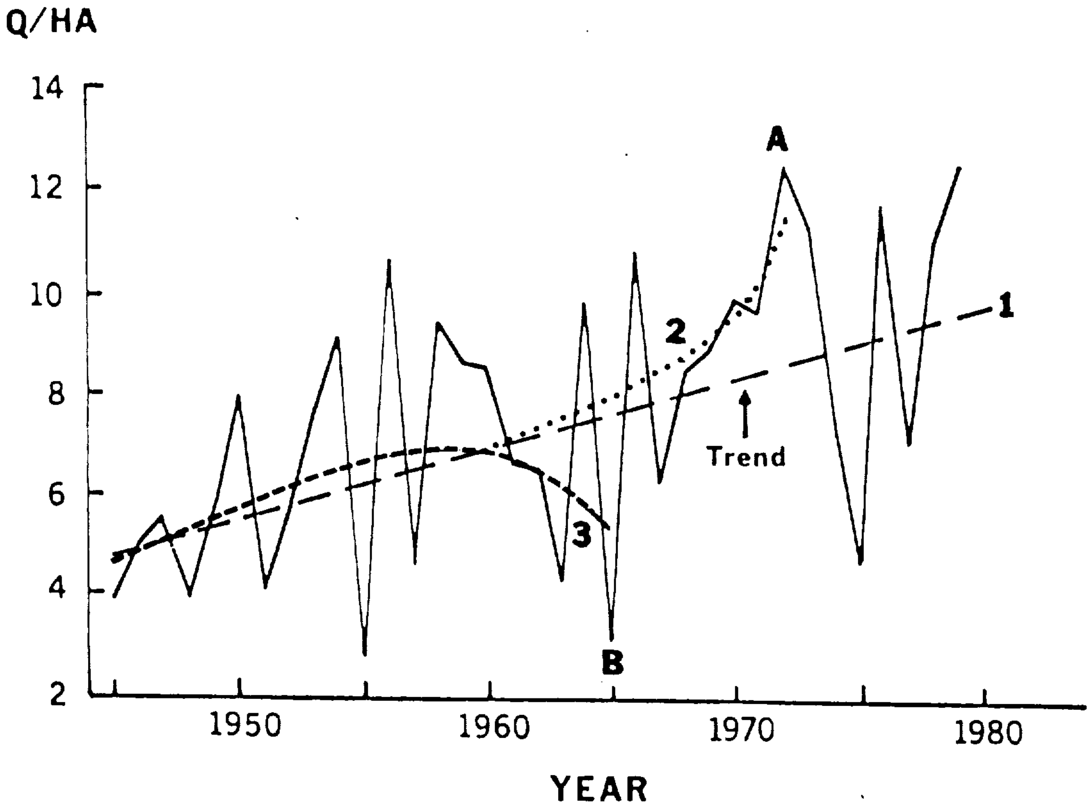

Figure 2. Cereal yield (quintals per hectare) in Kazakh economic region, USSR and trends 1, 2 and 3. (Points A and B define the end of periods for which trends 2 and 3 were estimated, respectively. Trend 1 was estimated from all of the data, 1945-1980).

from yield series involves some errors in estimation of yield-weather relationships. Accordingly, the separation of the trend from weather variables would be helpful in obtaining a better description of the weather-induced variation of yield in the analogue model. This separation would also allow us to apply various methods for reducing the number of weather variables in the model without losing some weather information useful for prediction. And, finally, this separation would be helpful in obtaining a more realistic representation of the weather-induced fluctuations of yield in the future. The concern is that absolute values of the weather-induced variation of yield have been gradually increasing as improvements in applied technology have taken place. Accordingly, a model based on historic data may underestimate the weather impact on yield late in the time-series and in the future. A better representation of the dependent variable could be in some relative form, although this approach also has some shortcomings. To overcome some of these drawbacks, an alternate technique will be presented later in this Section.

One way to reduce the number of variables in a weather-crop model consists of representing the weather and climatic factors through the two major parameters: precipitation and temperature. These variables reflect two major crop requirements: water and heat. However, adopting only the two major weather parameters as independent variables in a model does not fully achieve a considerable reduction in the number of variables in a model, because precipitation and temperature variables are usually created for numerous time periods (weeks, months, crop stages, etc.). Because of these problems, relatively long time-averaged weather variables are sometimes used to reduce the number of weather variables. Use of the simple arithmetic average of weather variables over time (spring, fall, pregrowing season) erroneously assumes that the different sub-periods within this particular time interval are equally important for the crop. Unfortunately, models developed on this assumption do not reflect the differing crop response to weather over time. Therefore, the weather in some important periods, which could provide useful signals for yield prediction, will be nullified or moderated by the weather in other periods. Figure 3 illustrates this idea of the non-uniform influence of the two major weather variables on cereal yields. This figure also shows that almost every month has some information which can be used as a signal for yield prediction. At the same time, as seen in this figure, these two weather factors are collinear. To reduce the number of weather variables in the model and to overcome collinearity, a special technique has been developed.

Improved Techniques for Weather-Crop Modeling

The equation for yield modeling used in this study can be written in the following general form:

$$Y = F\,[f_1(TR),\ f_2(W)] \qquad (1)$$

$$f_1(TR) = g_1(t) + e_{TR}$$

$$f_2(W) = g_2(P,T) + e_W$$

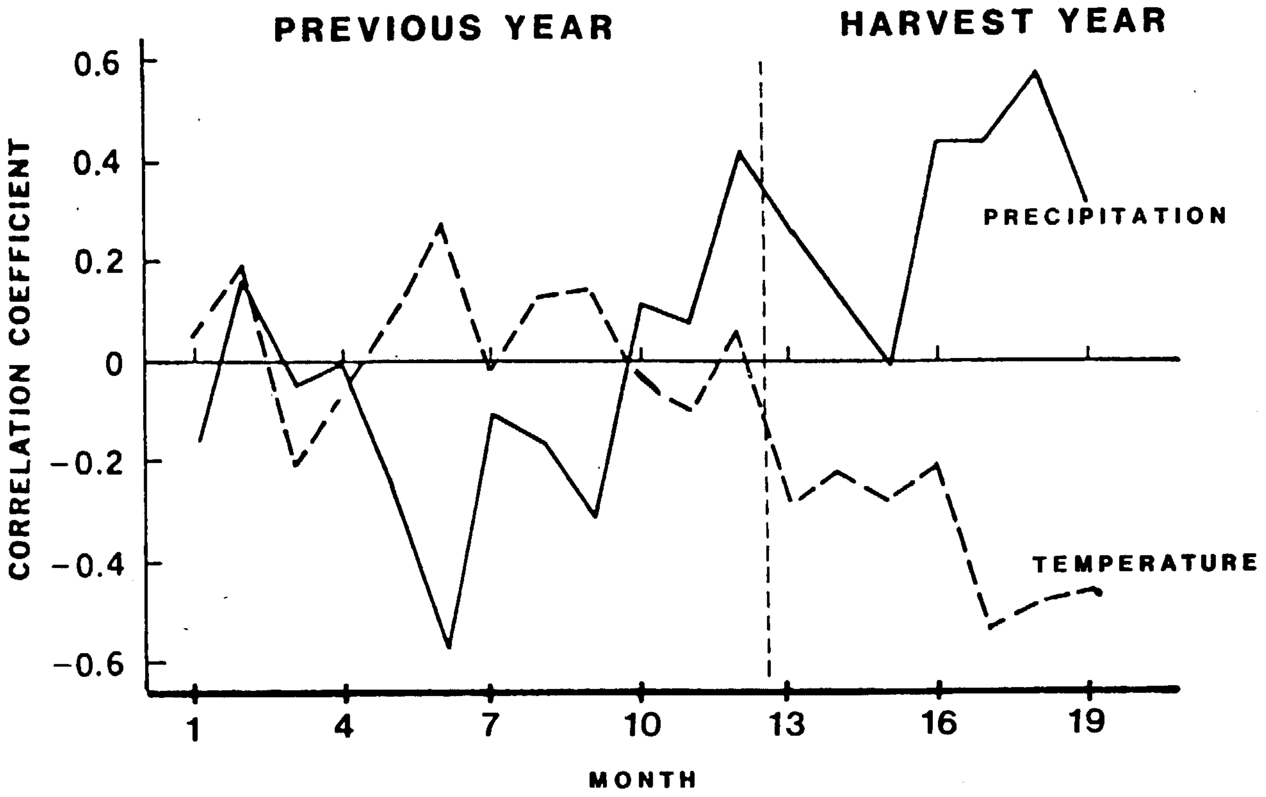

Figure 3. Correlation coefficients between end-of-season cereal yields and monthly values of precipitation and temperature in Kazakh economic region, USSR.

where Y is the reported yield, f_1 is a function of the yield's dependence on trend (TR) and f_2 is a function expressing the dependence of departures from the trend function (e_{TR}) on weather (W). Trend is usually approximated as a function of year number $g_1(t)$. Departures from trend are represented by functional combinations of precipitation and temperature weighted by the values of the physiological response of crop to weather fluctuations $g_2(P,T)$, and e_W are errors connected with estimating this function.

Equation 1 differs from the commonly used form of the relationship of yield to trend and weather which can be written as:

$$Y = a_o + a_1 TR + \sum_{i=2}^{K} a_i W_i + e \qquad (2)$$

where a_0, a_1, a_2, ..., a_k are regression coefficients and e is the model error.

The trend estimate in equation (1) was carried out with the application of least square regression techniques. For a better approximation of the weather induced fluctuation of yield in the last few years the following approach was applied. If the yields of the last several years in a yield series were approximately equally and uniformly distributed on both sides of the trend, then these years should be included in the data for trend specification. Otherwise these years should not be included in the calculation. In such a case the trend would be projected based on data prior to these years.

The trend can be approximated by either linear or nonlinear equations depending on the longevity of the yield series. For a yield series longer than 20-25 years, a second and sometimes a third degree of polynominal approximation might be considered the best from the standpoint of the long-range relationship among technology, climate and cropping power (achievable yield of crop) [19]. In our study, a second degree trend approximation was used.

The weather dependent function (f_2) in equation (1) can be calculated in two different ways: the dependent variable can express the departures from trend either as a difference ($Y - f_1(TR)$) or as a ratio ($Y/[f_1(TR)]$). The ratio is the preferred expression, since the absolute effect of growing conditions on yield increases as the trend level increases. This corresponds to the idea of increasing variability of yield over time with the improvement of technology [10], and this is very important for eliminating constant errors which can appear in the case of a long period of model application.

To study the influence of weather on yield, an analysis of the regression and correlation coefficients was performed. These coefficients were also used to weight weather variables for different months. For this purpose the following expressions were used:

$$W_{ij} = K_{ij} \Big/ \sum_{j=1}^{n} |K_{ij}| \qquad (3)$$

or

$$W'_{ij} = (-1)^{I_{ij}} K_{ij}^{2} \Big/ \sum_{j=1}^{n} K_{ij}^{2} \qquad (4)$$

where W_{ij} and W'_{ij} (which have the same sign as the corresponding K_{ij}) are the weights for weather variable i in month j, K_{ij} is a regression coefficient for variable i in month j or a correlation coefficient between departure of yield from trend and weather variable i in month j, n is the number of calendar months involved and I_{ij} is odd if $K_{ij} < 0$ and even otherwise.

Two weather variables, total precipitation and average temperature for each month averaged over the entire region, were used in this study. Months followed in the order from January (number one) through December (number twelve) of the year preceding the year of crop harvest and continued from January (number thirteen) through July (number nineteen) of the crop harvest year.

Weather information for the previous year was included since our research had shown that this information can serve as a predictor for the assessment of the crop productivity [11]. Expression (4) differs from (3) in that it assigns a relatively higher weight to weather variables of those months that have stronger influence on yield and a relatively lower weight to those months with a weaker influence.

Using these weights, index variables which are weighted average temperature or total precipitation over several months or over the entire nineteen-month period of aggregation were calculated using the following expression:

$$V'_{i(m_o-m_t)} = \Sigma_{j=m_o}^{m_t} W_{ij} V_{ij} \qquad (5)$$

where $V'_{i(m_o-m_t)}$ is the index variable for the particular weather variable i from the initial month ($j=m_o$) through terminal month ($j=m_t$), W_{ij} (or alternatively W'_{ij}) is as defined previously and V_{ij} is weather variable i for month j.

To compare methods for weighting weather variables over time, several types of weights were tested (Table 1). The first three types considered are the simple linear (Pearson) correlation coefficients, the partial (multiple) correlation coefficients and the standardized regression coefficients. These coefficients characterize some measure of the physiological response of yield to variation in weather. The fourth and fifth types of weights are time-dependent. The fourth type was designed to reflect a proportional increase in the value of weights with closeness to harvest. Thus, the index variable calculated based on these coefficients had higher input from the weather of the latest months and lower input from the weather of the earliest months. This is a somewhat logical way to combine weather variables, although it does not use any information about the real physiological value of the weather. The fifth type gives more weight to the earlier months rather than to the later ones. This type of weighting was used to examine whether it would lead to lower correlations between the obtained index variables and yield. The regression analyses obtained when these weighting alternatives were used, when the analysis was based on a simple average (equal weights) of weather variables and when the variables for each month are included in the equation separately rather than averaging the weather variables can be compared.

Table 1. Weights[1/] for weather variables in Kazakh economic region, USSR

Source of Weights	Version	Variable	Sign	Previous Year												Harvest Year						
				Month Number												Month Number						
				1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Pearson CC	1	P	+&-	-.04	.03	-.01	0	-.06	-.13	-.03	-.04	-.07	.02	.02	.09	.06	.03	0	.09	.09	.12	.06
		T	+&-	.01	.05	-.06	-.02	.03	.07	-.01	.03	.04	-.01	-.03	.02	-.07	-.03	-.07	-.06	-.14	-.13	-.12
	2	P	-					.18	.40	.08	.12	.22										
			+										.04	.03	.15	.10	.06	0	.16	.14	.21	.11
		T	+		.23			.11	.33		.15	.18										
			-													.12	.04	.12	.09	.23	.21	.19
Partial CC	1	P	+&-	.03	.10	-.07	-.03	.05	-.08	-.01	-.08	-.02	0	.05	.02	-.02	-.06	-.03	.01	.10	.15	.08
		T	+&-	.10	.11	-.01	.05	.01	.06	-.07	.04	.03	-.06	-.04	-.04	-.10	-.09	.03	.05	.04	-.06	.01
	2	P	+	.16	.58			.26														
			-			.22	.11		.28	.04	.27	.08										
			+																.03	.29	.45	.23
		T	+	.30	.34		.14	.04	.17													
			-							.16			.12	.09	.08	.22	.20				.13	
			+															.25	.42	.33		
Standardized RC	1	P	+&-	.04	0	-.09	.08	-.02	-.12	-.04	.01	-.02	.05	-.03	.03	-.05	-.04	.05	.08	.12	.11	.02
		T	+&-	0	.09	-.10	-.04	.08	-.01	.03	.06	.01	-.07	.07	.09	-.03	-.01	-.04	.03	0	-.12	-.15
	2	P	-					.14	.65	.21												
			+															.14	.22	.30	.27	.06
		T	+		.28			.25						.20	.27							
			-			.45	.21						.33									
			-																		.45	.55
Time Dependent C Increased	1	P&T	+	0	.01	.02	.02	.03	.03	.04	.04	.05	.05	.06	.06	.07	.07	.08	.08	.09	.09	.10
	2	P&T	+													.04	.07	.11	.14	.18	.21	.25
Decreased	1	P&T	+	.10	.09	.09	.08	.08	.07	.07	.06	.06	.05	.05	.04	.04	.03	.03	.02	.02	.01	0
	2	P&T	+													.25	.21	.18	.14	.11	.07	.04

CC is correlation coefficient
RC is regression coefficient
C is coefficient
P is precipitation
T is temperature
1/ The weights are determined by equation (3).

For each type of weighting alternative two versions of index variables were calculated and tested. For the first version, index variables were obtained by averaging monthly weather parameters over the entire nineteen-month period. The sign of the weight was taken into consideration. For the second version, several sub-periods were defined which separated periods with weights of opposite signs. In this case the index variables for each sub-period were assigned a + sign and each was entered into the regression analysis as a separate variable.

Finally, the models based on index variables were constructed. Based on the R^2 and the mean square error (MSE) these models were compared one to each other and also to models with simple average (equal weights) of weather variables or to models with the variables included for each month separately. Table 2 presents the values of the MSE and R^2 for the tested models with weighted index variables over the entire 19-month period (Version 1), and over some sub-periods (Version 2), using cereal yield and weather data for Kazakh economic region. Models based on the weighted weather index variables (precipitation-P, temperature-T) show the largest values for R^2 and the smaller values for MSE. The best model in terms of smallest MSE was based on index variables weighted by standardized regression coefficients. The quadratic form of this model had a slightly larger MSE (156.0 vs 152.2). The second best model was based on index variables weighted by Pearson correlation coefficients. MSEs for the linear and quadratic forms of this model were 253.3 and 248.3, respectively.

The model based on the simple average of the (P,T) variables over the entire 19-month period had a very low R^2 (.102) and large MSE (895.9) for the linear form. These results were expected as the influence of these variables on yield over time is not uniform both in value and sign (Figure 3). Thus, averaging of these variables over time eliminates most of the signals useful for prediction.

The R^2 increases sixfold (.633) and MSE decreases by one-half (453.5) when the whole period is broken into four sub-periods in accordance with the sign of influence of weather on crop yield. Although this approach is physiologically grounded, the estimates for this model (R^2 and MSE) are worse than for a model with variables weighted by means of either the Pearson correlation or standardized regression coefficients.

When weights increase as the harvest time is approaching, the R^2 reaches 0.382, MSE=616.5. Comparing this case with the model when the weights decrease shows a very great difference. The R^2 value for the last model is very low, 0.031, MSE=966.8. This is expected since this model does not reflect the response of yield to weather variation at all. But when weather of the less important part of the nineteen-month period was eliminated from the index variables, the performance of the model improved. The R^2 for the model, based on index variables of the last ten months sub-period increased to 0.545, MSE=454.1. It is necessary to notice that the R^2 for the model where the weather parameters are considered to be independent was slightly higher than the other values of R^2, 0.880. Even so, the slightly higher R^2 cannot be considered as significant since this model had 26 variables against two for other models. Also, the MSE for this model was over twice as large as for models with the first three types of weather variable weighting.

Table 2. R-square and mean square error (MSE)[1/] for tested models (cereals in Kazakh economic region, USSR)

Source of coefficients for weighting	Variables	Linear Regression			Quadratic Regression		
		Number of variables	R^2	MSE	Number of variables	R^2	MSE
		Version 1					
Pearson CC	Sum of P_{1-19}, T_{1-19}	2	0.746	253.3	5	0.775	248.3
Partial CC	" " "	2	0.538	461.3	5	0.570	475.0
Standardized RC	" " "	2	0.848	152.2	5	0.859	156.0
Time Dependent C							
Increased	" " "	2	0.382	616.5	5	0.414	647.2
Decreased	" " "	2	0.031	966.8			
Simple average	" " "	2	0.102	895.9	5	0.115	978.0
Without averaging	P7-P19, T7-T19	26	0.880	531.9			
		Version 2					
Pearson CC	P_{5-9}, P_{11-19}, T_{5-9}, T_{11-19}	4	0.728	290.5	8	0.768	286.5
Partial CC	$P_{1,2,5}$, $P_{3,4,6-9}$, P_{16-19}, $T_{1,2,4-6}$, $T_{7,10-14}$, T_{15-17}	6	0.569	494.3	12	0.642	526.7
Standardized RC	P_{5-7}, P_{15-19}, $T_{2,5,11,12}$, T_{18-19}	4	0.726	292.8	8	0.746	314.9
Time dependent C							
Increased	P_{13-19}, T_{13-19}	2	0.545	454.1			
Decreased	" " "	2	0.436	562.9			
Simple average	P_{5-8}, P_{9-11}, P_{12-14}, P_{15-19}, T_{5-8}, T_{9-11}, T_{12-14}, T_{15-19}	8	0.633	453.5			
Without averaging	P13-19, T13-19	14	0.796	331.4			

[1/] MSE is expressed as the square of the percentage deviation from trend.

Models based on index variables averaged over only the most important sub-periods show slightly lower R^2 and higher MSE than do models based on index variables for the entire nineteen month period. However, these differences are not statistically significant. None of the models show substantially larger R^2 when quadratic terms are included in the equations, while the MSE may actually increase slightly. Calculating the weights according to expression (4) sometimes results in an increase in the values of R^2 for the model, but this increase is not statistically significant.

The proposed technique of aggregating weather variables on the whole showed good results in significantly reducing the number of variables in the models without losing important information required for better prediction. The better indices for weighting weather variables were based on the standardized regression coefficients followed by the Pearson correlation coefficients. Models based on the entire nineteen-month period show slightly better performance than models based on sub-periods.

Another method of decreasing the number of variables in a yield model is based on computation of agrometeorological indices or indices of the soil moisture budget. Results of yield modeling based on some of these indices were compared with the results obtained with the application of the proposed approach in this study. Models developed based on originally measured precipitation and temperature have better statistical assessments (R^2 and MSE) than those which are based on calculated indices. More detailed considerations concerning the comparison of these two types of models are given in Appendix III.

Historical Data

This study was conducted for five major barley producing regions of the USSR. They included the following economic regions: South-West (1), Belorussia (2), Volgo-Vyatka (3), North Caucasus (4) and Kazakh (5). These regions are shown in Figure 4. Barley normally occupies an area of from 1.2 million hectares in Volgo-Vyatka to 5.6 million hectares in Kazakh. The development of models for areas smaller than an economic region would involve difficulties connected with the lack of crop and weather data. Barley crop statistics on production, area and yield are available by economic regions for the period of 1945-1978. This information is usually collected, stored and distributed by the Central Statistical Administration of the USSR. Barley production data, like those for any other crop, are actually measured after harvesting in "bunker" weight, i.e., weight with some water and foreign matter content in excess of the required standard. This non-grain weight normally amounts to no more than 5-10 percent of the actual production.

Weather information for computing regional total monthly precipitation and mean monthly temperature is available from meteorological publications of the Hydrometeorological Service of the USSR for the period of 1945-1978. The regional precipitation and temperature were obtained by averaging of weather elements taken from several weather stations in each of the regions. Stations were selected on the basis of the longevity and high quality of their records and representation of local weather in different parts of regions. The number of stations selected depended on the size of the region; 5 were selected in South-West, Belorussia and Volgo-Vyatka economic regions, 6 in North Caucasus, and 12 in Kazakh economic region (Fig. 4, Table 24).

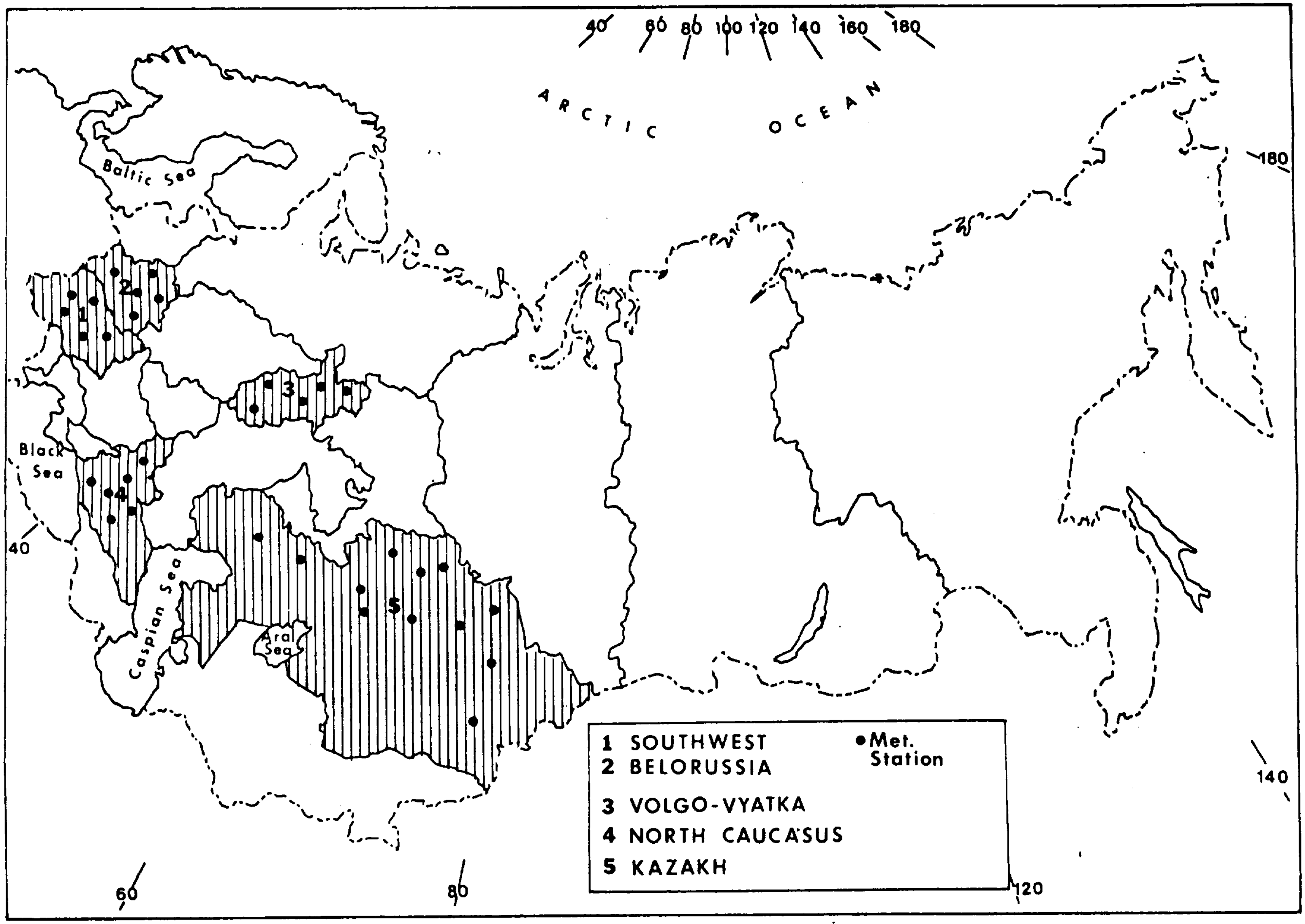

Figure 4. Economic regions studied and location of selected meteorological stations, USSR.

PECULIARITIES OF THE USSR'S NATURAL RESOURCES IN CONNECTION WITH BARLEY REQUIREMENTS

Planted barley area is second only to spring wheat in the USSR and accounts for 26 percent of the total area used to produce grain. Since barley yields are commonly higher than those of spring wheat, barley accounts for 28 percent of the total production of all grain. This is a larger share of the total than either winter or spring wheat.

Distribution of Barley

Barley is very important to Soviet agriculture. It is suited to a wide range of environmental conditions and is particularly tolerant of the different climatic and soil zones of the USSR. Figure 5 shows the proportion of grain land planted to barley by the USSR economic regions. The three most important regions (Kazakh, Ukraine, and Volga) each produce more than 10 percent of the total barley production for the USSR. Five other economic regions--Belorussia, Central, Central-Chernozem, North Caucasus, and West Siberia--each produce from 5 to 10 percent of the total USSR barley production.

In many areas barley is used extensively as a replacement for winter wheat damaged during the winter and early spring. Figure 6 shows changes in barley and winter wheat harvested area in economic regions selected for barley-yield modeling. The five recent years (1974-1978) were selected for analysis as there was no trend (or the trend was insignificant) in the area planted to barley during this period. As seen in Figure 6, changes in barley and winter wheat harvested area occur in opposite directions in regions with a very high occurrence of winterkill caused by frost (South-West and North Caucasus), or caused by alternate freezing and thawing during the winter (Volgo-Vyatka). In these regions, the entire area of damaged winter wheat is usually replanted with barley. There is no such tendency in Kazakh, as barley is not used as a substitute for winter wheat, which is grown in a different part of this region. The recent trend towards a substantially smaller area planted to winter wheat and a larger area planted to barley in Belorussia does not result from the same phenomenon of using barley as a substitute for damaged winter wheat.

Spring barley is not only very suitable for the environment, but it is also higher yielding than spring wheat and oats (Figure 7). That is why it is a very valuable crop and is used in the USSR as a substitute for damaged winter wheat.

The policy toward growing barley has changed several times over the past 35 years in the USSR. The changes were connected mainly with the political and economic situation in the country and also with the potential natural resources. This changing policy towards barley is reflected clearly in the areas of barley harvested in the selected economic regions from 1945 through 1978 (Figure 8). In the 1950's the area of barley dropped substantially, from 1/3 to 3/4 as compared with the late 1940's, in Belorussia, Volgo-Vyatka and North Caucasus economic regions; in the South-West region, barley area dropped by 15 percent. This decrease was related to the expansion of winter wheat plantings which resulted from the introduction of new very productive

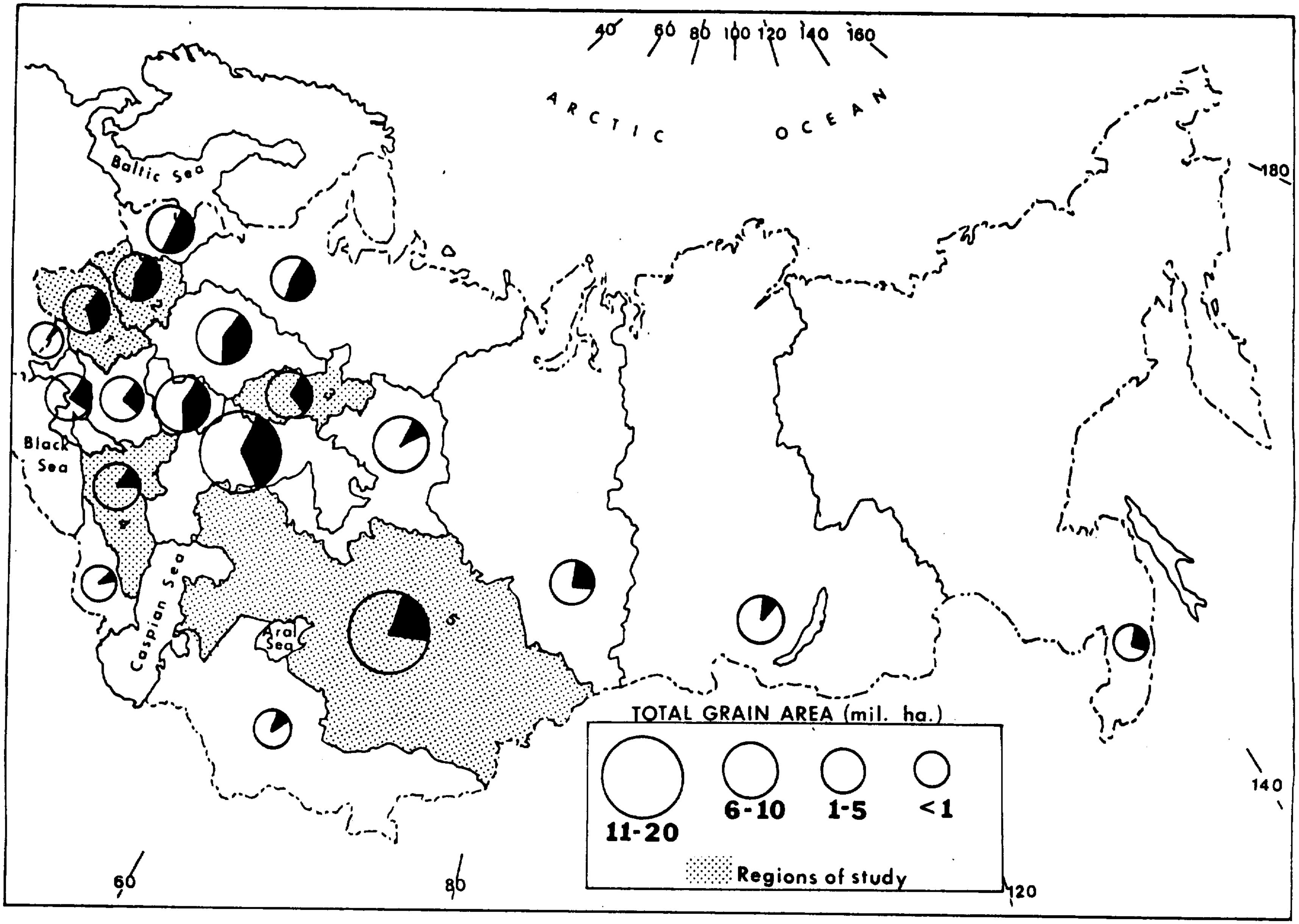

Figure 5. Proportion of grain land planted to barley, by economic regions, USSR, 1978.

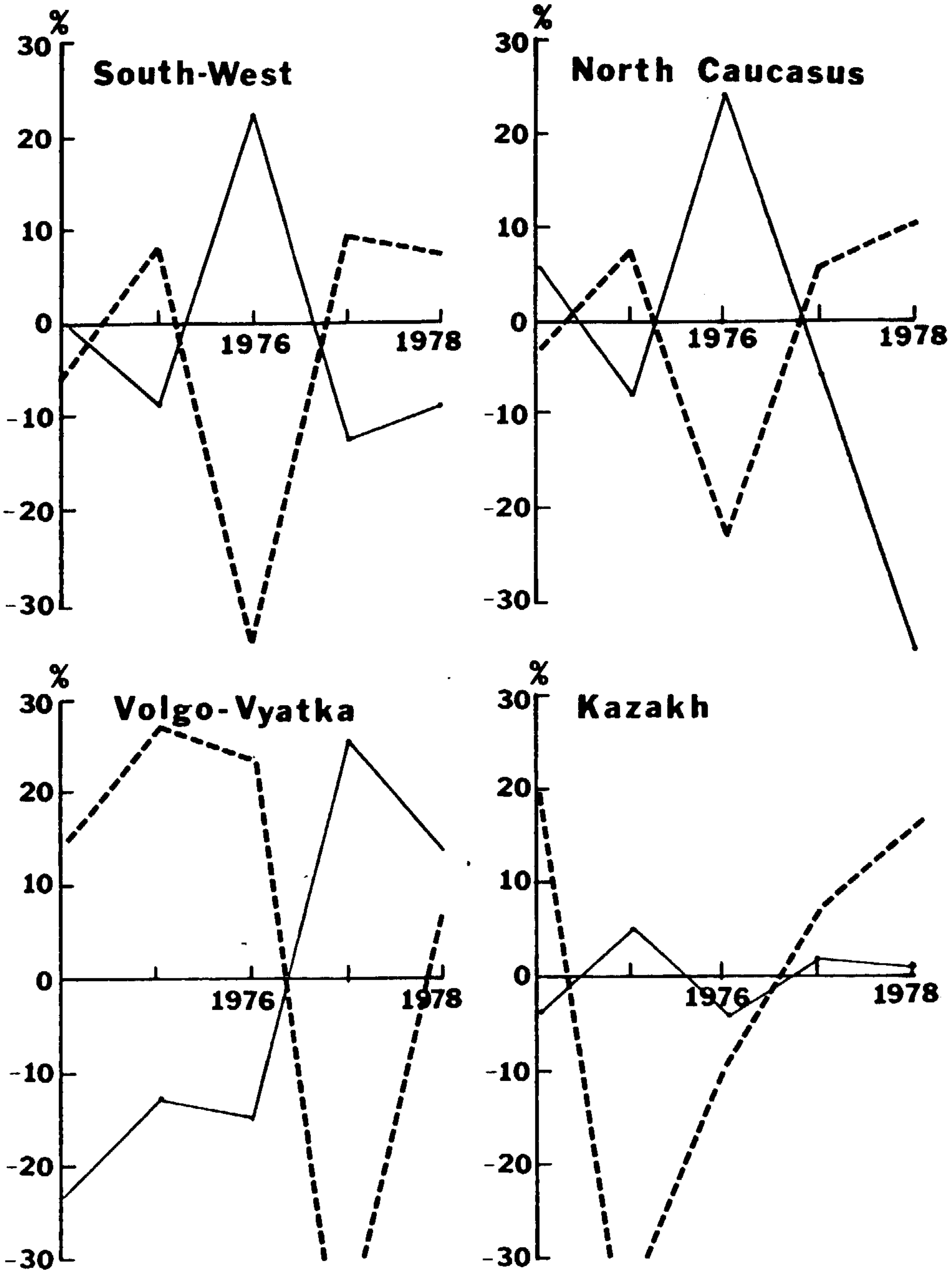

Figure 6. Changes in harvested area of spring barley (solid line) and winter wheat (dashed line) shown in percentages from average area of each crop during 1974-1978 in selected economic regions, USSR.

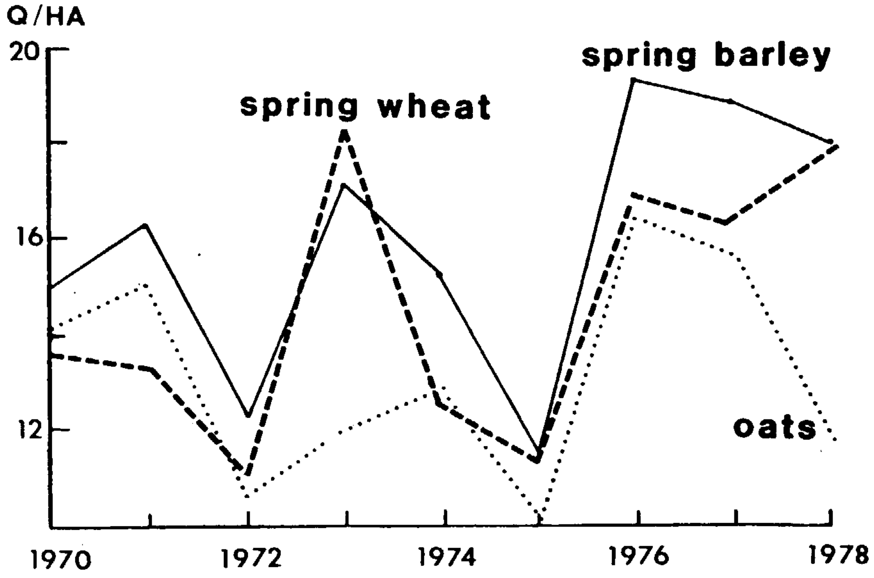

Figure 7. Yields of spring barley, spring wheat, and oats in Volgo-Vyatka economic region, USSR in 1970-1978.

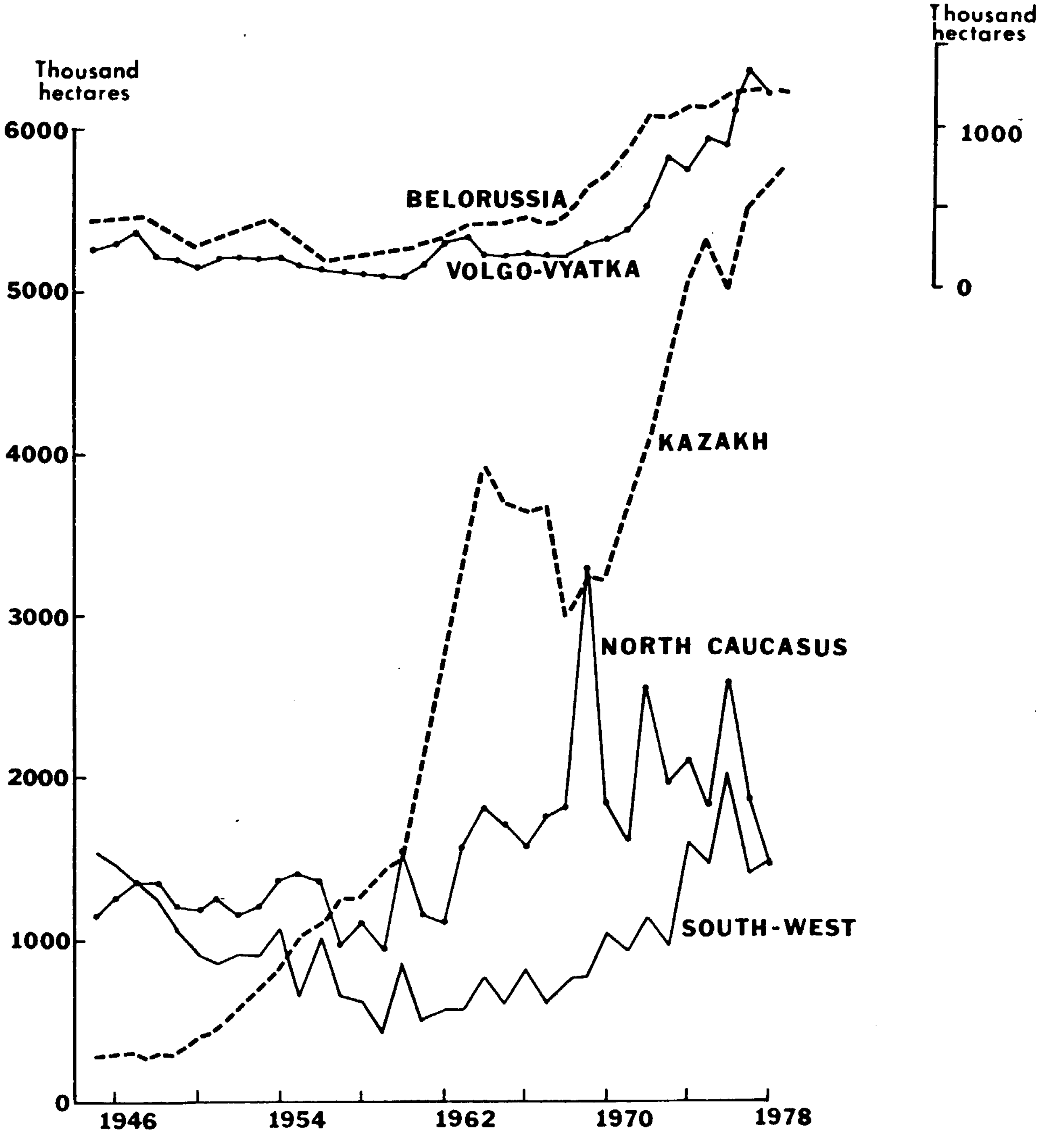

Figure 8. Harvested area of spring barley in selected economic regions, USSR in 1945-1978.

varieties of winter wheat at that time and reflected the USSR policy of rapidly increasing grain production. This decrease in the area of barley was also connected with the enormous expansion of plantings of corn, a crop which later proved to be impractical for most of the newly planted areas. A dramatic increase in barley area occurred in the 1970's, particularly in the northern half of the USSR. As seen in Figure 8, the Volgo-Vyatka and Belorussia barley areas expanded from 100-300 thousand hectares in the late 1950's to 1000-1200 thousand hectares by the late 1970's. In Central economic region (not included in this study) the area of barley increased almost 27-fold over the same 20 year period. In Kazakh economic region, the area of harvested barley increased rapidly over the whole period after World War II. This was because in the climatic conditions of Kazakh, barley and spring wheat were the most suitable crops. The increase in the barley area of Kazakh in the late 1950's and early 1960's can be attributed to the well known period of virgin lands plowing. During this period a lot of new land, including land not very suitable for agriculture, was brought into production. The very sharp increase in the planted barley area in Kazakh in the 1970's was connected mainly with the development of waste land, and also with the replacement of oats.

We should notice that the rate of increase in the harvested area of barley declined substantially in the late 1970's in Belorussia and North Caucasus compared to the earliest years. This suggests that the amount of land suitable for barley in these regions has been fully utilized. Of course, there are some waste lands left in these regions, but their utilization would require large investments to raise the fertility of soil and to bring them into efficient operation [15].

The varieties of barley cultivated in the USSR are very different and their improvement and distribution are usually adjusted to the climate, soil and economics of different zones of the USSR. Some varieties with high potential yield, such as Odesskii 36, Donestskii 4, Krasnodarskii 35, are distributed in the Ukraine, North Caucasus and some other regions of the southern USSR. In the northern part of the country widely distributed varieties are Moscowskii 121 and Kazanskii 64. The recent effort of scientific institutes to improve barley varieties in this zone has launched such new varieties as Luch, Sever 1, Viking, Agat, Dar, Vyatich and others which have high potential yields, capability for early ripening and other useful properties [16].

Principal Natural Resources Required by Barley

Barley normally matures earlier than other spring small grain crops. The length of the growing period of barley depends on the variety. But it is possible to single out some combination of weather patterns which regulates the growth and development of any variety of barley.

Although barley seeds start to germinate at about 1-3 degrees C, the optimal temperature for germination is 15-20 degrees C. Under optimal temperature conditions and with a good water supply, the period from sowing to emergence is five to seven days. From emergence to tillering usually takes another 10 to 15 days. This also depends on the water supply and temperatures. Barley is most sensitive to temperature and, especially, soil moisture in the period from shooting to shortly after heading. However, compared to other small grain crops, barley has the lowest transpiration ratio (297 to 468 as compared with 359 to 554 for spring wheat) [18]. High temperatures during flowering and pollination have a very adverse effect on barley yields. However, temperatures as high as 48 degrees C and above during the ripening of the barley injure it less than is the case for wheat and oats. The water requirement of barley in the period of ripening is substantially reduced.

Thus, from the standpoint of water and heat consumption, barley is considered to be more tolerant of weather extremes than wheat and oats. But, even considering barley's advantages, recent data show that in the climatic conditions of the Soviet Union, total barley production might vary as much as 30 million tons due to variations in weather. In the economic regions of the USSR chosen for modeling, the variation of barley output during two recent consecutive years, 1975 and 1976, ranged from 138 to 189 percent in the dry area and from 50 to 65 percent in the wet area.

Agroclimatic and Soil Conditions

Agroclimatic resources of economic regions of the USSR differ greatly. But in general, it is possible to identify several types of conditions for most of the regions.

The most peculiar feature of an agricultural area in the southern part of the USSR, including North Caucasus and Kazakh, is the combination of fertile chernozem and chestnut soils, an abundance of sunshine and heat, and a deficit of water, especially during the main part of the growing season. Water deficits, accompanied by high temperature and sometimes by dry winds, are the factors primarily responsible for limited productivity of barley as well as other crops. Figure 9 shows the balance of water and degree days, the two most important climatic factors which control the productivity of barley [2]. In North Caucasus (4) and Kazakh (5), due to the great amount of heat, potential evaporation exceeds precipitation by 200-400 mm per year. Droughts occur in this region 50 percent and in some places 70 percent of the years. Dryness of the USSR territory increases in the direction from north to south and southeast. This is because the climate acquires more continental features to the south and southeast. During the summer dry tropical air from Middle Asia and the Middle East usually comes to this territory, sometimes bringing droughts and desiccating winds [18]. Therefore, the amount of precipitation declines quite rapidly in the southeast. Typically when the South-West economic region (1) has 500-600 mm (some places 700 mm) precipitation per year, precipitation in the eastern part of North Caucasus and western part of North Kazakh regions falls to about one-half of these figures. Dry periods in Kazakh and North Caucasus can occur from May through September and continue from 10 days to 3 months. In some places droughts occur every 2-4 years. Sometimes they cover very large areas (as in 1946, 1963 and 1975). Unfortunately, it is impossible to rely on winter precipitation as a source of soil moisture in this area. There is very little precipitation during the

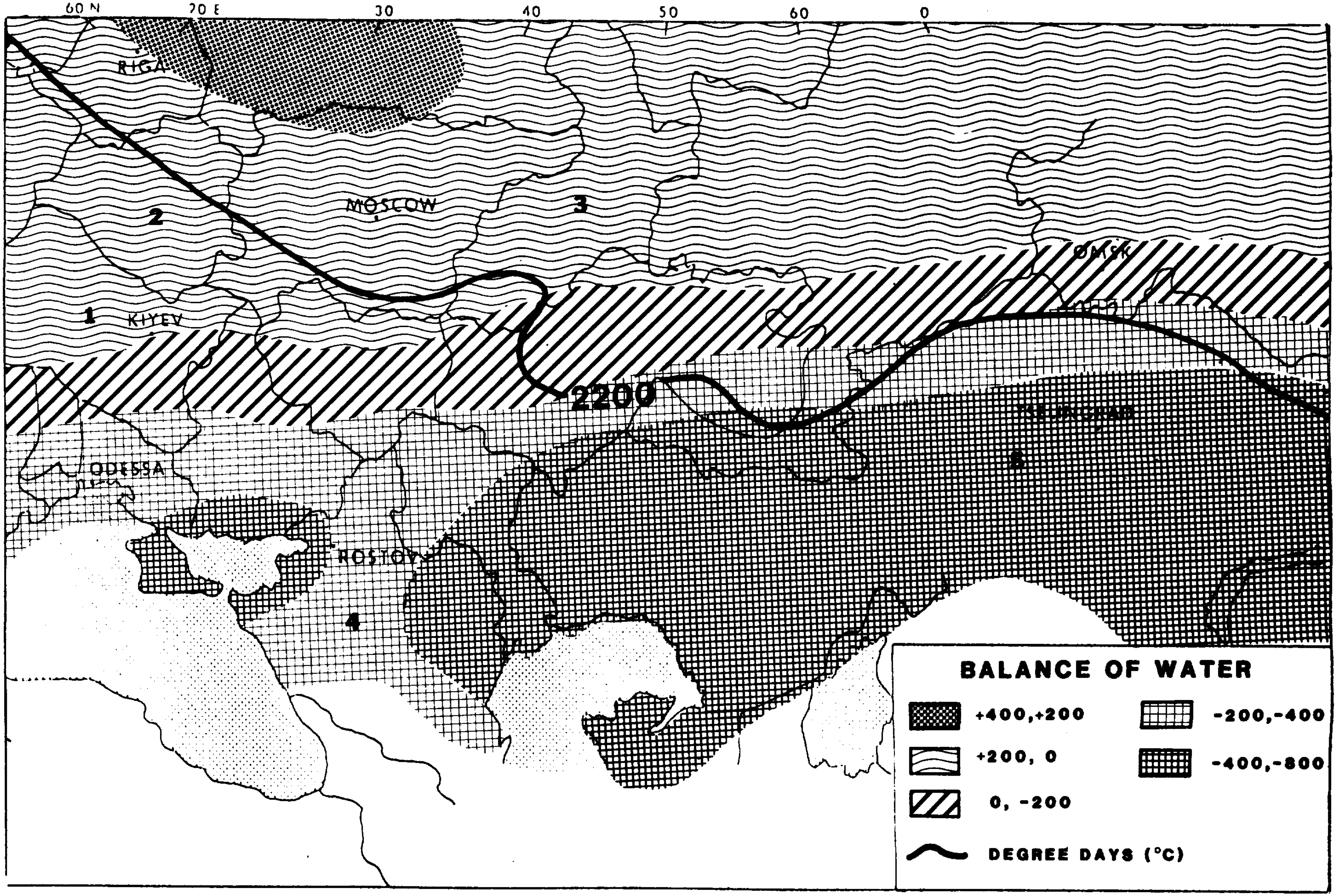

Figure 9. Agroclimatic resources expressed as a balance of water (difference between precipitation [mm] and potential evapotranspiration [mm]) and degree days in the USSR [2].

winter and as a result, the snow cover is not deep (Figure 10). In some years when rainfall in the second half of summer starts earlier than normal, this area yields a bumper barley crop (as in 1976 and 1978).

Soil moisture normally meets crop requirements in most of the northern area where the main factor limiting crop production is heat. The total amount of heat during the growing period is limited, especially in the areas north of 55 degrees latitude. One of the important climatic peculiarities of this territory is the existence of cold polar outbreaks. In summer they bring late or early frosts which can damage barley crops or delay their development and growth. But sometimes when the polar frontal zone in the troposphere is displaced far to the north, it causes the formation of specific weather patterns and finally the distribution of an anticyclone in the southern part of the USSR. Anticyclones block the maritime Atlantic westerlies and also polar outbreaks. Then, the weather becomes warm and dry. In such conditions the occurrence of droughts is likely. In this usually wet area, these conditions occur on the average once or twice in ten years. In the last ten years the occurrence of droughts here was more frequent (1972, 1975, and 1981).

Some particular soil and agroclimatic characteristics of the economic regions studied are summarized below and presented in Table 3 and Figures 11 and 12.

South-West, North Caucasus and Kazakh economic regions are in the Chernozem Zone of the USSR. Location of this Zone approximately corresponds to the area with inadequate or unstable water balance (negative water balance in Fig. 9). This Zone is known as an area with predominant varieties of chernozem soil in the USSR. Chestnut and grey forest soils are also present in some areas. Belorussia and Volgo-Vyatka economic regions are in the Non-Chernozem Zone. Location of this Zone approximately corresponds to the wet area (positive water balance) shown in Fig. 9. Typical soils of this Zone are various types of podzolic and swampy soils. Considerable swampiness of the territory and waterlogging of soils are the typical features of most of the land resources of these regions. In the southern part of these two regions there are also chernozem and grey forest soils.

South-West Economic Region (1)

Average annual precipitation is 600 mm. In western and northern parts of the region this amount reaches 700 mm; in the southeastern part it drops to 500 mm. Maximum precipitation falls in the summer period (70-85 mm per month). Dry periods rarely occur in the southeastern part. Degree days (above 10 degrees C daily mean air temperature) accumulated over the warm period range from a low of 1600 in the southwest to 3000 in the southeast part of the territory. Summer temperatures can reach 19 degrees C and winter temperatures drop to -6 degrees C. In some years temperatures can decrease to -26 degrees C. Snow cover lasts around 3 months (in the southeastern part around two months).

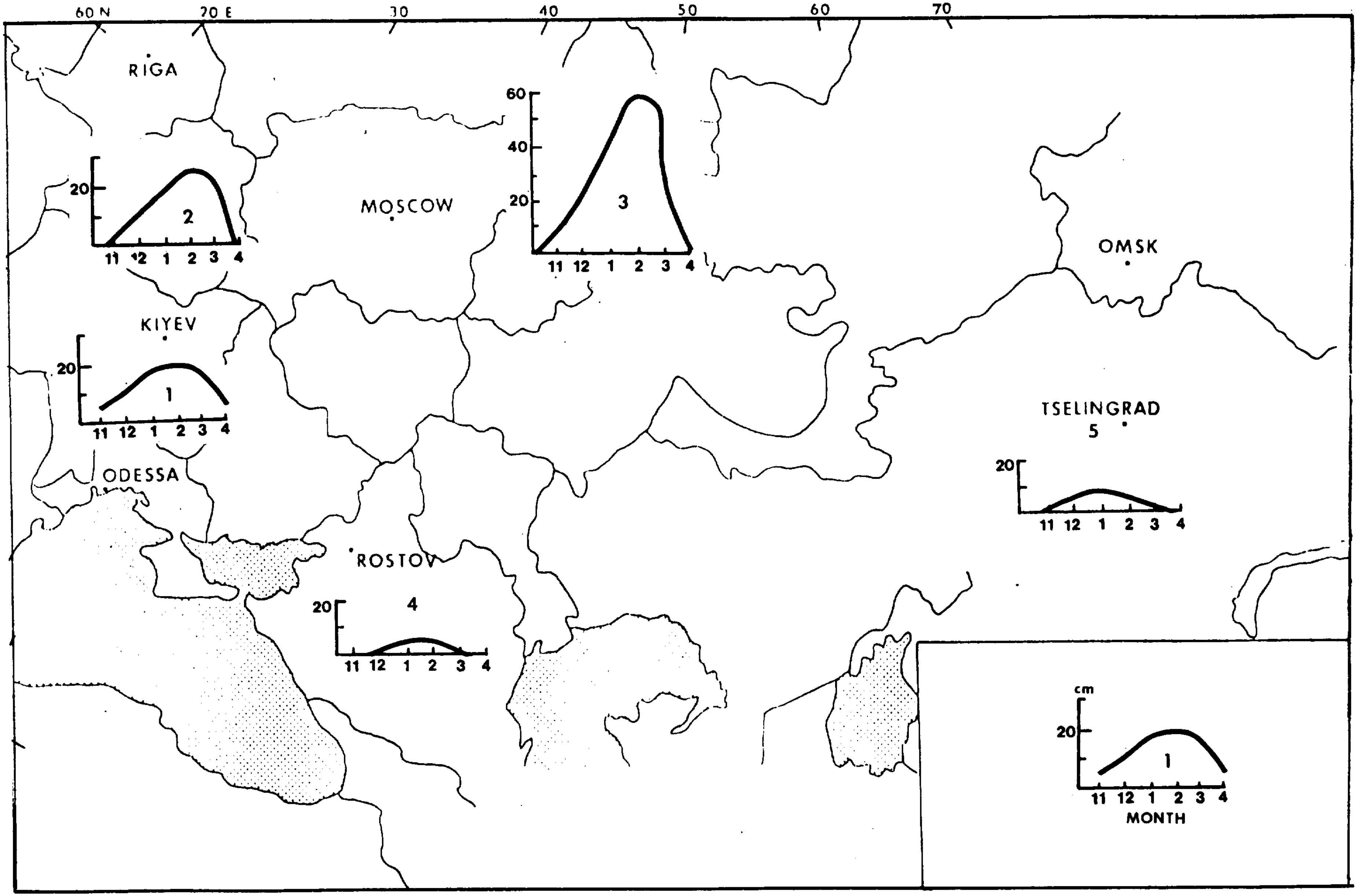

Figure 10. Average distribution of snow cover, USSR (meteorological stations).

Table 3. Agroclimatic conditions of selected economic regions in the USSR [2]

Economic Region	Solar Radiation in June (kal/cm^2)	Spring daylight duration (hours)	Yearly Precipitation (mm)	Dry period and its length (days)	Degree days for period with daily t > 10° C	Average Temperature of the Warmest Month (°C)	Duration of Period with t > 0° C (days)	Average Temperature of the Coldest Month (°C)	Average Minimal Temperature (°C)	Duration of Snow Cover (days)
South West	13-15	14-15	500-700	rare	1600-3000	17-19	150-165	-4 to -6	-24 to -26	80-100
Belorussia	13-15	14-16	500-700	rare	1900-2500	16-18	110-150	-5 to -10	-23 to -28	80-120
Volgo-Viatka	13-15	16-18	500-650	rare	1500-2800	16-18	102-125	-10 to -14	-32 to -40	150-180
North Caucasus	16-18	14-15	400-600	May-Sept. 50-100	2000-3500	23-24	165-180	-4 to -8	-20 to -26	< 80
Kazakh	16-19	14-16	250-400	Apr.-Sept. 50-180	2000-2600	19-24	110-135	-14 to -18	-30 to -40	120-150

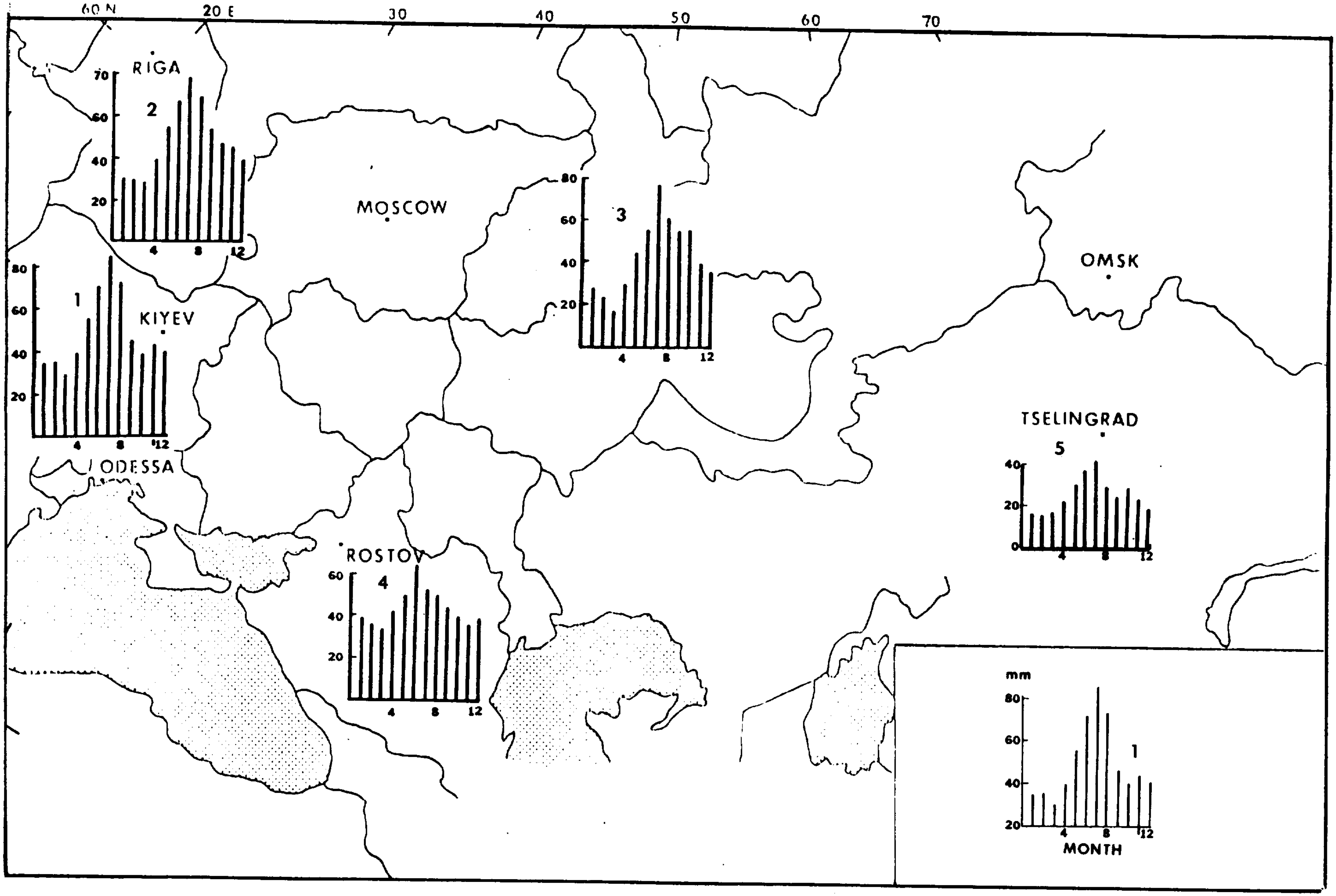

Figure 11. Distribution of average regional precipitation during the year, USSR.

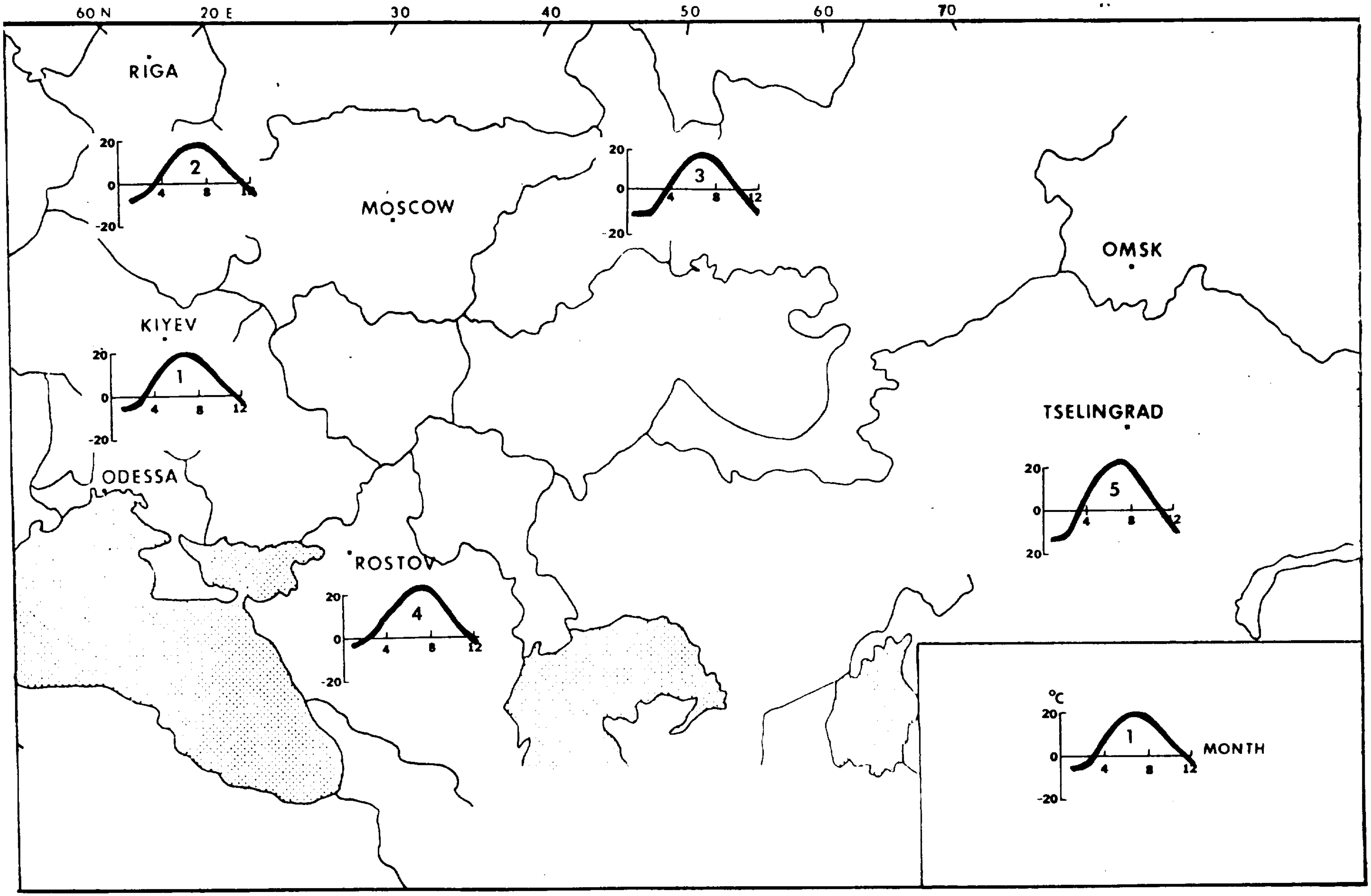

Figure 12. Distribution of average regional temperature during the year, USSR.

Belorussia Economic Region (2)

Total annual precipitation is about 600 mm, but in the southern part it drops to 500 mm. The summer period receives the largest portion of annual precipitation (65-75 mm per month). A substantial amount of precipitation also occurs in the fall. Winter and early spring receive the smallest portion of annual precipitation. Cumulative degree days (1900-2500) can provide good development for different varieties of barley. The warmest months, June, July and August, have average temperatures of 16-18 degrees C. The duration of the period with temperatures above 0 degrees C is limited in the northern part of the territory (110 days). The longevity of snow cover is 3 to 4 months, except in the southern part which receives less snow.

Volgo-Vyatka Economic Region (3)

Total annual precipitation is slightly less (550 mm) than for Belorussia, particularly during the first eight months. However, monthly rainfall during the summer period is still substantial (55-75 mm). In the fall the amount of precipitation decreases more slowly than in other regions. The northern part of this region has as few as 1500 degree days. In the southern part the degree days increase to 2300. The period with positive (above freezing) temperatures in this region is the shortest of the five regions studied (102-125 days). Winter conditions in this region are very severe. Temperatures in the coldest month can fall as low as -40 degrees C. Snow usually covers fields for 5 to 6 months of the year.

North Caucasus Economic Region (4)

Total precipitation for this region averages 500 mm per year. This amount is fairly small compared with large heat resources (2800 to 3600 degree days), and the resulting high rate of evaporation. The summer period of maximum rainfall is clearly defined, but the transition in the precipitation rate from spring to summer and from summer to fall is not as sharp as in the higher moisture regions. The warmest month is July with average temperatures of 23-24 degrees C. June and August are slightly cooler. Dry periods can occur any time during the May to September growing season and may continue for two or three months. The period with above freezing temperatures is fairly long (5-6 months). Winter in this region is generally moderate, but the temperature can drop as low as -26 degrees C. The duration of snow cover is less than 3 months.

Kazakh Economic Region (5)

Of the five regions, Kazakh has the lowest amount of precipitation (around 300 mm per year) in the area where barley is grown. Forty percent of this amount falls during July and August. But this amount is inadequate to the crop's requirements because of the existing heat resources (2000-2600 degree days). Accordingly, the potential evaporation substantially exceeds the amount of precipitation. Droughts can occur in the period of April through September and may continue for up to six months. The hottest month is July (19 to 24 degrees C), the coldest is January (-14 to -18 degrees C). Snow can cover fields for 3 to 5 months, but its average depth does not exceed 10 to 12 cm, and it has usually melted by the middle of April, long before the planting of barley starts.

In general it is possible to divide these regions into two groups. South-West Belorussia and Volga-Vyatka have a good water supply, a moderate or deficient heat supply, and droughts are rare. North Caucasus and Kazakh have plenty of sunshine and heat but very limited water resources. However, agroclimatic conditions vary greatly within these two groups, which makes the modeling of barley yield a very specific task for each region.

DEVELOPMENT OF BARLEY YIELD MODELS

Dynamics in the historic barley yield series were analyzed by examining two components of yield variability: long term and short term. The long term component reflects technological progress in growing crops; the short term component reflects annual changes in weather [12, 19]. Technological progress primarily has much more important impact on long term yield levels rather than on short term yield changes. Natural conditions can change substantially from year to year and thus have a very important impact on short term yield shifts.

Technology-Related Changes in Yield

The technological component of yield change was examined for two types of yield variability, temporal and spatial. The temporal type of technological component of yield variability shows how much yield change differs from the beginning level of yield in a region. The spatial type shows the contribution of a particular region to the nation-wide rate of yield change. They are represented by the following expressions:

$$T_i = \frac{Y_{E,i} - Y_{B,i}}{Y_{B,i}} \quad \text{, for the temporal} \qquad (5)$$

$$S_i = \frac{Y_{E,i} - Y_{B,i}}{Y_{E,USSR} - Y_{B,USSR}} \quad \text{, for the spatial}$$

where $Y_{B,i}$, $Y_{E,i}$, $Y_{B,USSR}$, $Y_{E,USSR}$ are yield (Q/HA) for economic region (i), for the Soviet Union (USSR), and for the beginning (B) and end (E) of the period.

Barley yields in the five economic regions from 1945 through 1978 are shown in Figures 13-17. Estimates for regression coefficients and some statistics for equations of trend are given in Table 4. The model fitted for trend is $\hat{Y}_t = a_0 + a_1t + a_2t^2$ where $\hat{Y}_t$ is the estimated yield, t = year-1944 and a_0, a_1 and a_2 are parameters estimated by least squares methods.

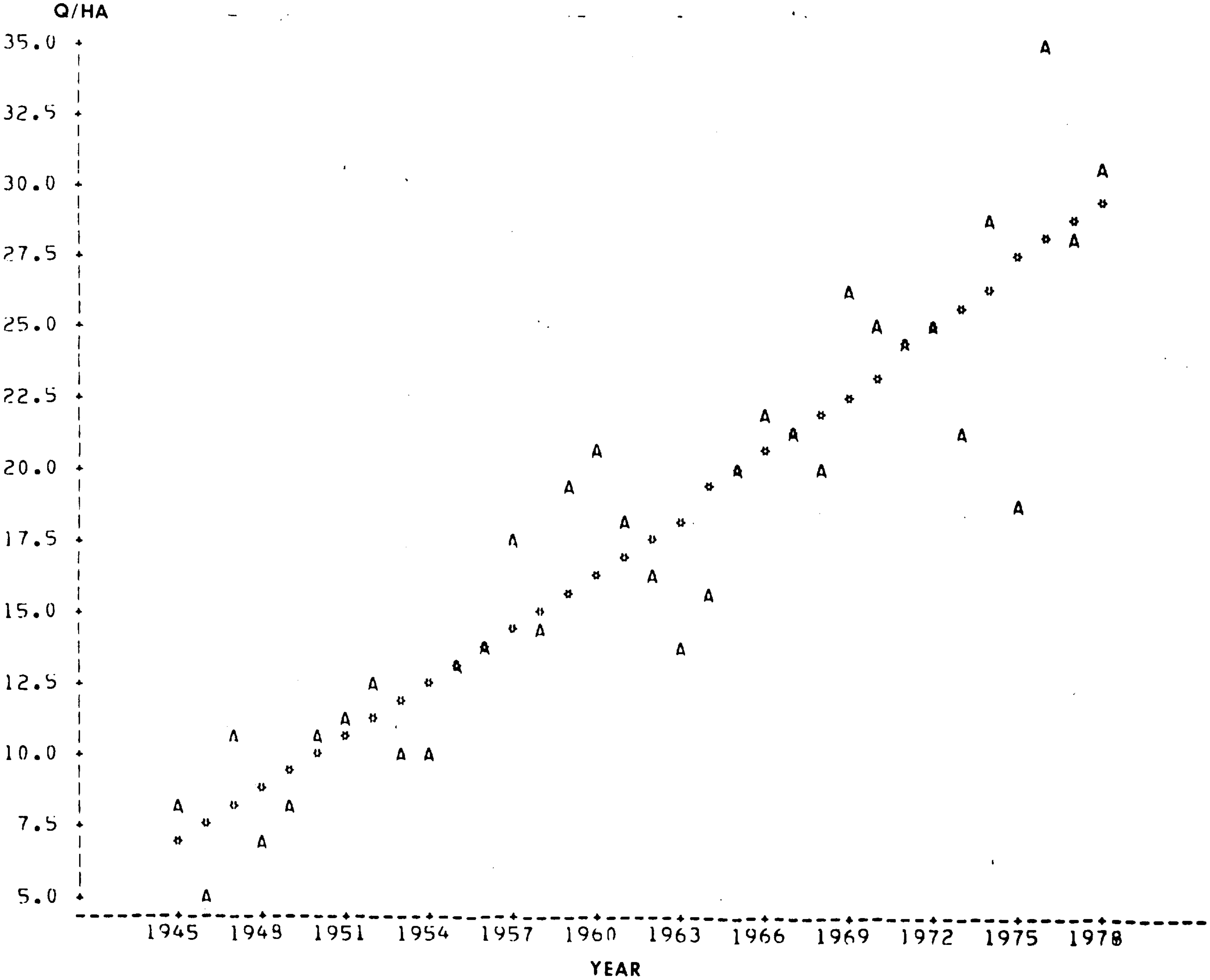

Figure 13. Barley yield (A) and trend (*) in South-West economic region, USSR, 1945-1978.

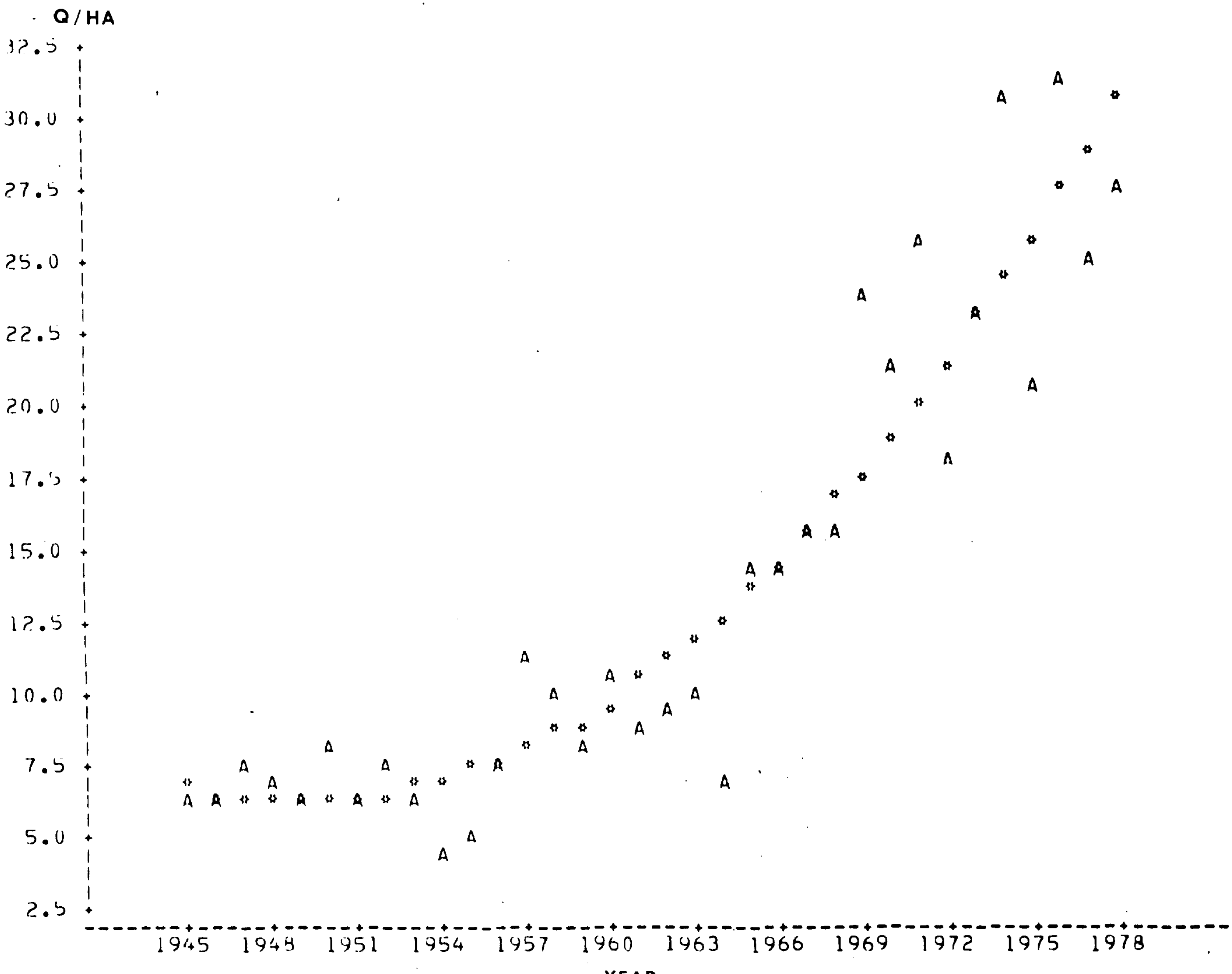

Figure 14. Barley yield and trend in Belorussia economic region, USSR, 1945-1978.

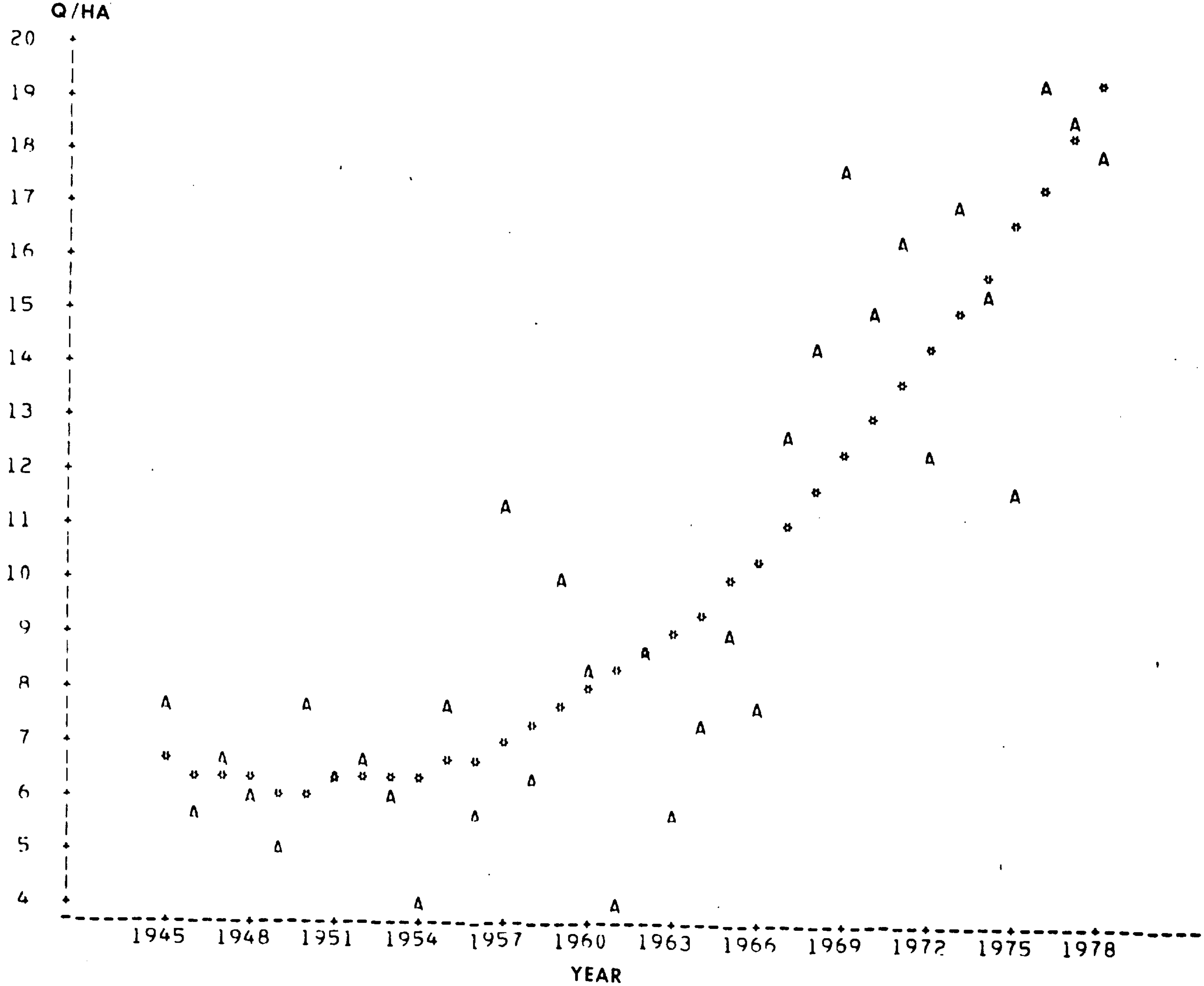

Figure 15. Barley yield and trend in Volgo-Vyatka economic region, USSR, 1945-1978.

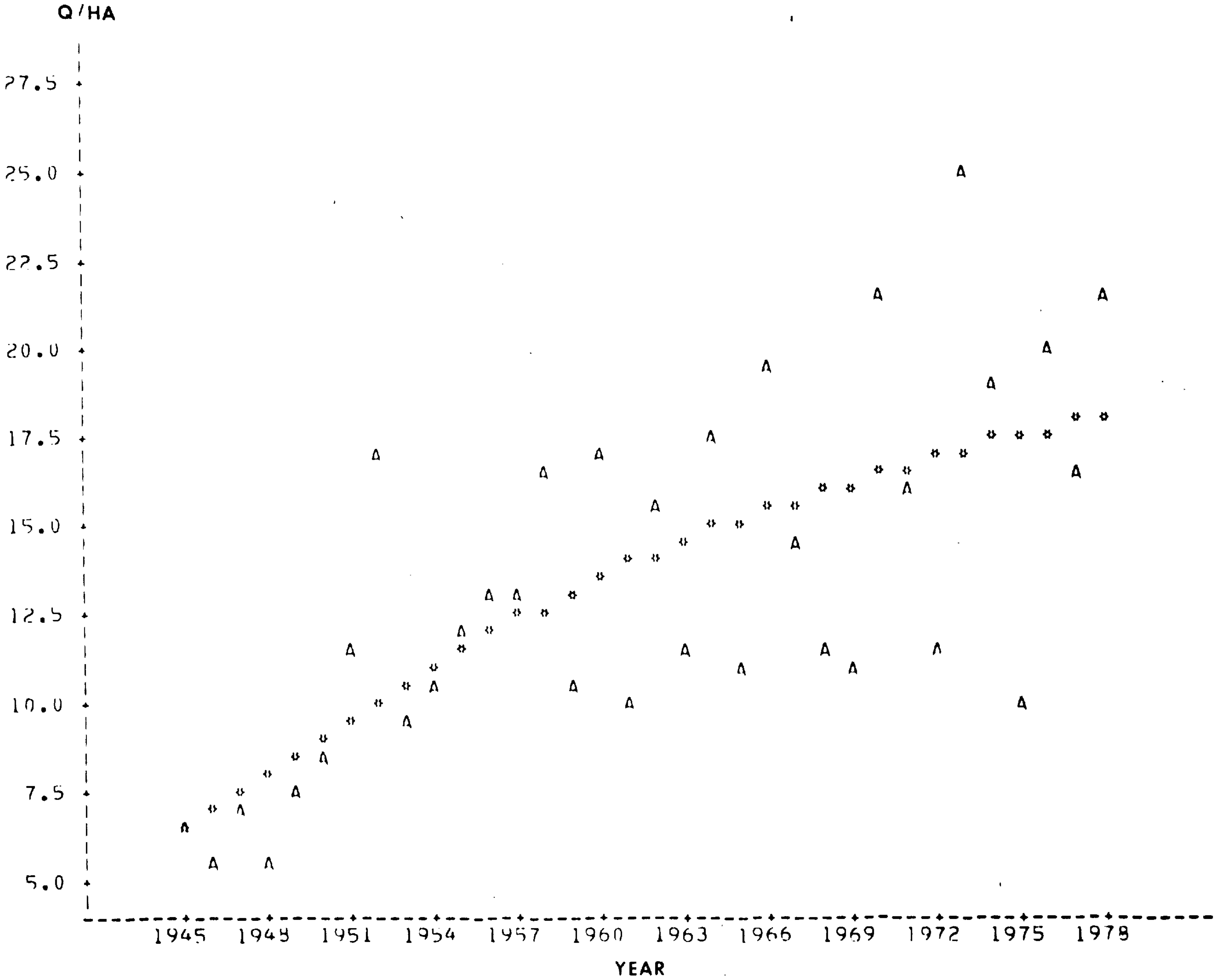

Figure 16. Barley yield and trend in North Caucasus economic region, USSR, 1945-1978.

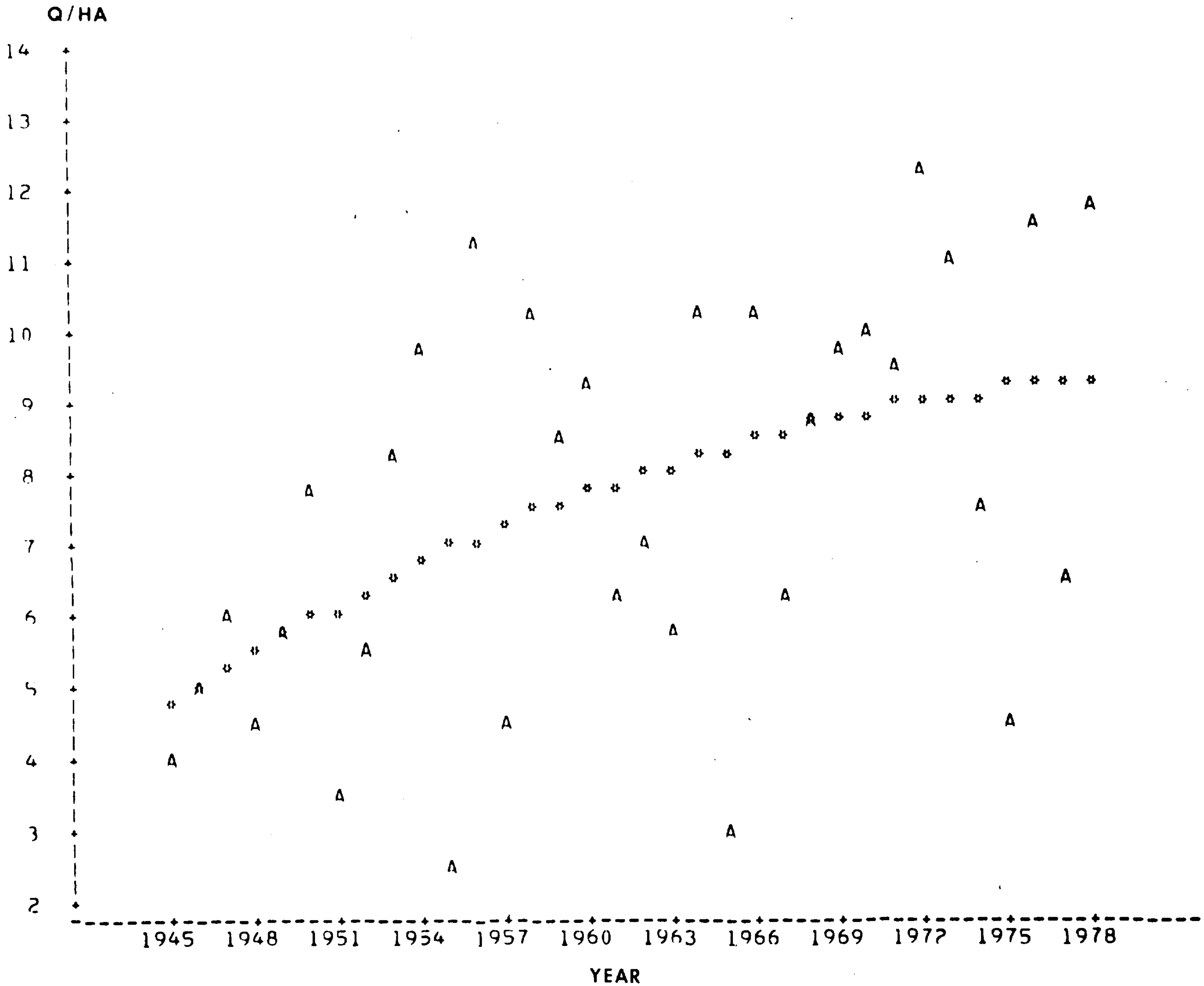

Figure 17. Barley yield and trend in Kazakh economic region, USSR, 1945-1978.

Table 4. Estimates for regression coefficients and some statistics for equations of the trend shown in Figures 13-17

Economic Region	Estimates for regression coefficients			R^2	Percentage of yield variance explained by:	
	a_o	a_1	a_2		Trend alone (technology)	Remainder (weather, other)
(1) South-West	6.460	.566	.0033	.850	85	15
(2) Belorussia	6.994	-.302	.0293	.880	88	12
(3) Volgo-Vyatka	6.694	-.188	.0164	.762	76	24
(4) North Caucasus	5.912	.576	-.0065	.496	50	50
(5) Kazakh	4.546	.247	-.0032	.244	25	75

As seen in Table 4, the R^2's are very high (0.76-0.88) in the wetter regions, but are considerably smaller (0.24 and 0.50) in regions with water deficits. The most interesting fact is that changes in the technology share of barley yield variance correspond to water deficits of the regions. In the Non-Chernozem area (Volgo-Vyatka, Belorussia) where applied technology and climate were complementary over the past 34 years, the technology-related yield variance explains 76-88 percent of the total yield variance. In the areas with substantial water deficits (North Caucasus, Kazakh), this proportion is reduced to 25-50 percent.

There were substantial increases in barley yields in all five economic regions during the entire 34 year period, but these increases were not uniform in relation to the natural and economic resources of the regions [12]. The figures of yield change are presented in Table 5. The beginning levels of yield in 1945, a period of poor technology, were very low, from 4.8 to 7.5 quintals per hectare. The differences among the regions in this period were not great. The ratio between the two regions with the highest and lowest barley yield was 1.6 in 1945. By 1978, this ratio had doubled (3.3) because of variations in the rate of yield increase in response to technological improvement for different climatic and soil conditions. In regions with very good natural resources, particularly water supply (South-West, Belorussia), the increment of barley yield was 22-23 quintals per hectare over the 34 years. This is in excess of 300 percent relative to the 1945 level of yield (temporal rate of yield growth). In regions where natural resources are limited (water deficit in North Caucasus or deficientheat in Volgo-Vyatka) the increment of barley yield totaled only 12-13 quintals per hectare or around 200 percent relative to the beginning level of yield in these regions. The lowest increment of barley yield, only 4.4 quintals per hectare, was in Kazakh which has a considerable deficit of water throughout the planted barley area. The temporal increase in barley yields did not even reach 100 percent.

The spatial rates of barley yield growth in these regions have the same type of change as the temporal rate. Southwest and Belorussia economic regions had the highest rates of increase, 1.8 times higher than the average for the USSR. For Volgo-Vyatka and North Caucasus, the rates were very close to the national average. And, Kazakh region again had the lowest rate of increase, only 35 percent of the average for the USSR.

Table 5. Estimates of barley yield change in the USSR and selected economic regions over the period 1945-1978

Region	Yield predicted from trend (Q/HA) for year			Yield increase (Q/HA)			Rate of yield growth (%)					
							Temporal (Y_{1945}=100)			Spatial		
	1945	1961	1978	1945-1978	1945-1961	1961-1978	1945-1978	1945-1961	1961-1978	1945-1978	1945-1961	1961-1978
South-West	7.0	17.0	29.5	22.5	10.0	12.5	321	143	178	177	172	181
Belorussia	7.5	9.7	30.6	23.1	2.2	20.9	308	29	279	182	38	303
Volgo-Vyatka	6.5	7.3	19.3	12.8	0.8	12.0	197	12	185	101	14	174
North Caucasus	5.5	13.8	17.9	12.4	8.3	4.1	225	151	74	98	143	59
Kazakh	4.8	7.8	9.2	4.4	3.0	1.4	92	62	30	35	52	20
USSR	5.0	10.8	17.7	12.7	5.8	6.9	254	116	138	100	100	100

The rates of spatial and temporal yield change were not uniform during the 34-year period for areas with different natural conditions. This can be seen when comparing these components for two parts of the 1945-1978 period, namely 1945 through 1961 and 1961 through 1978 (Table 5). For North Caucasus and Kazakh economic regions, which have a limiting natural water supply, both the temporal and spatial yield growth rates declined in 1961-1978 compared to 1945-1961. In these regions, the applied technology in the second half of the period and natural resources were not entirely compatible. Regions which have good water resources South-West, Belorussia, Volgo-Vyatka, had a very low rate (both temporal and spatial) of barley yield increase during the first part of the period. In the second half of the period this rate increased dramatically in Belorussia and Volgo-Vyatka. Such a sharp increase means that the natural resources of these regions were much more complementary to the technology applied in the later period. In the South-West economic region natural resources were also complementary to the applied technology. However, the increments of the rate of barley yield growth for the later period showed more modest rates of increase (temporal and spatial) compared to the earlier period. This may indicate that the applied technology in South-West region could be approaching a level sufficiently compatible with the existing climatic conditions in the region and that further growth in the rates of yield increase are unlikely. The quadratic coefficient (a_2) in the polynomial equation of the barley yield series of trend for this region has a positive sign (Table 4). However, the magnitude of this coefficient (0.0033) is much less than that of the same coefficient for Belorussia or Volgo-Vyatka region. In North Caucasus and Kazakh economic regions, which have negative coefficients (-0.0065 and -0.0032, respectively), the process of deceleration of the rate of the technology-induced yield growth has begun. Thus, there have generally been two results when advanced technology was applied. In some cases, the applied technology and natural resources were not entirely compatible. Accordingly, yield increases have been limited in the second part of the period. In other cases, natural resource factors have been complementary to the improved technology and yield levels have responded.

Weather-Related Yield Fluctuations

An impression of the amount of variability in barley yield which results from changes in weather can be obtained from the coefficients of variation and extreme yield deviations from trend. These are shown in Table 6. As seen in this table, the coefficients of variation for regions with good water resources (South-West, Belorussia) are approximately two-thirds as large as for regions with deficient water resources (North Caucasus, Kazakh). But extreme absolute fluctuations in yield over the past 34 years were observed in all the regions. They were not as large in Kazakh and Volgo-Vyatka economic regions, which had lower average yield levels. One of the noticeable, and important, facts is that most of the extreme values in variation of barley yield around the trend occurred over the second part of the considered period. It supports the idea previously discussed [10], that with the improvement of technology and enhancement of the general yield level, absolute values of the weather-induced yield fluctuation also increase.

Table 6. Variations in barley yields from trend and extreme deviations from trend in selected years, by selected economic regions, USSR, during 1945 through 1978

Economic Region	Coefficient of variation (%)	Extreme deviations of yield (from trend)			
		Q/HA	Year	Q/HA	Year
(1) South-West	16.8	6.9	1976	-8.3	1975
(2) Belorussia	21.6	6.1	1974	-5.8	1964
(3) Volgo-Vyatka	23.6	5.3	1969	-5.1	1975
(4) North Caucasus	27.5	8.1	1973	-7.3	1975
(5) Kazakh	32.9	4.2	1956	-5.3	1965

The frequencies of various deviations from trend are shown in Table 7. The largest deviations greater than plus or minus 5 Q/HA of barley yield from trend considered separately do not have frequent occurrences. The frequency of either extreme in any of the five regions has not exceeded 10 percent. However, as shown in the table, there have been substantial fluctuations from trend of lesser magnitude.

Table 7. Frequency of years (%) with barley yield variation from trend in selected economic regions, USSR, 1945-1978

Economic Region	Deviation of yield from trend (Q/HA)						
	Less than -5.0	-5.0 - -3.1	-3.0 - -1.1	-1.0 - +1.0	1.1 - 3.0	3.1 - 5.0	Greater than 5.0
(1) South-West	3	9	21	32	20	12	3
(2) Belorussia	6	9	20	32	22	9	2
(3) Volgo-Vyatka	3	9	18	38	26	6	0
(4) North Caucasus	8	9	20	24	18	12	9
(5) Kazakh	3	9	21	26	35	6	0

The variability of yield departures from trend depends almost entirely on the general climate of individual regions and weather variations. Of all variables, precipitation is the most important factor. Variability of cereal yields was observed to be well correlated with the climatic norm of precipitation for the main part of the growing season [12]. In the case of barley for the discussed regions, the same type of relationship was obtained. Among the examined regions, the Kazakh economic region had the largest yield variability (32.9%, Table 6) due to a very low norm of precipitation during April-September (230 mm). South-West economic region had the smallest variability (16.8%) due to a very large norm of precipitation (465 mm).

Estimate of Weather Input Into Yield Variation

Influence of weather on crops is not uniform during the growing and pre-growing period. Some of the factors affect crops directly within a short period of time, but others have a delayed influence on crops and the crop response may not be apparent until later.

A general understanding of the weather effect on barley yields can be obtained by comparing the weather data for years with extremely different yields. Precipitation and temperature for years in which extreme yields were reported are shown in Table 8. For all regions except South-West, differences in yields can be explained almost entirely by the differences in precipitation and temperature over spring and summer periods. Normally, weather patterns with increased precipitation and decreased temperature during March-May and June-August create favorable conditions for barley growth and hence for high barley yield. The opposite weather patterns have adverse effects on barley growth and yield. These two general types of weather patterns are typical for regions with either poor or good water supply. In South-West economic region, the year with the largest negative deviation of yield from trend (1975) turned out to be extremely wet (Table 8). However, in 1963 and 1973, years which had the second and third greatest negative deviation of yield from the trend in this region, the deficit of rainfall and increased temperature in the period of March-August were the main factors which reduced barley yield.

The analysis of correlation coefficients for barley yield (barley yield is expressed in terms of its departure from trend as $Y/\hat{Y}_t$, where Y is the reported yield and $\hat{Y}_t$ is the trend analysis predicted value) with monthly precipitation and average temperature confirmed the conclusions mentioned above. These coefficients are shown in Figures 18 a and b. Judging by these coefficients, three general statements can be made. First, in regions with water deficits (Kazakh, North Caucasus), the dependence of barley yield departures from trend on weather is defined much better than for regions with good water supplies (South-West, Belorussia, Volgo-Vyatka). In the water deficient regions one can distinguish several periods when barley yield shows considerable sensitivity to the change in precipitation and temperature. The regions with good water supplies do not show such a clear sensitivity. Accordingly, the results of modeling the yield departure from trend for the latter group could be expected to be less reliable than for the first group. The second general statement is that there are some periods of the year when the direction of response of barley productivity to weather variations is quite similar even for regions with different climatic conditions. Such periods are May through September (harvest year) with a negative response of barley yield to average monthly temperature, and fall and beginning of winter (year preceding harvest) with a negative but less defined response of yield to temperature. The third statement is that even though the importance of weather for barley productivity is not uniform over the entire 19-month period in magnitude or in sign, almost every month's weather provides some information which can be used as a signal for barley yield assessment.

The general climatic features of a region also cause some unique responses of barley yield to weather. In regions with a deficit of water, the weather in May and June of the harvest year considerably defines productivity of barley.

Table 8. Precipitation and temperature in years with the extreme values of barley yield deviation from trend in selected economic regions, USSR

Economic Region	Type of Year	Year	Yield Deviation from trend (Q/HA)	Total precipitation (mm)			Average Temperature (°C)		
				Dec.-Feb.	March-May	June-Aug.	Dec.-Feb.	March-May	June-Aug.
South-West	Favorable	1976	6.9	101	147	186	-5.6	7.0	16.4
	Adverse	1975	-8.3	88	176	265	-0.4	10.5	19.3
Belorussia	Favorable	1974	6.1	68	60	278	-4.2	4.8	15.6
	Adverse	1964	-5.8	66	121	144	-8.2	3.8	18.0
Volgo-Vyatka	Favorable	1969	5.3	56	93	208	-16.8	1.1	15.2
	Adverse	1975	-5.1	80	79	159	-8.4	7.5	16.7
North Caucasus	Favorable	1973	8.1	67	139	177	-2.0	10.2	19.9
	Adverse	1975	-7.3	119	123	113	-1.4	12.0	23.3
Kazakh	Favorable	1956	4.2	56	83	91	-15.2	4.1	20.2
	Adverse	1965	-5.3	50	59	74	-9.2	6.5	20.9

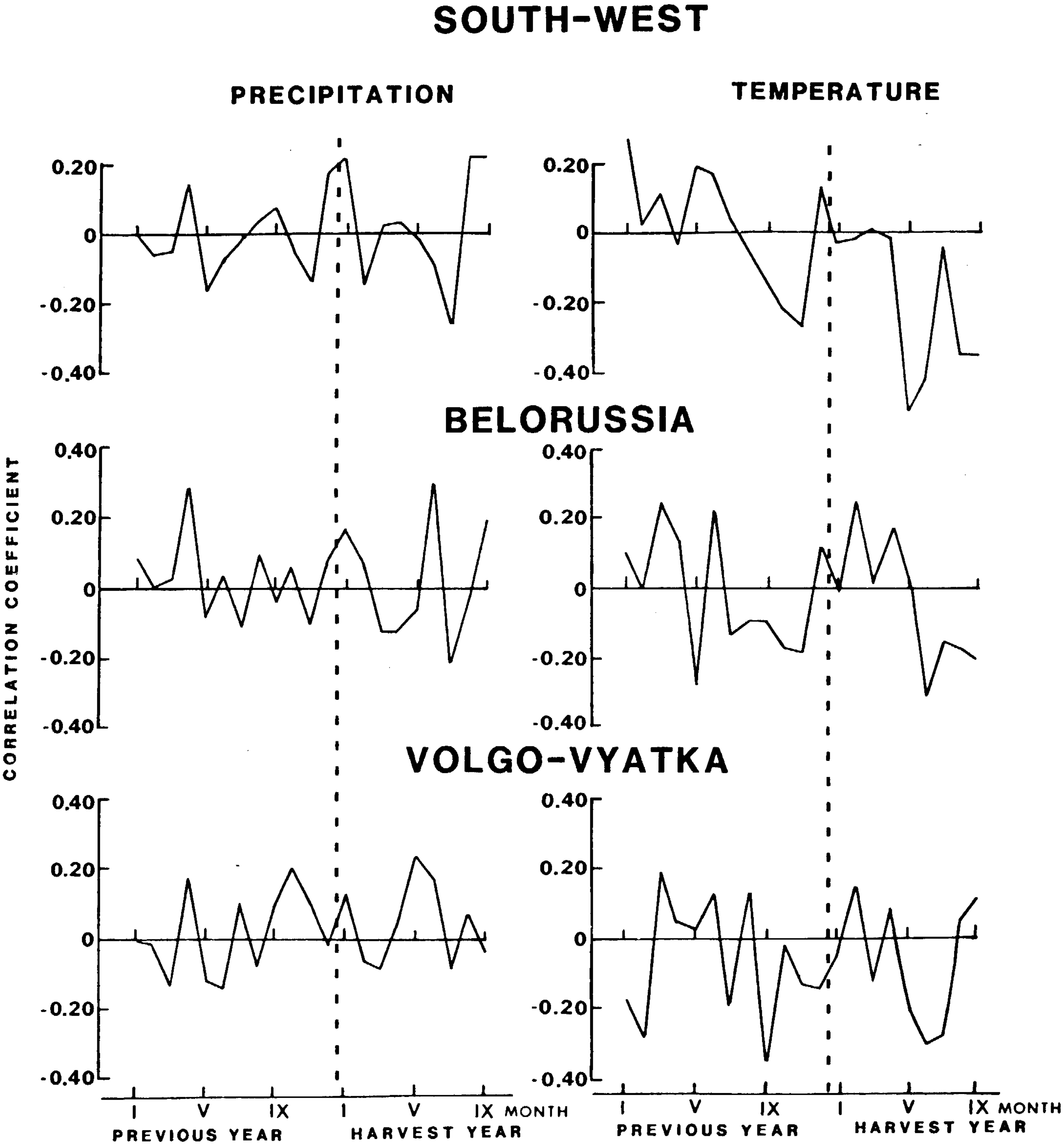

Figure 18a. Monthly coefficients of correlation between departures of barley yield from trend with precipitation and temperature in good water regions (South-West, Belorussia and Volgo-Vyatka) of the USSR.

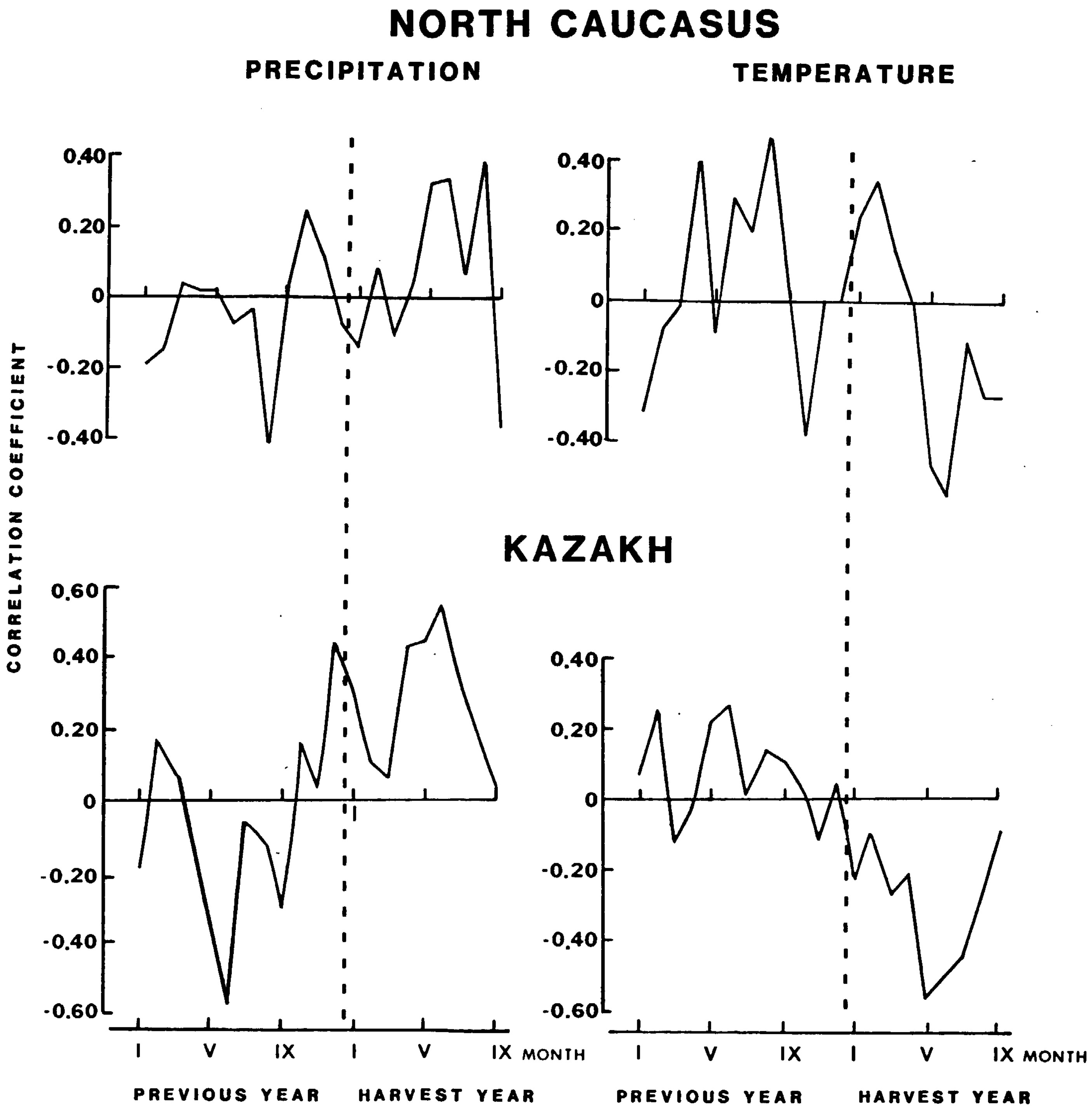

Figure 18b. Monthly coefficients of correlation between departures of barley yield from trend with precipitation and temperature in bad water regions (North Caucasus and Kazakh) of the USSR.

But in the Kazakh economic region, where the deficit of water is more substantial than in the North Caucasus region, April rainfall also plays an important role for barley yield. The importance of precipitation in July for barley in Kazakh is considerably less than for the previous months, but July precipitation is still important there, as it is the month in which barley forms the head and starts to fill grain. High temperatures during May, June and July (for North Caucasus, May and June) are unfavorable for head and grain formation and cause a decrease in barley yield. Fall precipitation in North Caucasus and fall and winter precipitation in Kazakh make some positive contribution to barley yield.

Although Belorussia and Volga-Vyatka economic regions have good water resources, the positive effect of rainfall on barley yield trend departures can still be seen even in June, and also in July in Volgo-Vyatka. High temperatures during this period have an adverse effect on barley yield. In the fall and winter seasons preceding the year of harvest, some positive influence of precipitation and negative influence of temperature on the magnitude of barley yield can also be seen in these regions.

In the South-West economic region the amount of rainfall during summer normally exceeds the requirements of barley for water. The excessive rainfall during this period usually has a negative effect on barley yield. But the average summer temperature there usually exceeds the optimum temperature for barley growth and causes a decrease in barley production.

Based on correlation coefficients, the relative significance of weather in different periods of the agricultural year has been computed. As seen in Table 9, April, May and June are the most important months for barley; the weather in these months accounts for 30-50 percent of weather-induced yield variability, whether a region has a deficit or an adequate supply of water. Periods prior to April and after June are less important for barley, but still, for some of the regions, weather in one of these quarters can explain 20-35 percent of yield variability. Accordingly, weather in April, May and June can be used as a predictor in models for most of these regions. The only exception is South-West economic region. There rainfall in April, May and June provides very little information about fluctuations in barley yields.

The values in Table 9 are average weights based on correlations over the thirty-four year period of observation. In a particular year, these figures can be quite different. Thus, in years with a very severe drought the importance of weather from April to June can increase considerably.

Special consideration should be given to the problem of using weather information from the previous year in forecasting barley yield. Sometimes extreme weather in one year is compensated for to some extent by the weather in the following year. An interaction of such an oscillating weather pattern with some of the biological processes intensifies their influence on crop growth and production. For example, good weather in one year causes high crop production. As a result, the crop takes out of the soil an increased amount of nutrients. In spite of the application of fertilizers, the amount of nutrients available for the next year's crop will be decreased since there is some delay in restoring lost soil fertility. Because of this decreased soil fertility, even slightly unfavorable or even normal weather usually causes noticeable yield decreases. The two-year cycle of weather-crop-soil interaction can especially intensify when weather conditions in the first year are extremely adverse.

Table 9. Estimate (%) of the weather significance for barley yield in selected economic regions, USSR, 1945-1978

Economic Region	Factor	October-December	January-March	April-June	July-September
South-West	Precipitation	23	23	10	44
	Temperature	25	3	41	31
Belorussia	Precipitation	15	25	32	28
	Temperature	26	16	30	28
Volgo-Vyatka	Precipitation	27	21	36	16
	Temperature	16	20	38	26
North Caucasus	Precipitation	19	14	31	36
	Temperature	14	25	37	24
Kazakh	Precipitation	20	16	47	17
	Temperature	7	21	43	29

As seen in Figure 18, there is a fairly noticeable correlation in almost all regions between summer temperature of the previous year with the next year's barley yield. This correlation is particularly well defined in regions with a deficient water supply (North Caucasus, Kazakh). The correlation between rainfall of the previous summer and barley yield is also very well defined in these regions. In regions with a good water supply, this correlation is not well defined. But even in these regions there are some possibilities for using weather of the previous year as a predictor. But we should bear in mind that if weather unfavorable for crop growth occurs two years in a row, then the weather of the previous year used as a predictor could distort the yield estimate.

Numerical Assessment of Weather-Yield Interaction

As shown above, monthly air temperature prevails in its influence on barley yield for some periods, at other times the precipitation prevails. Thus, their combined effect on barley yield has been estimated. First, the number of variables attributed to different months was decreased in accordance with the methodology described earlier. Precipitation and temperature for each month were weighted together according to the values of the Pearson correlation coefficients for barley yield departures from trend with these variables. Table 10 shows these weights.

The product of each month's precipitation (or temperature) and its corresponding weight are summed to compute the index for precipitation (or temperature). These variables have a fairly strong relationship to the ratio of yield departures from trend (Table 11). Comparison of the correlation coefficients between barley yield departures from trend and the precipitation and temperature index variables (Table 11) with the correlation coefficients of yield departures from trend with precipitation and temperature variables for separate months (Fig. 18) shows a substantial increase in the strength of the relationship for aggregated index variables. This is true even for regions with good water supply. In regions with a deficit of water, correlation coefficients for yield departure from trend with the index-variables are very high (0.70-0.85, Table 11).

Based on the index-precipitation and index-temperature, different versions of models designed to reflect the relationship to barley yields were developed and tested to obtain the best combination of predictors. This test was based on statistical assessments used in regression analysis and also on the condition index to diagnose collinearity [4]. The results of this test are shown in Table 12.

As seen in this table, R^2 and RMSE (root mean square error for yield departures expressed as percentage of trend) do not differ considerably for different sets of predictors within the same economic region. However, the differences in the condition index are very great. Normally, a condition index less than ten indicates freedom from collinearity [4]. A large condition index indicates a strong relationship between at least some of the supposedly independent variables. In such a case estimated regression coefficients may have inflated values which make the results of modeling unreliable.

Although the model based on five predictors, P, T, P^2 and T^2 and product of P and T has the largest R^2, it also has the largest condition index. Eliminating only one of the predictors decreases this index substantially. It is interesting that the best set of predictors in terms of the optimal combination of condition index, R^2 and RMSE is almost the same in each region even though the climatic conditions differ.

Predictors which seem to provide the best results for Belorussia and Kazakh include P, P^2 and T^2. For Volgo-Vyatka and North Caucasus, P and T appear best, and P and T^2 constitute the preferred combination for South-West. We should also notice that the various combinations of precipitation and temperature variables as compared with the single variables (P or T, see Table 11) had larger R^2's.

Table 13 presents regression coefficients for models which appear to have the optimal combination of the variables with respect to the R^2, RMSE and the condition index. Unfortunately, models for regions where water is not a limiting factor for barley growth explain only about 45-55 percent of the variance of the ratio yield departure from trend. But for regions where water is a much more limiting factor, this percentage increases to 70-80 percent, which substantially enhances the reliability of predictions based on these models.

Table 10. Weights ($K*10^{-4}$) used to combine monthly precipitation and temperature into index-precipitation and index-temperature for 19-month period in selected economic regions, USSR

Economic Region	Weather Parameter	Previous Year												Harvest Year						
		JA	FE	MA	AP	MY	JU	JL	AU	SP	OC	NO	DE	JA	FE	MA	AP	MY	JU	JL
South-West	Precip.	-26	-346	-282	949	-953	-430	-88	248	425	-269	-750	965	1182	-812	124	155	-63	-506	-1428
	Temp.	967	122	419	-106	714	625	128	-179	-526	-820	-994	477	-102	-60	27	-69	-1951	-1558	-159
Belorussia	Precip.	222	-197	127	1477	-524	411	-350	506	-409	214	-432	157	504	101	-755	-481	-438	1810	-884
	Temp.	329	24	770	645	-1063	622	-571	-364	- 20	-737	-615	331	-168	811	53	893	8	-1235	-742
Volgo-Vyatka	Precip.	310	204	-458	826	-676	-766	573	-471	473	825	331	193	1074	- 71	-140	46	1183	935	-445
	Temp.	-504	-936	595	21	263	471	-502	396	-1237	102	-516	-508	- 85	672	-470	227	-518	-1116	-852
North Caucasus	Precip.	-614	-527	246	140	212	-278	- 55	-1458	359	677	650	- 63	-405	495	-282	330	1518	1380	310
	Temp.	-695	-146	- 14	1010	-272	662	535	1072	81	-745	51	215	638	834	330	77	-1132	-1311	-178
Kazakh	Precip.	-432	332	140	-160	-658	-1233	-133	-242	-639	330	77	922	637	241	120	903	966	1161	673
	Temp.	180	681	-335	- 27	584	735	42	391	301	87	-300	133	-632	-248	-728	-573	-1541	-1274	-1208

Table 11. Correlation coefficients and corresponding coefficients of determination (R^2) for barley yield departure from trend with index-precipitation and index-temperature in selected economic regions, USSR

Economic Region	Index-Precipitation			Index-Temperature		
	Correlation Coefficient	R^2	Probability*	Correlation Coefficient	R^2	Probability*
South-West	0.425	.181	0.0140	0.697	.486	0.0001
Belorussia	0.542	.294	0.0011	0.514	.264	0.0022
Volgo-Vyatka	0.446	.199	0.0092	0.648	.420	0.0011
North Caucasus	0.708	.501	0.0001	0.810	.656	0.0001
Kazakh	0.849	.721	0.0001	0.706	.498	0.0001

*Probability of obtaining a larger correlation coefficiently (or R^2) estimate under the hypothesis that no real relationship exists.

Table 12. Alternative model versions which use various precipitation and temperature index variables to predict barley yield departures from trend in selected economic regions, USSR

Economic Region	Version	Predictors used in model					Estimates for equations		
		P	P^2	T	T^2	P*T	R^2	RMSE %	Condition Index
South-West	1	✓	✓	✓	✓	✓	0.560	11.6	295.5
	2	✓	✓	✓		✓	0.554	11.4	82.4
	3	✓	✓	✓			0.553	11.3	26.5
	4	✓	✓		✓		0.558	11.2	24.5
	5	✓		✓			0.550	11.1	19.8
	6*	✓			✓		0.557	11.1	10.2
Belorussia	1	✓	✓	✓	✓	✓	0.454	16.0	487.5
	2	✓	✓	✓		✓	0.454	15.7	32.4
	3	✓	✓	✓			0.454	15.4	22.2
	4*	✓	✓		✓		0.454	15.4	11.7
	5	✓		✓			0.428	15.5	18.7
	6	✓			✓		0.424	15.6	9.4
Volgo-Vytka	1	✓	✓	✓	✓	✓	0.530	18.4	109.4
	2	✓	✓	✓		✓	0.476	19.0	61.7
	3	✓	✓	✓			0.474	18.7	29.2
	4	✓	✓		✓		0.416	19.7	28.8
	5*	✓		✓			0.463	18.6	11.9
	6	✓			✓		0.407	19.5	8.5
North Caucasus	1	✓	✓	✓	✓	✓	0.750	13.8	28.9
	2	✓	✓	✓		✓	0.744	13.7	26.6
	3	✓	✓	✓			0.731	13.8	23.7
	4	✓	✓		✓		0.654	15.7	24.9
	5*	✓		✓			0.726	13.7	6.4
	6	✓			✓		0.646	15.6	5.8
Kazakh	1	✓	✓	✓	✓	✓	0.800	15.4	190.9
	2	✓	✓	✓		✓	0.794	15.4	62.8
	3	✓	✓	✓			0.788	15.3	16.1
	4*	✓	✓		✓		0.787	15.4	11.2
	5	✓		✓			0.768	15.8	14.2
	6	✓			✓		0.769	15.7	8.8

*Model with the best combination of the estimates for the equation.

Table 13. Regression coefficients and associated statistics for models in selected economic regions, USSR

Economic Region	Coefficients for					Condition Index	R^2
	P	P^2	T	T^2	Intercept		
South-West	0.578			-1.093	151.476	10.2	0.56
Belorussia	0.422	0.060		-0.841	121.380	11.7	0.45
Volgo-Vyatka	0.857		18.069		135.929	11.9	0.46
North Caucasus	1.456		14.031		74.626	6.4	0.73
Kazakh	6.485	-0.153		-1.444	87.104	11.2	0.79

Correlation coefficients of barley yield trend departures and index-precipitation and index-temperature can be used as weights to combine the two indices into one index-variable, PT. To obtain the index-PT variable, the index-precipitation and index-temperature variables were each expressed in standardized form. The calculations were based on an adaptation of equation (5). Weights for these two variables, calculated for two types of coefficients (equations 3 and 4 are adapted to weights for separate variables rather than months), are shown in Table 14.

Table 14. Weights used to combine standardized index-precipitation and index-temperature into index-PT variable, based on the Pearson correlation coefficient and the coefficient of determination (r, r^2) in selected economic regions, USSR

	r		r^2	
Economic Region	Index-precip.	Index-temp.	Index-precip.	Index-temp.
South-West	.379	.621	.271	.729
Belorussia	.513	.487	.526	.474
Volgo-Vyatka	.408	.592	.321	.679
North Caucasus	.466	.536	.433	.567
Kazakh	.546	.454	.591	.409

As seen in Table 14, weights calculated based upon r^2 define a greater input of the combined weather variable (index-precipitation or index-temperature) into index-PT for the variable whose influence on barley yield is stronger.

Statistical assessments for models based on the index-PT variable are presented in Table 15. Comparison of model assessments presented in Table 15 with those presented in Table 13 shows that differences between R^2 for the models are not significant. But the difference between condition indices is quite considerable. For the models presented in Table 13, condition indices were slightly above the limit of 10 which defines the existence of some collinearity (North Caucasus is the only exception). For the models in Table 15 these condition indices were far below this limit and ranged from 1.77 to 2.50. For South-West, North Caucasus and Kazakh economic regions, weighting of index-variables proportional to the r^2 showed a slight increase in R^2 for the model. For Belorussia and Volgo-Vyatka economic regions, it showed a slight decrease in model R^2.

Table 15. Estimates for regression coefficients and some statistical assessments for models in selected economic regions, USSR

Economic Region	Weight based on r					Weight based on r^2				
	Estimates of regression Coefficients					Estimates of regression Coefficients				
	Intercept	PT	PT^2	R^2	Condition Index	Intercept	PT	PT^2	R^2	Condition Index
South-West	100.013	14.213	-0.759	0.55	1.77	100.878	15.780	-2.121	0.62	1.89
Belorussia	99.159	16.324	1.471	0.43	1.94	99.395	15.219	1.251	0.38	2.05
Volgo-Vyatka	99.610	17.802	4.302	0.48	2.30	100.809	18.165	2.678	0.45	2.31
North Caucasus	96.164	22.235	4.580	0.75	2.50	98.565	24.924	1.551	0.78	2.42
Kazakh	104.871	30.852	-5.550	0.78	2.24	104.105	30.444	-4.452	0.80	2.22

EVALUATION OF MODELS AND PREFERRED MODEL RECOMMENDATION

The criteria for model evaluation have been very well developed and documented [36]. The referenced document considers evaluation of the following basic characteristics: yield reliability, objectivity, consistency with scientific knowledge, adequacy, timeliness, minimum costs, simplicity and accuracy of current measures of modeled yield reliability. Various indicators of model evaluation and particularly yield reliability are the primary topic of this section.

Results of Dependent Testing

To begin with model reliability was estimated by comparing the actual yield with the calculated yield based on the same data used for model development (dependent test). Figures 19-23 show the actual barley yield and yield estimates based on the models presented in Tables 4 and 13. In general, these figures show good agreement between actual and calculated barley yield both for regions with good and deficient water supplies. In 85 to 94 percent of years the actual and predicted yields agreed in the direction of change from the previous year. The average bias for models is close to zero for Belorussia, North Caucasus and Kazakh economic regions and differs slightly from zero in South-West and Volgo-Vyatka regions. The model underestimates barley yields on the average by 0.2 quintals per hectare for the South-West region, and overestimates yield by an average of 0.3 quintals per hectare for the Volgo-Vyatka region (Table 16). The distribution of departures of model predicted yields from reported yields presented in Table 16 also shows good model performance in the dependent test.

Table 16. Average bias and probability (%) of years of various departures of model predicted from reported yields in selected economic regions, USSR (dependent test)

Economic Region	Average Bias (Q/HA)	Departure of model predicted from reported yields (Q/HA								
		Less than -4	-4, -3.1	-3, -2.1	-2, -1.1	±1	1.1, 2	2.1, 3	3.1, 4	Greater than 4
South-West	-0.21		6	9	15	49	18	3		
Belorussia	-0.03	6	3	0	18	55	9	3	3	3
Volgo-Vyatka	0.33			3	18	52	15	6	3	3
N. Caucasus	-0.06	3	3	9	9	49	15	12		
Kazakh	0.02				18	64	18			

More than 50 percent of the differences between actual and calculated barley yield were within the range of ± 1 quintal per hectare. The highest percentage of differences in this range occurred in the Kazakh economic region (64%). The Kazakh model, in general, showed the best performance.

Independent Testing Methods

Good model performance in the dependent test is, however, not enough to make a decision about model reliability. In this respect the independent test information is more valuable. To obtain independent test data of model performance, a modified bootstrap technique has been used. This technique consists of developing a model from an earlier base period and applying this model to the data of

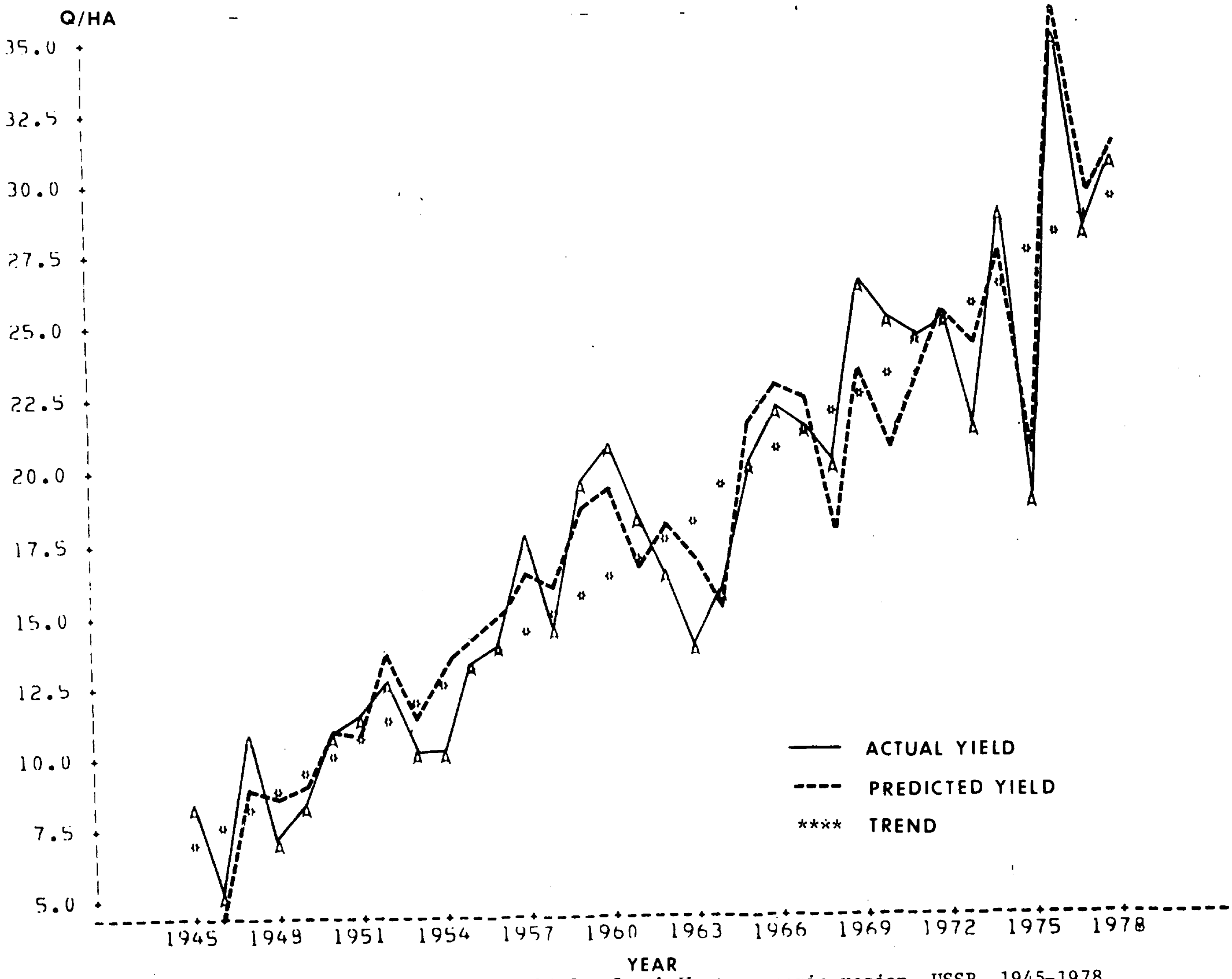

Figure 19. Actual and predicted barley yield for South-West economic region, USSR, 1945-1978.

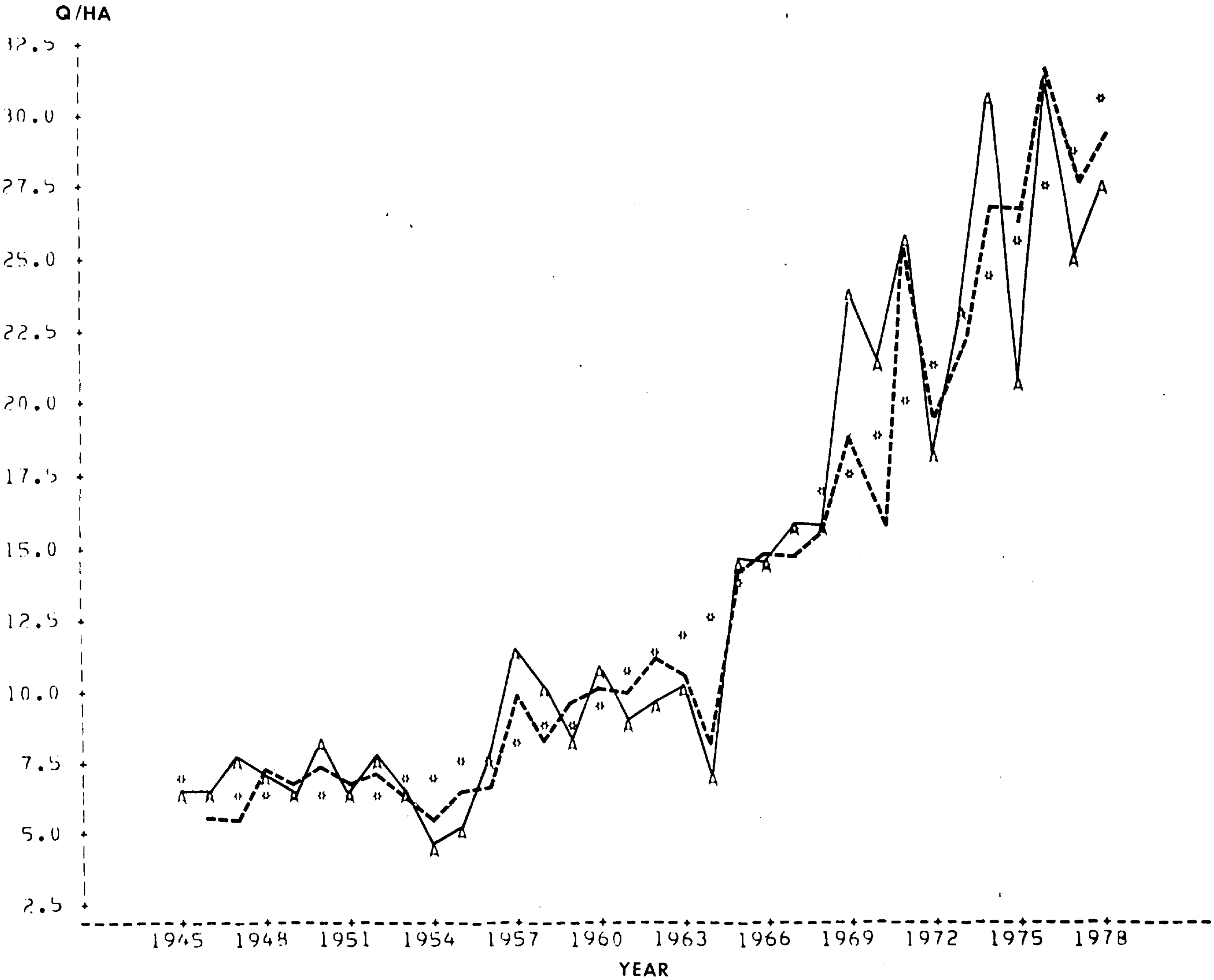

Figure 20. Actual and predicted barley yield for Belorussia economic region, USSR, 1945-1978. (The legend is on Figure 19.)

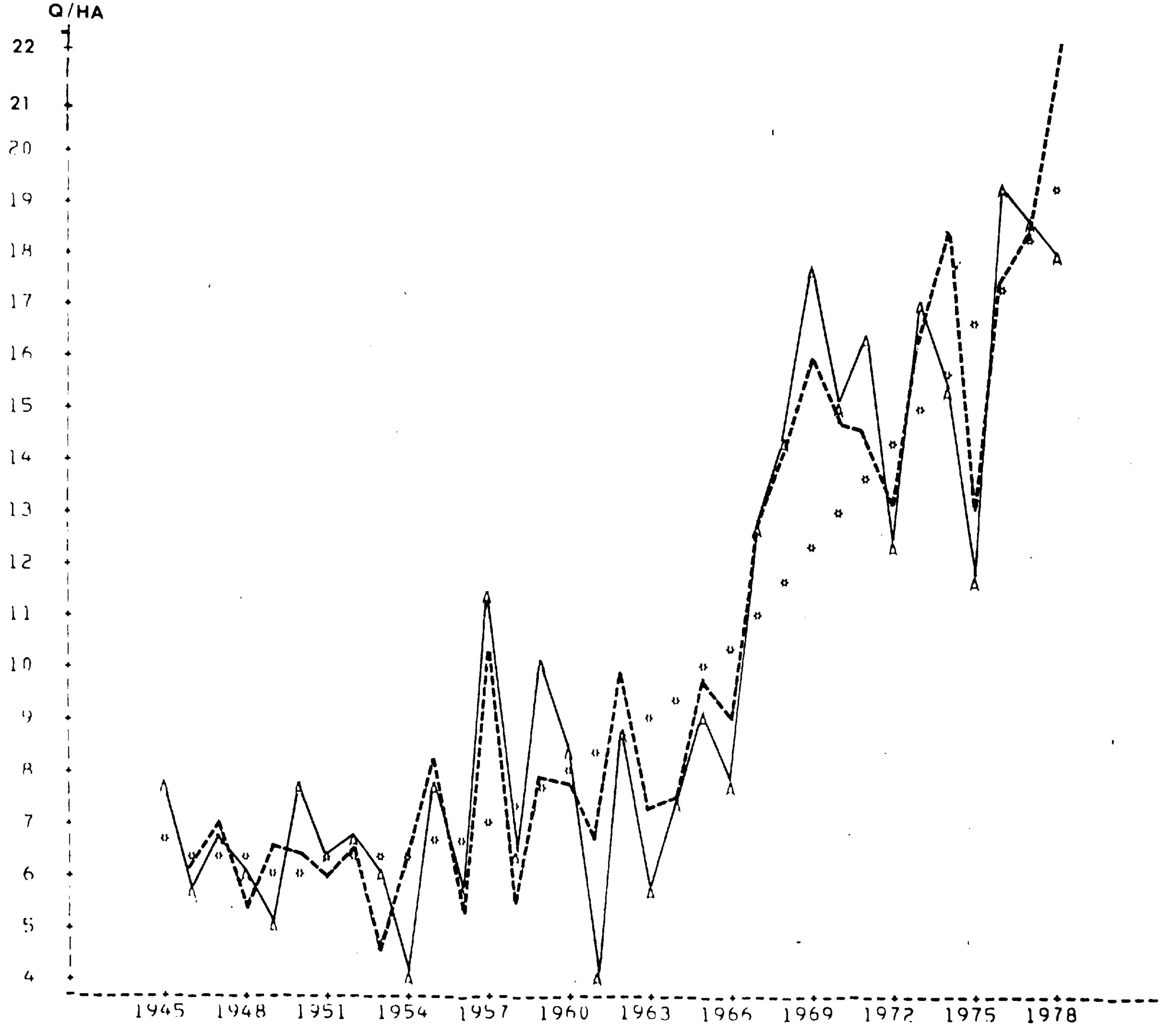

Figure 21. Actual and predicted barley yield for Volgo-Vyatka economic region, USSR, 1945-1978. (Legend is on Figure 19.)

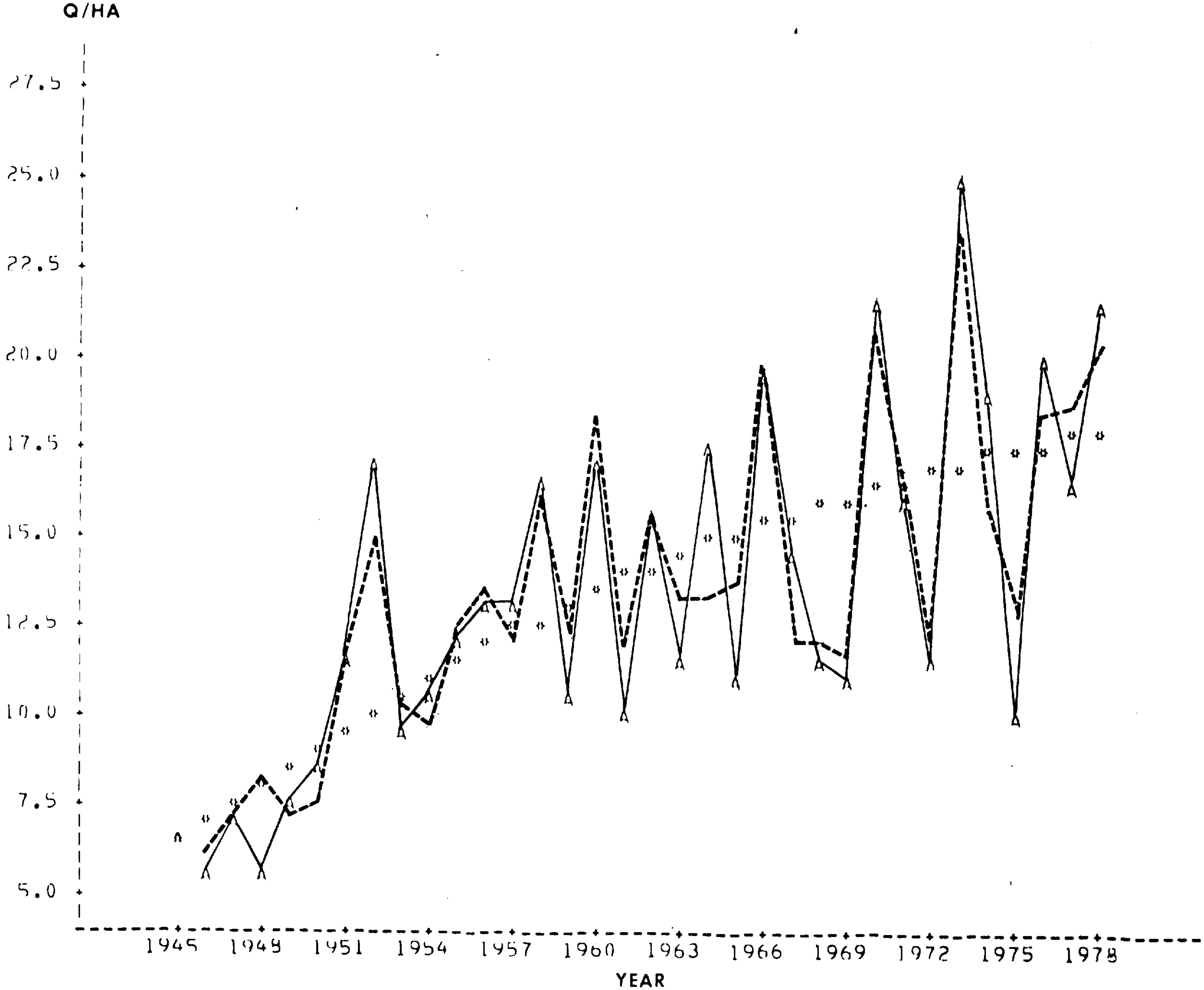

Figure 22. Actual and predicted barley yield for North Caucasus economic region, USSR, 1945-1978. (Legend is on Figure 19.)

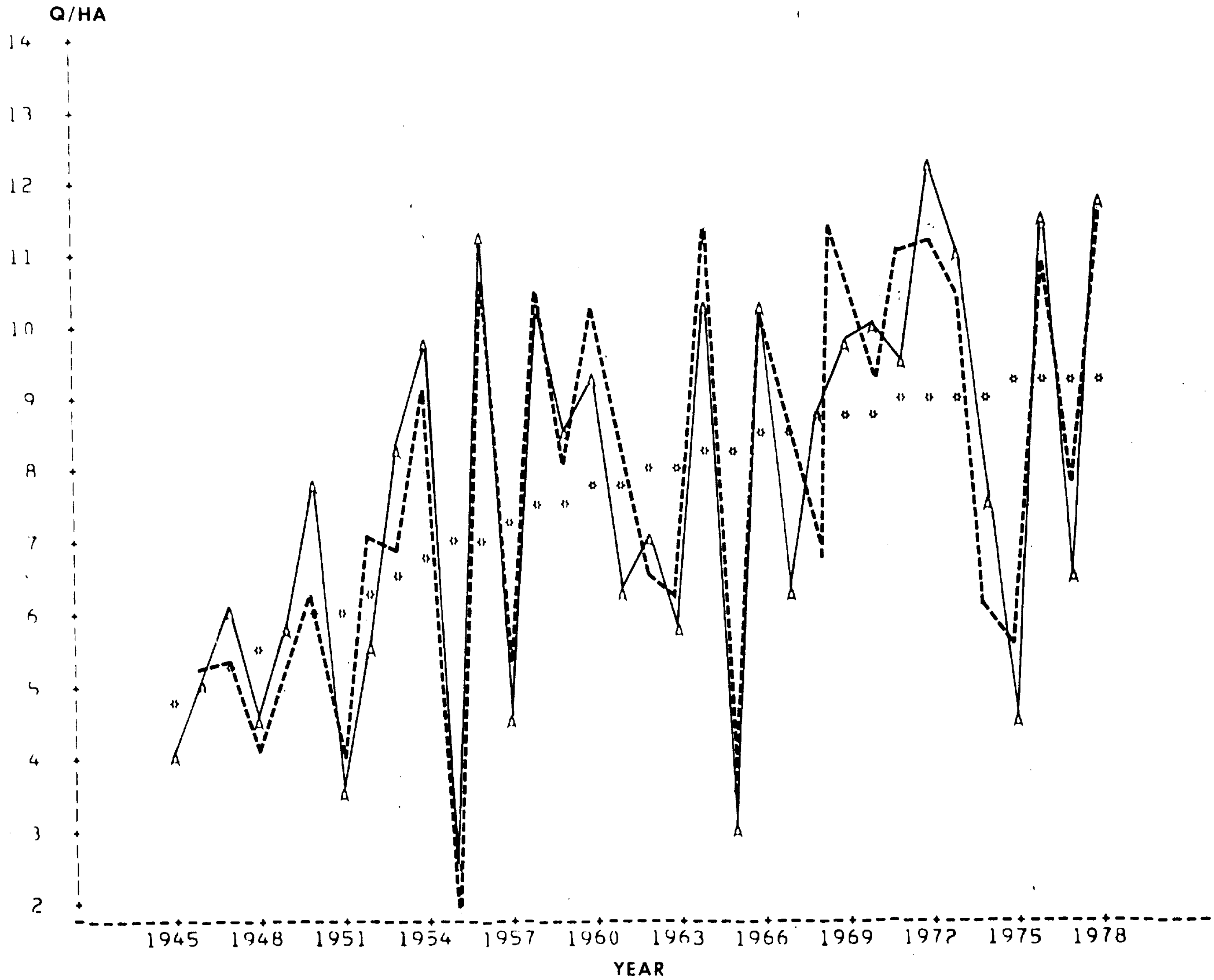

Figure 23. Actual and predicted barley yield for Kazakh economic region, USSR, 1945-1978. (Legend is on Figure 19.)

the following year. The modification was that the yield departures from trend were computed using the 1945-1978 trend base. The period of 1971-1978 was used for the independent tests in this study. Thus, a weather-crop model was developed for departures of yield from the 1945-78 trend, using P and T data for the 1945-70 period, and applied to 1971. Then the model was fitted to data for 1945-71 and applied for 1972, and subsequent models were developed and applied in this manner until a model based on 1945-77 data was applied for 1978.

Six models were independently tested. These models included two sets of index-variables. Two of the models use three independent index variables, P, P^2 and T^2, developed by weighting months proportional to the Pearson correlation coefficient and its square (r and r^2). In Table 17 these models are numbers 2 and 3, respectively. Two other models use the independent index variables PT and $(PT)^2$ double-weighted by means of the Pearson correlation coefficient. The first weighting was made by weighting the index-precipitation and index-temperature in proportion first to r and then to r^2. These models are designated in Table 17 as model numbers 4 and 5, respectively. In the South-West, Volgo-Vyatka and North Caucasus economic regions, one additional model was included for testing. The additional model contained P, T^2 variables for South-West region (model number 1a, Table 17) and P, T variables for Volgo-Vyatka and North Caucasus regions (model number 1b).

To estimate the performance of models with different sets of variables in independent tests, some indicators of yield reliability were generated based on the approach developed in [27, 28, 36]. In this approach, all indicators are based on the comparison of two components: predicted yield ($\hat{Y}$) using a bootstrap technique and actual yield (Y). The difference between these two yields is expressed in absolute terms (Q/HA) as $d_i = \hat{Y}_i - Y_i$ and in relative terms (%) as $Rd_i = 100\ d_i/(Y_i)$, where i = 1, ..., 8 indicates the year.

Based upon d_i and Rd_i the following measures of model performance are computed:

mean square error (MSE = $1/8\sum_{i=1}^{8} d_i^2$),

the root mean square error (RMSE = $(\mathrm{MSE})^{1/2}$),

the relative root mean square error (RRMSE = $((\mathrm{RMSE}/\bar{Y})*100)$,

the variance (Var = $1/8\sum_{i=1}^{8}(d_i-\bar{d})^2$),

the standard deviation (SD = $(\mathrm{Var})^{1/2}$),

the relative standard deviation (RSD = $((\mathrm{SD})/\bar{Y}+\bar{d}))*100$, bias $(B=1/8\sum_{i=1}^{8} d_i)$,

and the relative bias RB = $(B/\bar{Y})100$.

Here $\bar{d}$ = B and $\bar{Y} = 1/8\sum_{i=1}^{8} Y_i$. The Pearson correlation coefficient between sets of actual and predicted yields is also computed. All these parameters (except the variance) are presented in Table 17.

Table 17. Indicators of yield reliability for developed models in selected economic regions, USSR, in tested period 1971-1978

		Model includes trend & weather variables		Model assessment for weather variables		Independent Test								
						Error			Standard Deviation		Bias			
Economic Region	Model Number	Index Variable	Type of Weighting	R	Condition Index for Weather	MSE $(Q/HA)^2$	RMSE (Q/HA)	RRMSE (%)	SD (Q/HA)	RSD (%)	Bias (Q/HA)	RB (%)	Correlation Coefficient for Y and Y
South West	1a	P, T^2	R_1	0.746	10.2	27.8	5.28	19.9	4.85	19.8	-2.08	-7.8	0.27
$\overline{Y}$=26.6 Q/HA	2	P, P^2, T^2	R_1	0.747	24.5	23.0	4.80	18.1	4.57	18.2	-1.48	-5.6	0.36
	3	P, P^2, T^2	R_1^2	0.784	20.3	22.0	4.69	17.6	4.13	15.8	-2.00	-7.5	0.48
	4	$PT, (PT)^2$	R_1, R_2	0.743	1.8	12.8	3.58	13.5	3.18	11.3	1.65	6.2	0.89
	5	$PT, (PT)^2$	R_1, R_2^2	0.787	1.9	3.2	1.80	6.8	1.76	6.5	0.39	1.5	0.95
Belorussia	2	P, P^2, T^2	R_1	0.673	11.7	12.7	3.56	14.1	2.86	10.7	1.36	5.4	0.66
$\overline{Y}$=25.3 Q/HA	3	P, P^2, T^2	R_1^2	0.650	23.9	11.0	3.32	13.1	3.04	11.4	1.35	5.3	0.70
	4	$PT, (PT)^2$	R_1, R_2	0.658	1.9	12.1	3.47	13.7	3.30	12.5	1.08	4.3	0.69
	5	$PT, (PT)^2$	R_1, R_2^2	0.617	2.0	13.0	3.61	14.3	3.19	11.9	1.54	6.1	0.72
Volgo-Vyatka	1b	P, T	R_1	0.680	11.9	9.7	3.11	19.4	3.10	19.7	-0.30	-1.9	0.61
$\overline{Y}$=16.0 Q/HA	2	P, P^2, T^2	R_1	0.645	28.8	15.5	3.94	24.6	3.49	20.3	1.62	11.0	0.06
	3	P, P^2, T^2	R_1^2	0.667	38.1	14.6	3.82	23.9	3.39	19.1	1.76	11.0	0.25
	4	$PT, (PT)^2$	R_1, R_2	0.696	2.3	10.4	3.23	20.2	3.05	19.1	1.05	6.6	0.53
	5	$PT, (PT)^2$	R_1, R_2^2	0.668	2.3	14.4	3.79	23.7	3.59	22.4	1.38	8.6	0.52
North Caucasus	1b	P, T	R_1	0.852	64.0	4.3	2.08	11.9	2.05	11.8	-0.15	-0.9	0.92
$\overline{Y}$=17.5 Q/HA	2	P, P^2, T^2	R_1	0.809	24.9	11.6	3.40	19.4	4.05	24.1	-0.09	-0.5	0.53
	3	P, P^2, T^2	R_1^2	0.771	18.5	21.6	4.65	26.4	4.64	26.8	-0.24	-1.4	0.17
	4	$PT, (PT)^2$	R_1, R_2	0.865	2.5	5.7	2.39	13.6	2.38	13.8	-0.21	-1.2	0.90
	5	$PT, (PT)^2$	R_1^2, R_2^2	0.880	2.4	4.8	2.18	12.4	2.11	12.5	-0.56	-3.2	0.80
Kazakh	2	P, P^2, T^2	R_1	0.887	11.2	1.3	1.15	12.4	1.15	12.5	-0.09	-1.0	0.90
$\overline{Y}$=9.3 Q/HA	3	P, P^2, T^2	R_2^2	0.896	19.5	1.1	1.07	11.5	1.07	11.5	0.04	0.4	0.92
	4	$PT, (PT)^2$	R_1, R_2	0.883	2.2	1.5	1.24	13.4	1.24	13.4	0	0.0	0.89
	5	$PT, (PT)^2$	R_1, R_2^2	0.897	2.2	1.1	1.07	11.5	1.07	11.6	-0.02	-0.2	0.92

In addition, paired-sample test statistics (parametric test based on the Student's t-test and nonparametric test based on the Wilcoxon Signed Rank test) were also used to identify the best performing model. This statistic was applied to the values of difference between absolute differences of actual and predicted yields for compared models j and k $(D_{(jk)i}=|d_{ji}|-|d_{ki}|)$ [27, 28, 34].

It should be understood that the modeling technique employed in this study requires that two steps be completed in order to predict a yield. The first step consists of the prediction of the departure of yield from trend based on weather. The second one considers extrapolation of trend. It would, of course, be desirable to conduct independent evaluations for both of these yield components. Since there are four to five models to be evaluated for each region, many calculations would be required to evaluate both components for each model independently. Therefore, the independent test was conducted only for the yield departure from trend weather component which explains the largest fraction of yield variability from year to year. To find out to what extent this approach could misrepresent the predicted yield, a special test for independent evaluation of both the trend and departure from trend components was conducted for the best performing model in Kazakh economic region.

Results of Independent Testing of Weather Component

The paired-sample test statistics (both parametric and nonparametric) show that none of the average differences between values of model performance could be considered statistically significant. Even so, for some regions there are fairly substantial differences in the values of the parameters of yield reliability which may indicate that one model is preferable to another one. Thus, in South-West, Belorussia, and Kazakh economic regions, models with a low condition index for weather have the highest correlation coefficient between Y and $\hat{Y}$. These models (No. 5 in each region) are among those with the lowest errors and standard deviations. The largest relative bias for these models is only 6.1 percent, and for two of the regions the bias is very small (-0.2 and 1.5 percent). These models also have the highest probability of the absolute value of the relative difference $(|Rd_i|)$ being less than 15 percent (except in Kazakh). A difference of less than 15 percent occurred in 62-88 percent of all years (Table 18). The percentage of coincidence in the direction of actual and predicted yields from year to year is also high and (except for South-West) exceeds this percentage for other models (Table 18). The results of the independent tests for Volgo-Vyatka and North Caucasus differ from the other regions but are not as clear in identifying a preferred model. In these cases, preference could be given to the models with the smallest condition index.

In general, models based on double-weighted PT index-variables showed the same or better results in independent tests than models based on single-weighted P and T index-variables. Also, for any particular region, all tested models had high correlation coefficients for equations relating the ratio of yield departures from trend with weather in the independent test (Table 17) and for equations relating calculated and actual yield in dependent tests (r=0.92-0.98). Even though models with the smallest condition index had very good combinations

Table 18. Model comparison based on the percentage of years with the relative difference (Rd) less than 10, 15, 20 percent and strength of agreement in the direction of yield change from previous year to the next for actual and predicted yields in selected economic regions, USSR

Economic Region	Model number	Model includes trend and weather variables		Percent of years with |Rd| less than			Percent of agreement in the direction of Y and $\hat{Y}$ change from year to year
		Index variable	Type of weighting	10	15	20	
South-West	1a	P, T^2	R_1	50	62	62	43
	2	P, P^2, T	R_1	38	62	62	43
	3	P, P^2, T^2	R_1^2	38	62	62	28
	4	$PT, (PT)^2$	R_1, R_2	62	75	88	100
	5	$PT, (PT)^2$	R_1, R_2^2	88	88	100	86
Belorussia	2	P, P^2, T^2	R_1	50	50	88	72
	3	P, P^2, T^2	R_1^2	62	62	88	100
	4	$PT, (PT)^2$	R_1, R_2	62	88	88	86
	5	$PT, (PT)^2$	R_1, R_2^2	62	88	88	100
Volgo-Vyatka	1b	P, T	R_1	50	62	62	57
	2	P, P^2, T^2	R_1	25	50	50	14
	3	P, P^2, T^2	R_1^2	38	38	50	43
	4	$PT, (PT)^2$	R_1, R_2	38	62	62	71
	5	$PT, (PT)^2$	R_1, R_2^2	50	62	62	57
North Caucasus	1b	P, T	R_1	50	75	88	86
	2	P, P^2, T^2	R_1	25	38	75	43
	3	P, P^2, T^2	R_1^2	25	25	50	86
	4	$PT, (PT)^2$	R_1, R_2	50	62	88	86
	5	$PT, (PT)^2$	R_1^2, R_2^2	62	75	88	100
Kazakh	2	P, P^2, T^2	R_1	50	75	75	86
	3	P, P^2, T^2	R_2^2	62	75	75	86
	4	$PT, (PT)^2$	R_1, R_2	50	62	88	86
	5	$PT, (PT)^2$	R_1, R_2^2	62	62	88	100

of the various measures of yield reliability, these models in regions with a good water supply (South-West, Belorussia and Volgo-Vyatka) performed much better when the year with the largest difference between actual and predicted yields was excluded. The elimination of only this one test year substantially decreased the bias in these three regions. Upon elimination of the worst test year, the correlation coefficients between $\hat{Y}$ and Y for some models increased, especially in the Belorussia economic region. In regions which tend to have a water deficit (North Caucasus and Kazakh), the elimination of the test year with the largest error caused a slight increase in bias, but at the same time, correlation coefficients did not change. Thus, in most cases, independently tested models showed good results. One should, however, bear in mind that in some years and regions the error of yield prediction can be as great as 50-60 percent of the average barley yield for a particular region (Table 19).

Once again emphasis should be placed upon the fact that the models with the value of condition index lower than the limit defining the existence of collinearity can be considered as the best performing. The estimates for the coefficients of these models' equations are shown in Table 20.

All models showed very high values of R^2, 0.902-0.953. Of course, this R^2 is for the model base period and is higher, sometimes considerably, than the corresponding R^2 of the independent test for the shorter period (Table 21). Considering separate assessments for technology and weather, substantial differences between the regions can be identified. In the dry area (North Caucasus and Kazakh), the relationship between yield and technology weakens (correlation coefficients decrease) as the yield-weather relationship strengthens (correlation coefficients increase). But in the areas with adequate water, the opposite result can be observed.

The assessments of yield reliability for the best performing models in the studied regions are shown in Table 21. Only in Belorussia did the average difference between predicted and actual barley yield amount to as much as 1.5 Q/HA. In the rest of the regions the absolute values of this difference were very low and did not exceed 0.4 Q/HA. But there are noticeable differences between regions on the other indices. Because a less variable yield is easier to predict than a more variable one, an additional measure is proposed to estimate relative standard deviation. In addition to expressing the standard deviation as relative to the sum of two parameters $\bar{Y}$ and $\bar{d}$, it is also expressed with respect to the sum of three parameters -- $\bar{Y}$, $\bar{d}$ and $\bar{V}$. Here $\bar{V}$ defines average level of year to year yield variability in a region. The expression for $\bar{V}$ is presented in Table 21. The relative standard deviation (SD) in prediction based on $\bar{Y}$, $\bar{d}$ and $\bar{V}$ provides more information about model performance in different regions. Application of SD relative to $\bar{Y} + \bar{d}$ showed little difference between models in Belorussia, North Caucasus and Kazakh. But, in the case of the standard deviation relative to $\bar{Y} + \bar{d} + \bar{V}$, we can see some important differences. However, these differences are not as large as the differences between RMSE and SD. Judging by this type of relative SD, the model with poorest performance is the one for the Volgo-Vyatka economic region, with the value of 15.7. The succession of regions corresponding to the decrease of this value and respective improvement of model performance following Volgo-Vyatka is Belorussia, North Caucasus, Kazakh and the South-West. This order is in agreement with the order of regions corresponding to the increase in the Pearson correlation coefficient between actual and predicted barley yield.

Table 19. Effect of the elimination of one of the observations with the largest error on the value of some indices of yield reliability in selected economic regions, USSR

Economic Region	Model Number	Largest Absolute Relative Difference (RD_i) (%)	Bias (Q/HA)		Pearson Correlation Coefficient	
			All Observations	Without Observation with the Largest Error	All Observations	Without Observation with the Largest Error
South-West	1a	-32.9	-2.08	-1.56	0.27	0.50
	2	-27.2	-1.48	-0.33	0.36	0.43
	3	-26.1	-2.00	-0.98	0.48	0.47
	4	25.2	1.65	0.63	0.89	0.89
	5	13.6	0.39	0.03	0.95	0.98
Belorussia	2	29.1	1.36	0.56	0.66	0.72
	3	29.1	1.35	0.68	0.70	0.78
	4	38.4	1.08	0.11	0.69	0.85
	5	39.3	1.54	0.60	0.72	0.87
Volgo-Vyatka	1b	38.0	1.40	0.63	0.61	0.69
	2	57.4	1.62	0.91	0.06	-0.33
	3	57.4	1.76	1.08	0.25	0.35
	4	38.0	1.05	0.23	0.53	0.58
	5	48.6	1.38	0.33	0.52	0.58
North Caucasus	1b	30.3	-0.15	-0.33	0.92	0.90
	2	62.7	-0.09	-1.00	0.53	0.54
	3	76.4	-0.24	-0.36	0.17	0.36
	4	37.2	-0.21	-0.78	0.90	0.89
	5	24.5	-0.56	-1.00	0.80	0.90
Kazakh	2	40.9	-0.09	-0.36	0.90	0.89
	3	43.2	0.04	-0.23	0.92	0.93
	4	32.8	0.00	-0.30	0.89	0.92
	5	36.4	-0.02	-0.26	0.92	0.90

Table 20. Estimates for the coefficients and assessments for the large area barley-yield models in selected economic regions, USSR 1945-1978

Economic	Y_t Estimates Based on Technology Trend			Y/Y_t Estimates Based on the Weather Index-variables					Type of	R^2		
Region (model)	Intercept	YR-1944	$(YR-1944)^2$	Intercept	IP	IT	IPT	$(IPT)^2$	Weighting	Model	Technology (trend only)	Weather (de-trended yield)
South West (mod 5)	6.460	0.566	0.0033	100.878			15.780	-2.121	R_1,R_2^2	.953	.850	.619
Belorussia (mod 5)	6.994	-0.302	0.0293	99.159			16.324	1.471	R_1,R_2	.943	.880	.434
Volgo-Vyatka (mod 1b)	6.694	-0.188	0.0164	135.929	0.857	18.069			R_1	.902	.762	.463
North Caucasus (mod 1b)	5.912	0.576	-0.0065	74.626	1.456	14.031			R_1	.910	.496	.726
Kazakh (mod 5)	4.546	0.247	-0.0032	104.105			30.444	-4.453	R_1,R_2^2	.902	.244	.804

Table 21. Assessment of yield reliability for the best performing models in selected economic regions, USSR

Economic Region (model)	Error of Prediction (Q/HA)			R^2	Relative SD based on	
	Bias	RMSE	SD		$\overline{Y}+\overline{d}$ (%)	$\overline{Y}+\overline{d}+\overline{V}$* (%)
South-West (Model 5)	0.39	1.80	1.76	.902	6.5	5.0
Belorussia (Model 5)	1.54	3.61	3.19	.518	11.9	9.5
Volgo-Vyatka (Model 1b)	-0.30	3.11	3.10	.372	19.7	15.7
North Caucasus (Model 1b)	-0.15	2.08	2.05	.846	11.8	8.2
Kazakh (Model 5)	-0.02	1.07	1.07	.846	11.6	7.8

$^{*}\overline{V}$ defines an index of year to year yield variability. It is expressed in the form: $\overline{V} = \sum_{i=1}^{N-1}(Y_{i+1}-Y_i)^2/(N-1)$ where Y is yield and N is number of years with observed yields, i=1,...,N.

Results of Independent Testing of Both Trend and Weather Components for Kazakh Economic Region

As was mentioned above, an independent evaluation of both trend and departure from trend was conducted for the Kazakh economic region. Based on the criteria of the extraction of a basic time period for trend development, as described in the section on crop yield modeling and the available data base, two basic periods were identified for the Kazakh economic region: 1945-1970 and 1945-1976. This attempts to independently simulate the decision that would have been made for each of the test years. That is, to conclude either (1) that recent year yields are approximately equally and uniformly distributed on both sides of the projected previous trend and these years should be included in specifying the new trend, or (2) that recent years are not equally or uniformly distributed with respect to the existing trend and, therefore, should not be included. Alternative (1) was selected for test years 1971 and 1977, and alternative (2) for 1972 through 1976 and for 1978 (see Figure 17, page 35). Thus, two trend equations were developed (Table 22). These equations were applied to extrapolate independently barley yield in the period of 1971-1978 (the first equation was applied up through 1976; the second one after 1976). After that, the departure of actual yield from independently fitted and extrapolated yields ($Y/\hat{Y}_t$) for the applicable period was calculated and the ratio yield departure from trend-weather model was developed. Using the bootstrap technique, this model was tested and evaluated for the period 1971-1978. One alternative period was also selected for development of the trend model (1945-1969 in place of 1945-1970) to estimate the effect of this minor change on the model results. The 1945-69 trend model had an intercept of 4.52 and slope coefficients of 0.250 and -0.0041, respectively.

Comparison of the regression coefficients for independent variables of the models developed for different periods can characterize model performance. As seen in this table the differences in regression coefficients for the trend models for the periods of 1945-1978, 1945-1970 and 1945-1976 are not great. However, the differences become somewhat larger if the trend model for 1945-1969 is used in place of the one for the 1945-70 period. A similar conclusion can be obtained by comparing the regression coefficients for the weather-yield departure models. Smaller differences in the weather model coefficients were found when the 1945-70 period was used to estimate trend and establish the departures from trend than when the 1945-69 period was used.

Table 22. Estimates for regression coefficients in Kazakh economic region, USSR

Period[1/] used for model development	Trend			Weather		
	Int.	YR-1944	$(YR-1944)^2$	Int.	PT	$(PT)^2$
1945-1970a	4.51	0.261	-0.0041	104.8	30.3	-4.2
1945-1971a	4.51	0.261	-0.0041	104.0	29.8	-4.1
1945-1972a	4.51	0.261	-0.0041	104.6	30.4	-4.2
1945-1973a	4.51	0.261	-0.0041	105.4	30.5	-4.5
1945-1974a	4.51	0.261	-0.0041	105.8	29.8	-4.4
1945-1975a	4.51	0.261	-0.0041	105.4	30.6	-4.6
1945-1976b	4.52	0.250	-0.0033	104.2	30.6	-4.5
1945-1977b	4.52	0.250	-0.0033	104.6	30.7	-4.6
1945-1978c	4.54	0.247	-0.0032	104.1	30.4	-4.4

[1/] Periods used for estimating the trend relationship are 1945-70(a), 1945-76(b) and 1945-78(c).

The weighting coefficients for weather variables also show only small differences between models developed by considering trend for different periods. Normally, the difference between coefficients exists only in the second or even the third decimal place. A difference in the first decimal figure can be noticed for some months when the 1945-69 period is used to establish trend and departures from it. But even in that case the months with the most substantial influence of weather on yield show the same values of this influence for all trend specification periods. The results of the fully independent test of models for the considered trend specification periods are shown in Table 23. Comparison of these results with the results previously obtained (for the period 1945-1978) shows that the differences in the values of the predicted yield components (trend and ratio departure from trend) and final yield are not very large. At the same time, it should be pointed out that the correct choice of the period for the development of trend model is a very important step in analogue weather crop modeling.

Concerning other evaluation criteria, we should mention that all obtained regional models are not costly to operate or to develop. They are consistent with current knowledge concerning weather-crop-technology relationships for large areas. The lead time of a yield estimate, based on actual weather data, is 2-3 months ahead of the barley harvest. These models are easy to understand. The models were developed to facilitate their redevelopment as soon as new weather and crop data are available.

Table 23

Fully Independent Test of Barley Yield Models Using Two Alternative Trend Specification Methods As Compared to the Weather Model Independent Test in the Kazakh Economic Region, USSR

Method		Weather Independent Test			Fully Independent Tests					
Trend Period		1945-78			1945-70, 1945-76			1945-69, 1945-76		
Application Period		1971-78			1971-76, 1977-78			1971-76, 1977-78		
Year	Reported Yield (Q/HA)	Trend Yield (Q/HA)	Departure of Yield from Trend (%)	Yield Estimate (Q/HA	Trend Yield (Q/HA	Departure of Yield from Trend (%)	Yield Estimate (Q/HA)	Trend Yield (Q/HA)	Departure of Yield from Trend (%)	Yield Estimate (Q/HA)
1971	9.5	8.9	124	11.0	8.6	126	10.8	7.8	132	10.3
1972	12.3	9.0	125	11.2	8.6	127	10.9	7.7	131	10.1
1973	10.9	9.0	112	10.1	8.6	114	9.8	7.6	118	9.0
1974	7.4	9.1	68	6.2	8.7	70	6.1	7.5	76	5.7
1975	4.4	9.1	66	6.0	8.7	69	6.0	7.4	78	5.8
1976	11.6	9.2	116	10.7	8.7	119	10.4	7.3	127	9.3
1977	6.4	9.2	73	6.7	9.2	72	6.6	9.2	72	6.6
1978	11.7	9.2	131	12.1	9.2	131	12.1	9.2	131	12.1
Bias ($\bar{d}$)				- 0.025			- 0.17			- 0.66
MSE				1.15			1.34			2.43

Summary of Preferred Model Development and Use

The algorithm explaining the succession of steps in yield calculation based on trend extrapolation and assessment of weather influence on the deviation of yield from trend is shown in Figure 24.

To calculate average precipitation and temperature for regions the actual monthly sum of precipitation and monthly average temperature at the weather stations shown in Table 24 are required. Simple averages are used. Average monthly weather data for regions must cover the period from January of the year preceding the year of crop harvest until July of the crop harvest year. Based on the weighting coefficients presented in Table 10 and applying equation (5), the index-weather variables (IP, IT) are calculated. These index variables are used along with the corresponding estimates for coefficients (presented in Table 20) to calculate the barley yield departure from trend for those regions where a single-weighting procedure is used (Volgo-Vyatka, North Caucasus). For regions where a double-weighting procedure is required (South-West, Belorussia, Kazakh), a new index-variable, IPT, must be calculated. For this purpose, the previously calculated IP and IT variables and weights (from Table 15) which correspond to the variables and the type of weighting in Table 20 are used. The calculation of IPT is carried out based on equation (5). And again, the estimated barley yield ratio departure from trend is calculated based on corresponding IPT and $(IPT)^2$ variables and estimates for coefficients in Table 20. Extrapolation of barley-yield trend for each region is based on the equations of trend (Table 20) and the empirical methodology presented in the section on crop yield modeling and the available data base. Using the assessments of yield based on trend extrapolation and deviation from trend due to weather fluctuation, the forecasted barley yield is calculated as the product of these two estimates.

The flow chart reviewing the method presented in this study for weather-crop-technology model development is shown in Figure 25. The process of modeling starts with an empirical procedure of examining the yield series and identifying the appropriate period for trend model development. The next step consists of developing the trend model and calculating the weather induced variation of yields around the trend in form of $Y/\hat{Y}_t$. Preparing the weather information requires averaging total monthly precipitation and average monthly air temperature at separate weather stations in an economic region. The index-weather variables are calculated based on the values of physiological response of crop to weather variation. These values constitute the weights of every month over the nineteen month period preceding harvest. The second weighting is required to aggregate different weighted-over-time weather variables (IP and IT) into one IPT weather variable. Then, the regression coefficients are estimated for the equation which expresses the dependence of the yield departures from trend $(Y/\hat{Y}_t)$ on various index-weather variables.

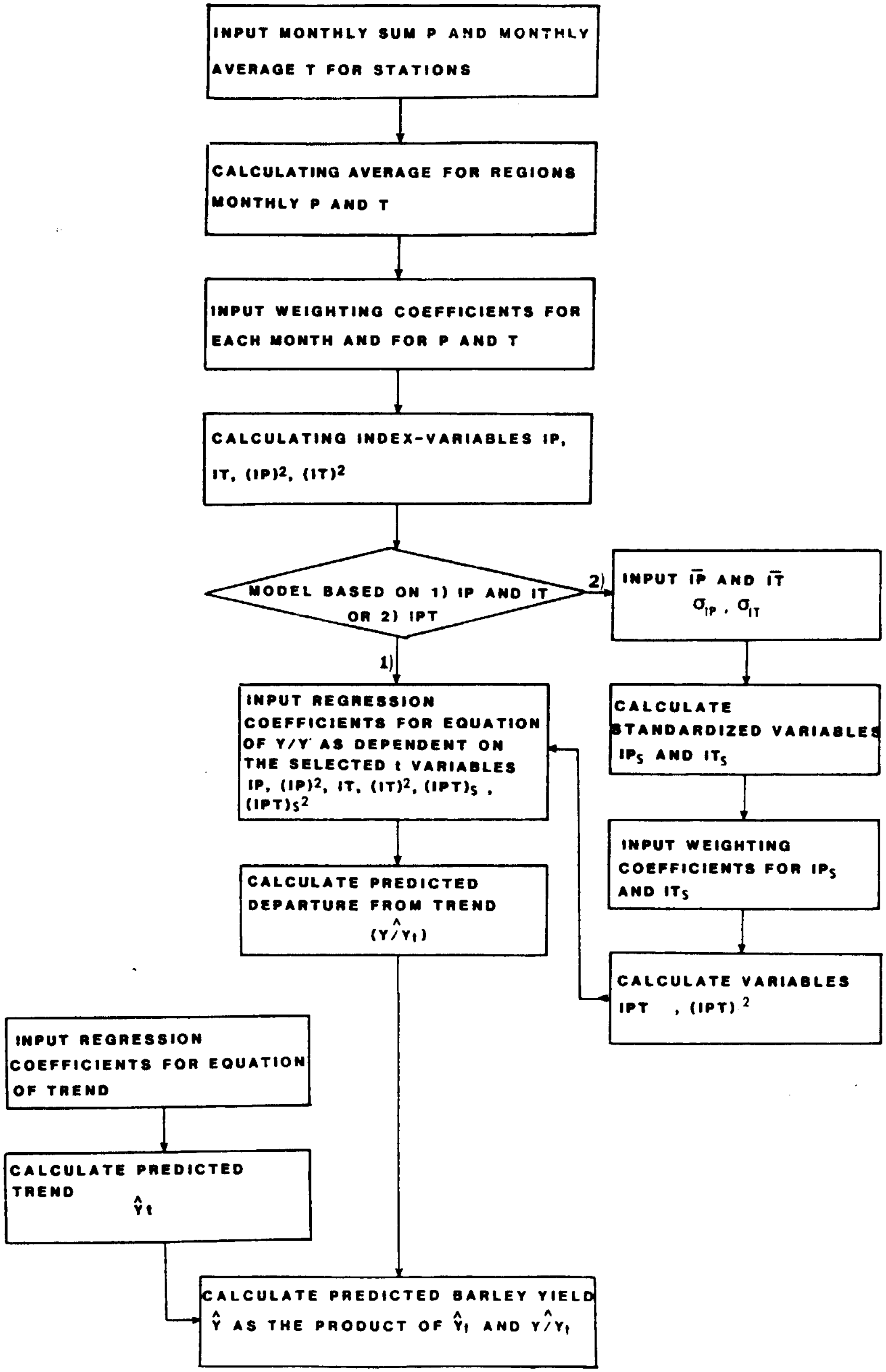

Figure 24. Flow chart explaining the calculation of barley yield based on developed model.

Table 24. List of meteorological stations in selected economic regions, USSR used for calculation of the average regional weather

Economic Region	WMO[1] Number	Name of station	Latitude	Longitude	Altitude
South-West	33088	Sarny	51 21N	26 37E	153
	135	Chernigov	51 29	31 17	137
	345	Kiev	50 24	30 27	179
	393	L'vov	49 49	23 27	325
	562	Vinnica	49 14	28 28	287
Belorussia	26666	Vitebsk	55 10	30 08	151
	825	Grodno	53 41	23 50	152
	850	Minsk	53 52	27 32	234
	863	Mogilev	53 54	30 19	180
	33041	Gomel	52 27	31 00	139
Volgo-Vyatka	27196	Kirov	58 39	49 37	164
	485	Joskar-Ola	56 38	47 50	103
	553	Gor'kij	56 13	43 49	82
	557	Sercach	55 31	45 30	93
	581	Cheboksary	56 09	47 17	183
North Caucasus	34731	Rostov-Na-Donu	47 15	39 49	77
	748	Proletarsk	46 42	41 44	25
	838	Tihoretsk	45 51	40 05	79
	929	Krasnodar	45 02	39 09	33
	949	Stavropol	45 03	42 01	467
	964	Arzgir	45 24	44 12	75
Kazakh	28879	Kokchetav	53 17	69 21	229
	952	Kustanaj	53 13	63 37	171
	966	Ruzaevka	52 49	66 58	227
	29807	Irtyssk	53 21	75 27	94
	35166	Kazgorodok	51 16	67 14	252
	188	Tscelinograd	51 08	71 22	348
	229	Aktjubinsk	50 17	57 09	227
	302	Saraevo	50 12	51 10	16
	358	Turgaj	49 38	63 30	123
	379	Karaganda	49 48	73 08	555
	36003	Pavlodar	52 17	76 57	146
	177	Semipalatinsk	50 21	80 15	206

---

[1] WMO is the World Meteorological Organization.

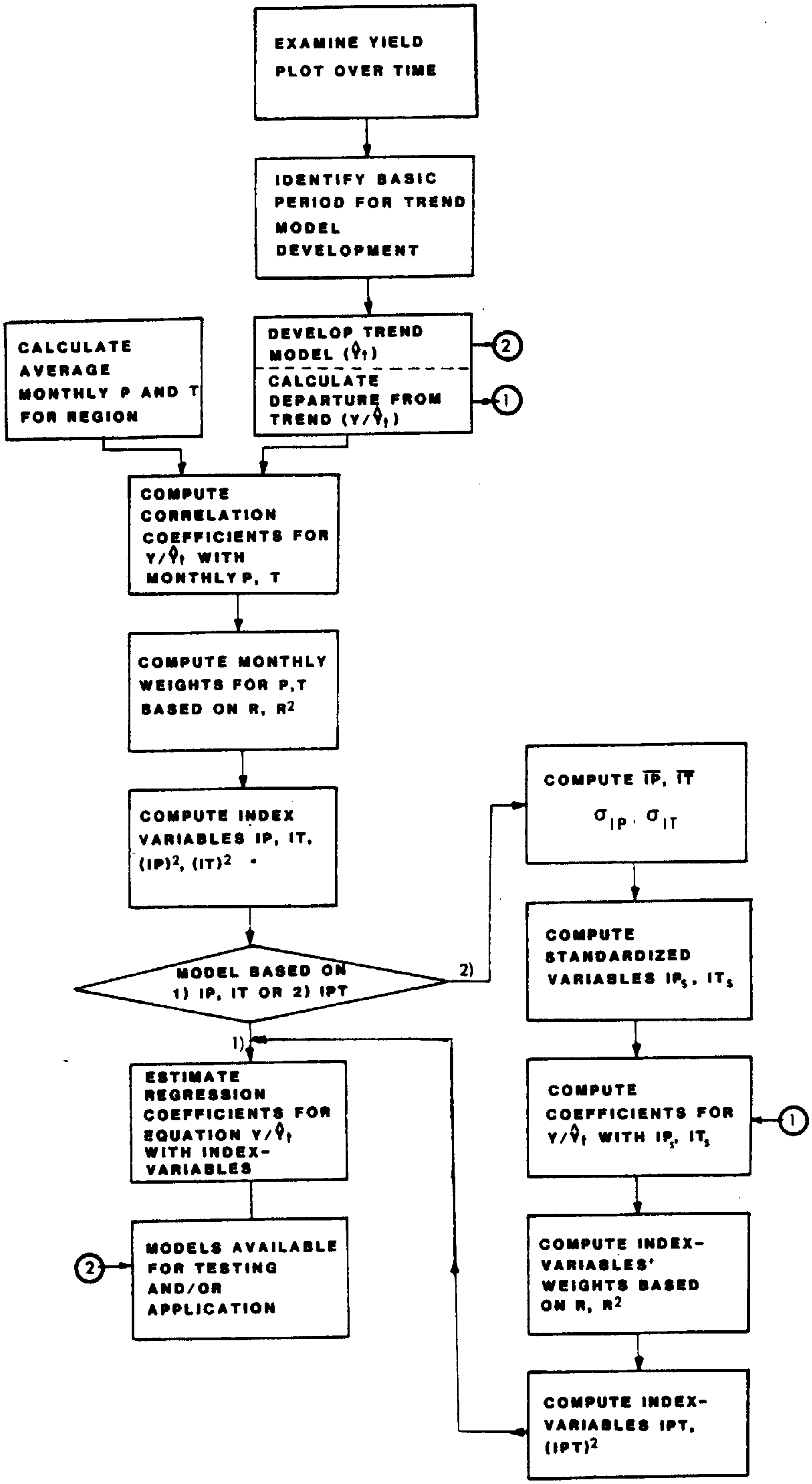

Figure 25. Flow chart explaining the development of crop yield model.

REFERENCES

1. Abashina, E. V., Prosvirkina, A. G., Sirotenko, O. D., "Simplified dynamic model for spring barley formation," Trudy IEM, No. 8 (67), 1976, pp. 54-67.

2. Agroclimatic Atlas of the World [in Russian] Edit. Gol'tsberg I.A., Gidrometeoizdat, Moscow-Leningrad, 1972.

3. A Study of the Caribbean Basin Drought/Food Production Problem: Final Report, Climatic Impact Assessment Division-Models Branch of NOAA/EDIS and University of Missouri-Columbia. Atmospheric Science Department, May, 1979, 200 pp.

4. Belsley, D. A., E. Kuh, R. E. Welsch, Regression diagnostics. New York: John Wiley and Sons, 1980.

5. Brounov, P. I., "The effect of the sun spots and meteorological factors on grain crop productivity," Speech at the Meeting of the Imperatorskoye Vol'noye Economicheskoye Obschestvo on October 31, 1898, St. Petersburg, 1899.

6. Dmitrenko, V. P., "The mathematical model for assessment of agricultural crops yield," Trudy UKRNIGMI, no. 122, 1973, pp. 3-13.

7. ________________, "On the methodology for estimating hydrometeorological conditions of the productivity of agricultural crops." Trudy UKRNIGMI, no. 128, 1973, pp. 3-23.

8. Dugas, W. A., B. S. Jackson, G. F. Arkin, "Evaluating relationships between data network distribution and simulated sorghum yields." 15th Conference on Agriculture and Forest Meteorology and 5th Conference on Biometeorology. AMS, Anaheim, California, April 1981, pp. 186-189.

9. Gladishev, A. N. and others. The Non-Chernozem Zone of the RSFSR. "Soviet Russia" 1975, 208 pp.

10. Kogan, F. N., "Weather, crop management and the variability of cereal yield," Meteorologiya i gidrologiya [Soviet Meteorology and Hydrology], no. 7, pp. 60-67, 1977.

11. ____________, "The grain crop productivity in the Chernozem Zone of the European Territory of the USSR and the possibility for its prediction, based on weather information," Trudy GMTs, no. 174, 1977, pp. 3-23.

12. ____________, "Geographical aspects of climate and weather limitations for cereal production in the USSR." 15th Conference on Agriculture and Forest Meteorology and 5th Conference of Biometeorology. AMS, Anaheim, CA, April, 1981, pp. 128-130.

13. __________, Grain Production in the USSR: Present Situation, Perspectives for Development and Methods for Prediction. Staff Report No. AGES 810904. USDA SRS RD, 1981, 81 pp.

14. __________, "Testing Techniques for the Reduction in Number of Variables in Crop-Weather Regression Models." Seventh Conference on Probability and Statistics in Atmospheric Sciences. AMS, Monterey, CA, November 1981, pp. 134-138.

15. Kulakovskaya, T. N., "Basic directions for raising productivity of land in the Non-Chernozem Zone." Doklady VASKHNIL No. 8, 1981, pp. 3-6.

16. Kurlovich, B. S., "Creation and introduction of new grain crop varieties." Selektsiya i semenovodstvo No. 8, 1981, pp. 7-10.

17. McQuigg, J. D. Economic impacts of weather variability. University of Missouri, Columbia, Missouri, 1975, 256 pp.

18. Nuttonson, M. Y., Barley-climate relationships and the use of phenology in ascertaining the thermal and photo-thermal requirements of barley. American Institute of Crop Ecology. Washington D.C., 1957, 280 pp.

19. Obuhov, V. M., Yield and meteorological factors [in Russian], Gosplanizdat, Moscow, 1949, 349 pp.

20. Palmer, W. C., Meteorological drought. U. S. Weather Bureau, Department of Commerce Research Paper #45, 1964.

21. Phinney, D. E., Meteorological data density and accuracy requirements for the yields estimation subsystem. Lockheed Technical Memorandum LEC-9572, 642-2097, Houston, TX, 1976.

22. Polevoy, A. N., Mizina, T. I., "Agrometeorological forecast of barley yield" in collection: Research in agricultural meteorology of the Non-Chernozem Zone, [in Russian], Gidrometeoizdat, Leningrad, 1978, pp. 81-88.

23. Prinsky; R., Soviet's Ability to Purchase Grain May Be Hampered by Tight Supply. The Wall Street Journal, 5/6, 1981.

24. Protserov, A. V., "Assessment of the water availability index for spring, wheat, oats and barley." In collection: The methods for the analysis and assessments of agrometeorological conditions" [in Russian], Gidrometeoizdat, Leningrad, 1957, pp. 49-53.

25. Ravelo, A. C., and W. L. Decker, "The Probability Distribution of a Soil Moisture Index," Agriculture Meteorology, 20, 1979, pp. 301-312.

26. Sakamoto, C. M., The Z-index as a Variable for Crop Yield Estimation. Agriculture Meteorology, 19, 1977, pp. 305-313.

27. Sebaugh, J. L., One, Two and Three Line Segment "Straw Man" Models. ESS Staff Report No. AGESS810514 USDA ESS SRD, Washington D.C., 1981, 57 pp.

28. Sebaugh, J. L., Evaluation of the CEAS Trend and Monthly Weather Data Model for Spring Wheat in North Dakota and Minnesota. Staff Report No. AGES811214 USDA SRS SRD, Washington D.C., 1981, 45 pp.

29. Selyaninov, G. T., Climatic principle of the agricultural specialization. Map of the agroclimatic zone of the USSR. In "Rastenievodstvo SSSR" [in Russian], V. 1 Part 1, Leningrad, 1933.

30. Sirotenko, O. D., "On the consideration of the autocorrelation for agrometeorological observations." Trudy IEM, No. 8, 1969.

31. _______________, "Physicomathematical modeling in agriculture meteorology," Meteorologiya i gidrologiya [Soviet Meteorology and Hydrology], no. 3, 1978, pp. 80-88.

32. South-East of the European territory of the USSR "Nauka" Moscow, 1971, 457 pp.

33. Thompson, L. M., "Weather and technology in the production of wheat in the United States." Journal of Soil and Water Conservation, 23, 1969, pp. 219-224.

34. Urbakh, V. U., The Biometric Methods [In Russian] "Nauka," Moscow, 1964, 415 pp.

35. Weather-Crop Yield Relationships in Drought-Prone Countries of Sub-Saharan Africa: Final Report, Climatic Impact Assessment Division-Models Branch of NOAA/EDIS, and University of Missouri-Columbia, Atmospheric Science Department, 1979, 177 pp.

36. Wilson, W. W., T. L. Barnett, S. K. LeDuc, F. B. Warren, Crop yield model test and evaluation criteria. AgRISTARS Yield Model Development Project, Document YMD-1-1-2, 1980.

37. Yao, A. Y. M., "The R-Index for Plant Water Requirements," Agriculture Meteorology, 6, 1969, pp. 259-273.

38. Yield Model Development. Project implementation plan for fiscal year 1981 and 1982. Lyndon B. Johnson Space Center, Houston, TX, 1981.

39. Yield Model Development Project Implementation Plan for Fiscal Year 1982 and 1983. Lyndon B. Johnson Space Center, Houston, TX, November 1981.

APPENDIX I

DESCRIPTION OF THE EXISTING BARLEY MODELS IN THE USSR

The beginning of weather-crop modeling in Russia goes back to the end of the previous century with the pioneer work of Brounov [5]. In the 1920's-1930's Obukhov continued this work based on updated information concerning crop productivity and environmental conditions. These early efforts mainly involved the use of calculators. Absence of computers and limitations of the regression technique did not allow the expansion of this type of modeling until the 1950's. In the early 1950's when these drawbacks were partly overcome and the first symptoms of shortages of food in the world became apparent, the development of yield models, both in the USSR and throughout the world, was highly stimulated. Since that time different approaches to crop modeling have been developed. So far, only two basic approaches have been applied by scientists in the USSR for modeling barley yield. They are analogue and bio-physical. A third approach, remote sensing, is in the process of development.

A simple analogue equation was developed by Polevoy and Mizina to forecast barley yield in the highly barley oriented Non-Chernozem zone [22]. The equation is:

$$Y/Y_{max} = -0.0042T_1 - 0.0077T_2 + 0.0008W_1 + 0.0031W_2 + 0.0015H + 0.5878 \quad (1)$$

where Y, Y_{max} are barley yield (Q/HA) predicted and "maximum geographical yield" respectively; T_1, T_2 are mean air temperature (°C) over periods of emergence-shooting (1) and shooting-heading (2), respectively; W_1, W_2 are mean available soil moisture (mm) in the top 20 cm of soil over periods 1 and 2 respectively, H is mean height of barley at heading. The "maximum geographical yield" represents some potential yield of a crop for particular climatic and technological conditions in a region. It can be defined by using the Goombol's technique [6], which is based on the empirical estimates of the parameters of yield's distribution.

Equation (1) shows that even in the Non-Chernozem zone, which in general has good natural water resources, water is still a factor limiting productivity of barley. Unfortunately, this method cannot be easily used as the information on the actual height of crop, soil moisture, and crop calendar, which are usually measured in the field, is not available.

A more complex analogue model has been developed by Dmitrenko [6,7]. This model is based on an assessment of the technology index expressed in form of the "maximum geographical yield" and also on an assessment of the index of weather productivity.

$$Y = Y' * C * \frac{m(T)}{M(T_o)} * \frac{m(P)}{M(P_o)}, \quad (2)$$

where $C = (1-p)\left[1- \frac{K-K_o^2}{K_o}\right]$ (3)

$$\frac{m(T)}{M(T_o)} = e^{-a\left(\frac{T-T_o^2}{10}\right)} \quad (4)$$

$$\frac{m(P)}{M(P_o)} = (1+ \frac{P-P_o}{P_o-P_{min}})^{a_1} * (1- \frac{P-P_o}{P_{max}-P_o})^{a_2} \qquad (5)$$

In these equations Y, Y' present model predicted yield and maximum yield (Q/HA), respectively; p is the density of plant population in spring, it can be calculated as a ratio $\frac{A_W}{A}$, where A_w is the area with winterkill or other damage to crops during winter and A is the whole area of crop; K, K_0 are the actual and optimal number of stems per plant, respectively; T, T_o are actual and optimal temperature (°C), respectively; P, P_0, P_{min}, P_{max} are actual, opitmal, minimal and maximal precipitations (mm), respectively; a, a_1, a_2 are empirical parameters; m(T) and M(P) are parameters for assessment of the weather productivity for the crop. In the terminology of the author of the model, weather productivity is a ratio of actual to optimal weather.

As is seen from equations (3)-(5), the idea is to define the proportionate reduction in the maximum geographical yield (Y') when actual parameters of weather and conditions of crop differ from optimal ones. The greater the difference between these parameters, the greater is the difference between the maximum yield and actual yield. In the ideal case of optimal weather and crop conditions, when $A_w = 0$, $K = K_0$, $T = T_0$ and $P = P_0$ then in accordance with equations (3)-(5)

$$C = \frac{m(T)}{M(T_0)} = \frac{m(P)}{M(P_0)} = 1$$

and the actual yield will be equal to the maximum yield. If weather and crop conditions are not optimal and $A_w > 0$, $K \neq K_0$, $T \neq T_0$ and $P \neq P_0$ then

$$C < 1, \frac{m(T)}{M(T_0)} < 1 \text{ and } \frac{m(P)}{M(P_0} < 1$$

In the case where any of these inequalities hold, the predicted yield will be less than the maximum geographical yield (Y < Y'). Using historical weather and yield data, the author of the method estimated optimal weather variables for every month of the year and for various stages of spring barley production (Table AI-1) [7].

Table AI-1. Optimal mean air temperature (T_0, °C) and optimal total amount of precipitation (P_0, mm), for spring barley by stage of growth in the Ukraine Republic, USSR

Barley Stage (approximate period)	Weather	
	T_o	P_o
Presowing (December-February)	2	100
Sowing-rooting (March-April)	4	100
Leaves formation (May)	10	120
Formation of reproductive parts (June)	18	90
Ripeness (July)	22	<15

Unfortunately, these estimated optimal values of precipitation and temperature cannot be considered as constants since technology changes. This model, like the previous one, requires such field measurements as the crop calendar stage and density of the crop, creating impossible problems in the remote application of this model.

A bio-physical model has been developed by a group of scientists under the guidance of Sirotenko [31]. The basic equations written in differential form follow below:

$$m_p^{j+1} = m_p^j + \frac{a_p^j \varepsilon}{1+R_R} \Phi^j - \frac{1}{1+R_R} K \qquad K = ({}_{\tau}^{i}R_0 + \beta_p^i)m_p \text{ for } 1, s, r$$

$$K = {}_{\tau}^{i}R_0 m_c^i - \sum_{I,s,r} \beta_p^i m_p^j \text{ for } c$$

$$W_j^{j+1} = W_i^j - TR_i^j - \sigma_i E^j - q_{i-1}^i - q_i^i$$

$$h^{j+1} = j^j + \frac{0.1E^j}{\Theta_{mfc} - \Theta_{mh}} \qquad (6)$$

where p, 1, s, r, c; and i = 1, 2, 3, 4

$$\Phi^j = \frac{\tau^j \psi_T^j L^j}{2(r^j - r_c)} \quad CO_2 + a\bar{I}^j r^j - \quad [(a\bar{I}^j r^u - CO_2)^2 + 4r_c \bar{I}^j CO_2)]^{1/2}$$

Here j is the number of the time step (of the day); i is the number of the soil layer; 1, s, r, and c are the leaf, stalk, root, and ear subscripts, respectively; w_i^j are the amounts of stored moisture; TR_i^j are the amounts of moisture transpired; σ is a logic variable ($\sigma_i = 1$ if $i = 1$ and $\sigma_i = 0$ otherwise); q_{i-1}^j, q_i^j are the moisture fluxes across the upper and lower boundaries of the i-th soil layer, respectively; h^j is the depth of the absorbing layer; E^j is the evaporation from the surface of the soil; τ is the length of the day; L^j is the area of the photosynthesizing phytomass; $\bar{I}^j$ is the daily average total radiation; ψ_T^j is the temperature coefficient of the total photosynthesis Φ^j; r is the diffusion of resistance on the path of the carbon dioxide, which depends on the soil moisture pressure; Θ_{mfc} and Θ_{mh} are the minimum field moisture capacity and the maximum hydrocapacity of a unit volume of soil, α_p^j; and β_p^j are biological functions; R_0, R_R, α, and r_c are constants; and CO_2 is the carbon dioxide concentration. In addition to equation (6), it is necessary to use many other equations enumerated in [1] for calculating such parameters as evaporation from the surface of the soil, transpiration of the crop, total area of green part of crop and elevation of the sun. Sometimes it is necessary to use special approaches and experimental information for the identification of many parameters of the model. All of these measures, however, do not improve the accuracy of the model.

Initial information required by this model is the following: dry weight of the plants from a unit of the planted area and available moisture stored in the soil by layers at the time step j = 0 (when the seedlings appear). It is also necessary to specify the latitude, cumulative degree days (in the form of effective temperature) starting from emergence which reflects biological time, some soil properties, daily average air temperature, air humidity, duration of sunshine, and the daily precipitation totals.

The application of this model for assessment of barley growth considers the calculation of biomass (total and broken down by leaves, stems, roots, ears), leaf area, evapotranspiration, and available soil moisture for ten-day periods. The attempts of the same group of scientists to simplify this model and to adjust it to the requirements of barley yield prediction on a large-scale basis led to the development of an arbitrary procedure. This procedure considers a search for an analogous indicator of the current weather pattern in the past history and determination of the corresponding changes in barley yield, estimated from the model, versus the climatic values of yield estimates [31]. The final quantitative estimates are made in a very simple way by evaluating the agrometeorological conditions that determine the fluctuations in barley yield (from historical data). Table AI-2 shows the technique for such quantitative assessments.

Table AI-2. Quantitative evaluation of agrometeorological conditions determining barley yield (Valdai Station [31])

Ten-day period	May	June			July		Yield Estimate, %	
	3	1	2	3	1	2	Ten-day	Accumulated
-	N	N	N	N	N	N	0	100
3	DC	N	N	N	N	N	-13	87
1	DC	MW	N	N	N	N	10	97
2	DC	MW	DW	N	N	N	-17	80
3	DC	MW	DW	MC	N	N	- 3	77
1	DC	MW	DW	MC	DC	N	-16	61
2	DC	MW	DW	MC	DC	DW	5	66

Note: N is normal conditions; DC is dry and cool; DW is dry and warm; MW is moist and warm; MC is moist and cold.

This simplified barley yield model has many of the weaknesses described in Appendix II. Some errors in the yield estimate arise due to imposed restrictions, some due to errors in measurements of such unusual parameters as dry matter at the period of barley emergency or agrohydrological properties of soil. The model does not reflect the technological input into crop productivity. The application of this model to forecast Soviet barley yield without in-country observations is connected with difficulties in determining plant biomass, water stored in the soil, agrohydrological properties of soil and others.

APPENDIX II

OVERVIEW OF MODEL DEVELOPMENT AND UTILIZATION

Presently, three different approaches to environment-crop modeling can be identified. They are the following: analogue (regression, statistical), biophysical (physiological, physical-mathematical) and remote-sensing (physical, distant). Each of these is based upon different theoretical principles.

Analogue modeling is a type of simulation which is based on the statistical description of the interaction between crop and environment. This is extracted from the historical data through the application of general knowledge of physiology, climate, soil, technology and also statistical tools. Bio-physical modeling considers a description of the crop environment system based on energy- and mass-exchange in the system. In general, this modeling involves solution of equations for such vital plant processes as photosynthesis, respiration, water demand, and mineral nutrition, and also for the basic hydro- and thermodynamic equations which approximate the entropy of the continuum. Remote-sensing modeling is founded on the quantitative estimates of the physical properties of crop canopies. It involves the measurement of spectral characteristics (reflectance, absorption, brightness) of plants with respect to such characteristics as plant productivity and distinguishing between crops.

Each of these three approaches has some advantages and disadvantages which can be discussed from different aspects. The following important aspects of model development and application can be distinguished:

> accuracy and precision of models and their time and spatial (size of area covered by the model) resolution;
>
> applicability of models for prediction of yield, management of crop productivity, or improving knowledge about crop-environment interaction;
>
> cost, rapidity, and simplicity of model development and application.

Accuracy, precision, and time and spatial resolution affect other aspects. The remote-sensing models have the possibility for coverage of large areas. Indeed the satellite imagery normally covers a very wide band of land. But the accuracy of the current remote-sensing models is not sufficiently high to be very effective in practice for the assessment of crop productivity. This is partially due to the imperfections in the spectral measurements and in the techniques for measuring physical properties of plant canopies, and partially due to the imperfection in methodology for assessment of grain output through estimating the plant biomass. The latter problem is very important, especially at the present time when short-stemmed small grain crops are being introduced. These crops have smaller biomass and higher cropping power (yield potential). In addition, the amount of satellite imagery which can be acquired for remote-sensing models may be limited by the number of cloud-free satellite overpasses. At present, this could be only a few times per season for the Landsat satellite but much more often for Metsat. the number of imagery acquisitions could be increased by using additional Landsat satellites, but this would be very costly.

The time resolution of the bio-physical models varies from one to twenty-four hours. These time intervals appear to be optimal for these models. The recent approach [1] of increasing the time interval from daily to ten days in order to simplify the operational use appears arbitrary and inefficient. The concern is that the time-averaged meteorological variables do not have the same high accuracy required for the assessments of photosynthesis, respiration and other processes which constitute the basic equations of these models. The spatial resolution for the bio-physical model does not usually go beyond the scope of a single field, where the detailed observations of such variables as radiation, available soil moisture, evapotranspiration, CO_2 flow, and biomass required as inputs to these models are normally measured.

The transition from a small field to large-area yield assessment in bio-physical models is usually connected with obtaining average weather data from a very dense meteorological network [8]. This also involves the development of some indirect methods to estimate crop characteristics and some other parameters on a large-scale basis. The accuracy of bio-physical models depends on how completely physiological and environmental processes are represented in them and how accurately these equations describe all these processes. Their accuracy also depends on the method of transition from biomass or other parameters of crop productivity which are generally considered to be an output of the bio-physical model to the actual production or yield, and the accuracy with which input variables can be measured. Unfortunately, none of the bio-physical models currently offers a reasonable sub-model for accurate approximation of such vital processes as plant nutrition or organic matter distribution between the plant's organs. Many processes, even those which have been studied fairly well (photosynthesis, evapotranspiration) have been represented in the models only with simple equations. This is because the complicated equations require the specification of many coefficients which can be estimated accurately only by laboratory experiments. But even the simple equations have some imposed restrictions which limit model performance. One major limitation is that the bio-physical model requires that a large number of model parameters be estimated. This is one more very important problem of the bio-physical type of modeling which affects the model's accuracy. Usually at least a couple of dozen parameters are required to be identified [1]. The disadvantages of transition from biomass to the actual production of the crop has been discussed above. Some methods developed to avoid this transition do not improve model accuracy, since these methods consider the application of experimentally defined portions of biomass passing to roots, stems, leaves and grain. And finally, it should be mentioned that the experimental definition of a great number of coefficients based on micro-scale studies is usually quite different from that which is based on large-scale studies.

Time resolution for analogue models normally covers time intervals from one day to one month. The optimal time-interval for this type of modeling could be defined at 7 to 10 days considering the accuracy of the models, possibilities for measuring crop response to environment and the model's application. The space resolution in analogue models is normally considered to be good, on the area of 0.3-10.0 million hectares. Although it is possible to make fairly reliable assessments of crop productivity based on models with resolution up to 20 million hectares [11]. The author's experience has shown that possibilities exist for developing models to predict yield for very large areas (more than 100 million hectares). In such cases a special technique for weighting weather variables over space must be applied.

Normally, the constructed analogue models reflect the general behavior of crops within the range of the past environmental conditions. Therefore, the longevity and high quality of historical data play the most important roles in the ability of the model to adequately describe the real behavior of crops and to produce accurate estimates of yield. There are some other problems (methodological, statistical, informational) which complicate analogue modeling and limit their accuracy.

When simplicity and cost are considered, the differences between the three types of models are great. The analogue models are the simplest and require the least expenditure of resources and time for their development. The bio-physical models are more complicated and costly than analogue models and require much more time for their development. The most costly models are those based on the remote-sensing approach. They cannot be considered as complex, but they do require timely acquisition of the remote sensing imagery. Also, algorithms for translating the spectral data first to biomass and then to probable yield have not yet been identified.

One of the most important aspects of modeling is the application of the models. Three major fields of model application can be singled out. They are: prediction of crop output, management, and improving knowledge about environment-crop systems. Taking into consideration the preceding discussion of the accuracy, time and spatial resolution of the models and also such characteristics as the simplicity, cost and time requirements for model development, the analogue models are, at least currently, the best for application in large-scale crop output predictions. This is so because these models have the best combination of the desired characteristics discussed above. Besides that, they do not require dense weather station networks for obtaining space-averaged weather variables. The accuracy of measuring (or estimating) crop production usually corresponds to the accuracy of space-averaged weather from a limited number of stations. One of the important advantages of analogue models relates to their capability to assess crop yield well in advance of harvest.

Unfortunately, the analogue models cannot be considered very useful either for managing crop productivity or for improving understanding of environment-crop interaction. In this respect, the bio-physical models have greater potential. But in the field of forecasting of agricultural production, especially for a large area and with a good lead time and accuracy, the bio-physical models are less efficient than are the analogue models.

The remote-sensing models can be considered potentially useful for the purpose of a large scale forecasting of agricultural production, particularly in identifying conditions of extreme drought. But, due to the lead time required, these models give way to the analogue ones. Remote-sensing models are also less efficient than the bio-physical models in both their application to crop production management and to deepening our understanding of environment-crop systems.

APPENDIX III

SOME CONSIDERATIONS IN THE APPLICATION OF INDICES FOR YIELD MODELING

In regions which do not have a long period of historical yield data to develop weather-crop models, different indices are widely used to estimate crop growth conditions and probable crop yield. In general, these indices specify the ratio or difference between actual available water and required water for optimal crop growth within existing weather conditions. Indices found to have application in practice for crop condition and production assessment would include hydro-thermal [29], R-index [37], Z-index [26], water availability [24], soil moisture [25], Palmer Drought [20], crop moisture ratio [35], and yield moisture [3].

As indices normally combine both precipitation and temperature, their application for weather-crop modeling could be considered useful in reducing the number of variables. Thus, we employed some of the mentioned indices and compared the results of weather-crop modeling based on these indices with the results obtained by using those proposed in this study. The R-index (RI), the Z-index (Z) and the soil moisture index (SMI) were tested. The R-index is a ratio of actual to potential evapotranspiration. The Z-index is defined by the difference between actual precipitation and "climatically appropriate" precipitation. The soil moisture index is derived from the ratio between plant available water and the maximum plant available water.

Figure AIII-1 presents the dynamics of correlation coefficients for barley yield with monthly precipitation, Z- and SMI-indices. In general, the Z- and SMI-indices' correlation coefficients (and also the RI's curve for some sub-periods that are not plotted on this figure) are quite similar in the response of yield to precipitation. This means that although indices are calculated based on both precipitation and temperature, they still reflect mostly precipitation dynamics in yield fluctuation rather than temperature. Of all tested indices, the correlation curve for Z-index most closely matches the correlation curve for precipitation.

Despite this close match, models based on originally measured precipitation and temperature have better statistical assessments (R^2 and MSE) than those which are based on calculated indices (Table AIII-1). Thus, only precipitation and temperature were taken into consideration in the process of development of barley-yield models in this study.

Table AIII-1. Comparison of models based on precipitation and temperature and calculated indices in Kazakh economic region, USSR, 1945-1978

Model developed based on variables for	R^2				MSE			
	P,T	Z	SMI	RI	P,T	Z	SMI	RI
Separate months:								
1-19	0.88	0.81	0.83	-	432	463	405	-
13-19	0.80	0.56	0.49	-	331	560	656	-
Sub-periods:								
5-8, 9-11, 12-14, 15-19	0.63	0.48	0.21	0.48	453	597	908	571
12-14, 15-19	-	0.37	0.16	-	-	672	1204	-

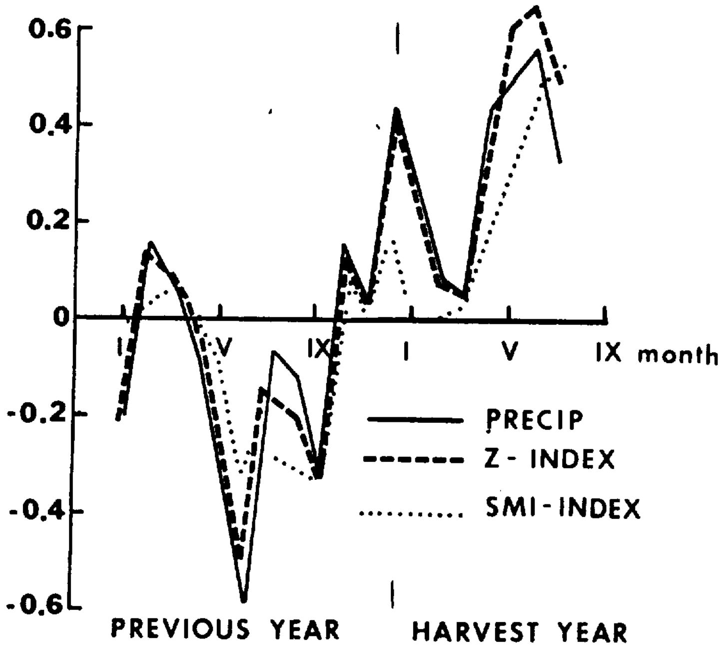

Figure AIII-1. Correlation coefficient for barley yield with precipitation, Z-index and SMI-index, in Kazakh economic region, USSR, 1945-1978.