Application of regression models to assess and predict the qualitative composition of the crop

I N Besaliev¹, D I Parfenov¹,², I P Bolodurina¹,², S S Akimov¹,²

¹Federal Research Centre of Biological Systems and Agro-technologies of the Russian Academy of Sciences, 29, 9 January Street, Orenburg, 460000, Russia
²Orenburg State University, 13, Pobedi Ave., Orenburg, 460018, Russia

E-mail: sergey_akimov_work@mail.ru

Abstract. The article examines the influence of various agro-climatic factors on grain quality in a continental climate. As a result of the study, one-factor models of the relationship of each of the eight factors with each specific indicator of grain quality have been developed. The degree of correlation of all the listed factors (temperature, precipitation, humidity, hydrothermal coefficient) with quality indicators was calculated, the most significant factors were selected, to which the average temperature, precipitation and moisture deficit were attributed, their multivariate models were calculated, describing the cumulative effect of each factor on a specific quality indicator. - nature, vitreousness, gluten and protein content. All equations are combined into a single system.

1. Introduction

To increase the efficiency of the implemented agrotechnical measures and the stability of agricultural production, it is necessary to fully take into account the peculiarities of the weather in a specific territory, study the correlation of weather factors with productivity, identify the optimal parameters of control actions that contribute to the formation of a high yield of a particular crop.

Currently, there is a problem of accurate assessment of weather factors, as well as a comprehensive interpretation of indicators of the quality of agricultural crops. This circumstance predetermines the need for the systematization of data collected by various farms and special services, as well as the development of software in order to automate the process of their processing.

The scientific novelty of the task is to assess the quality of the crop being grown even before the harvesting stage, the calculation of which would be based on measurable weather factors. This approach will make it possible to assess the class of the growing crop immediately, before the expiration of the ripening period, which will make it possible to make scientifically grounded decisions on the distribution of the crop in accordance with the quality parameters of crops in time.

The inherent ability of cultivated plants to regulate their own internal environment determines their extremely significant relationship with soil, climatic and weather conditions. Any agricultural culture has its own specific interdependence of the size and quality of the crop on certain factors: temperature drops, humidity, lighting, the content of mineral nutrition elements, and many others. For this reason, the change in yield from year to year for most crops today depends on uncontrolled environmental conditions.

Choosing the right variety for specific locations is one of the key factors that contribute to efficient grain production. The quality of grain depends on the impact of multiple conditions, some of which
are still not amenable to human control (natural conditions), others are able to regulate (agrotechnical conditions).

To improve the assessment and forecasting of grain quality indicators, various mathematical methods, Internet and information technologies are currently widely used. The use of mathematical and regression models for forecasting in agriculture is largely due to the introduction of highly informative research methods into practice.

The assessment of grain quality indicators was carried out by many researchers. Zhao Jun-fang, Guo Jian-ping, WU Ding-rong, Fang Shi-bo, E You-hao conducted a study of the grain quality indicator using nitrogen and nitrate leaching as a key parameter, different regions of the Netherlands served as the study site, and the main conclusion - about the great prospects of "organic management" [1].

Research conducted by R.R. Ismagilov, A.A. Nigmatyanov and A.M. Shayakhmetov showed a very strong dependence of the yield and quality of grain on the average daily air temperature in the conditions of the Republic of Bashkortostan [2].

In the work of G N Sandakova and etc. studies were carried out to create an integral indicator of wheat quality, taking into account agrometeorological factors. The resulting indicator met the criteria of accuracy and reliability, based on a set of regression equations describing the dependence of quality on weather factors [3].

G. R. Brankovic, D. Dodig, M. Z. Zoric, G. G. Sultan-Momirovic, V. Dragicevic, N. Duric using the method of regression modeling considered the interaction of agro-climatic factors in Serbia on the final characteristics of wheat grain, such as yield and quality [4].

In the study by Leshchenko MA, the results of the analysis of the correlation dependences between the SDS sedimentation index and the main qualitative characteristics of durum winter wheat grain are presented. The analysis of correlation dependences revealed the presence of reliable positive links between SDS indicators with the nature of the vitreous body and the nature of the vitreous body; A close positive correlation was found between the nature of grain and glassiness [5].

Feng H. and etc. In my work, we assessed soil quality indicators under the influence of five-year diversified and monocultural farming systems, where it was determined that the quality of the harvest directly depends on the quality of soil factors [6].

Asseng S, Bar-Tal A., Bowden JW, Keating BA, Van Herwaarden A., Palta JA, Huth NI, Probert ME, through a simulation model of a crop production system, which includes various agrochemical and climatic factors, predicted changes in yield and quality composition of wheat grain for forest-steppe zones [7].

In the studies of Abdollahi L., Hansen E. M., Munkholm L. J., Rickson R. J., an overall assessment of the quality of the soil on wet sandy loams was carried out, the effect of soil cultivation, as well as location and rotation, was assessed. Conclusions were made about the relationship with grain quality [8].

Haboudane D., Miller J.R., Tremblay N., Zarco-Tejada P.J., Dextraze L. use the method of integrated narrow-band vegetation indices to predict the yield and composition of crops for use in precision agriculture [9].

Sharma S., Singh P., Kumar S. in their work investigated the reaction of soil carbon deposits, enzymatic activity and yield to the inclusion of nitrogen and straw in the sowing system of rice and wheat cultivation in Northwest India [10].

Thus, most studies agree that various agro-climatic factors affect the quality indicators of grain.

2. Materials and methods
Based on the above bibliographic analysis, and using modern scientific developments in the field of durum wheat quality, we will compile a list of weather factors that can have the most significant impact:

- X1 - average air temperature, °C;
- X2 - minimum air temperature, °C;
- X3 - maximum air temperature, °C;
- X4 - amount of precipitation, mm;
- X5 - average relative air humidity, %;
- X6 - minimum relative air humidity, %;
- X7 - average deficit of air humidity, Pa;
- X8 - hydrothermal coefficient, mm / °C.

The main indicators of the quality of durum wheat are selected:
- Y1 - nature, g/l;
- Y2 - glassiness, %;
- Y3 - protein content, %;
- Y4 - gluten content, %.

The data for analysis were obtained as a result of cooperation with the Federal State Budgetary Scientific Institution "Federal Scientific Center for Biological Systems and Agricultural Technologies of the Russian Academy of Sciences."

The analysis used modern methods and tools of applied statistics.

To assess the correspondence of the empirical distribution to the normal law, the following were used:
- criterion of the Kolmogorov-Smirnov type;
- criterion omega-square;
- Shapiro-Wilk criterion.

The analysis of the relationship of factors was carried out using:
- Pearson correlations;
- Spearman's correlations.

The assessment of the strength of the connection between factors and indicators was carried out using the Chaddock shala.

The elimination of multicollinearity in the estimated factors was carried out using the method of principal components.

In the work, we used the statistical material of the examination of the grain quality of spring durum wheat in the Orenburg region for many years, as well as the results of laboratory studies of the grain quality of this crop, carried out in the experimental laboratory of the Federal Scientific Center for Biological Systems and Agricultural Technologies of the Russian Academy of Sciences.

3. Results and discussion

In Table 1, we will display all the obtained one-way regression models.

|   | Y1                  | Y2                  | Y3                  | Y4                  |
|---|---------------------|---------------------|---------------------|---------------------|
| X1| -0.35x^2 + 11.53x + 684.9 | -0.15x^2 + 5.49x + 30.71 | 0.02x^2 - 0.63x + 18.68 | 0.03x^2 - 0.62x + 31.29 |
| X2| -0.61x^2 + 12.23x + 720.81  | -0.2x^2 + 4.21x + 59.12  | 0.03x^2 - 0.59x + 17.16  | 0.04x^2 - 0.7x + 31.09  |
| X3| -0.34x^2 + 15.53x + 602.59  | -0.13x^2 + 6.87x - 4.42  | 0.02x^2 - 0.72x + 21.12  | 0.02x^2 - 0.82x + 34.25  |
| X4| -0.01x^2 + 0.78x + 764.4   | -0.004x^2 + 0.25x + 77.56 | 0.0002x^2 - 0.37x + 15.83 | 0.0001x^2 - 0.04x + 30.06 |
| X5| -0.05x^2 + 5.71x + 614.44  | -0.025x^2 + 2.61x + 15.28 | 0.0002x^2 - 0.33x + 26.48 | 0.001x^2 - 0.3x + 41.48 |
| X6| -0.08x^2 + 5.76x + 672.7   | -0.03x^2 + 2.09x + 49.56  | 0.004x^2 - 0.37x + 23.19  | 0.004x^2 - 0.5x + 40.81  |
| X7| -0.3x^2 + 7.7x + 738.55  | -0.18x^2 + 4.96x + 48.96  | 0.02x^2 - 0.23x + 14.53  | 0.02x^2 - 0.05x + 25.3  |
| X8| -27.95x^2 + 44.51x + 763.96 | -8.71x^2 + 10.95x + 78.24  | 0.98x^2 - 2.56x + 15.92  | 1.5x^2 - 3.84x + 30.41  |

The models indicated in Table 1 made it possible to carry out a number of calculations to determine the most optimal inclusion of various factors in the final models. On the basis of one-factor models, it
is possible to determine the fundamental possibilities of combining factors in terms of their influence on a specific key indicator.

Among other things, non-linear correlation coefficients were calculated (Table 2).

Table 2. Nonlinear correlation coefficients for each of the combinations "initial factor - key parameter"

|     | Y1   | Y2   | Y3   | Y4   |
|-----|------|------|------|------|
| X1  | 0.199| 0.172| 0.458| 0.575|
| X2  | 0.282| 0.194| 0.221| 0.184|
| X3  | 0.273| 0.523| 0.788| 0.652|
| X4  | 0.335| 0.381| 0.262| 0.213|
| X5  | 0.189| 0.284| 0.578| 0.562|
| X6  | 0.264| 0.459| 0.416| 0.473|
| X7  | 0.225| 0.247| 0.436| 0.449|
| X8  | 0.323| 0.285| 0.265| 0.198|

The obtained nonlinear correlation coefficients were significant in most cases. On the Chaddock scale, for eight paired coefficients, the relationship can be characterized as "noticeable" (0.5-0.7) (Figure 1).

An analysis of the set of initial factors suggests the presence of multicollinearity among some indicators (Spearman's correlation coefficients for temperature indicators range from 0.87 to 0.96). Taking into account this circumstance, it is determined the need to exclude some factors in order to obtain the most stable estimates of the obtained regression parameters.

The model also excludes the following factors that had the least impact::
- $X_2$ - minimum air temperature, °C;
- $X_3$ - maximum air temperature, °C;
- $X_5$ - average relative air humidity, %;
- $X_6$ - minimum relative air humidity, %;
- $X_8$ - hydrothermal coefficient, mm / °C.

The exclusion of these factors allows us to achieve one more goal: this process makes it possible to significantly simplify the applied factor models for key quality parameters, making them convenient for calculation.

Thus, as a result of the study, a system of four regression equations was obtained. The calculated parameters most fully characterize the dependence of quality indicators on the initial factors. The system looks like this:

\[
\begin{align*}
Y_1 &= 0.766x_1 - 0.215x_2 - 2.747x_3 + 799.478 \\
Y_2 &= 0.001x_1 - 0.057x_2 + 0.242x_3 + 79.193 \\
Y_3 &= 1.721x_1 - 0.987x_2 + 14.113x_3 + 98.234 \\
Y_4 &= 0.951x_1 + 1.128x_2 - 2.783x_3 + 651.347
\end{align*}
\]

4. Conclusion
As a result of the study, one-factor models of the relationship of each of the eight factors with each specific indicator of grain quality have been developed. The degree of correlation of all the listed factors (temperature, precipitation, humidity, hydrothermal coefficient) with quality indicators was calculated, the most significant factors were selected, to which the average temperature, precipitation and moisture deficit were attributed, their multivariate models were calculated, describing the cumulative effect of each factor on a specific quality indicator. - nature, vitreousness, gluten and protein content. All equations are combined into a single system.

5. Acknowledgment
The studies were supported by a grant from the President of the Russian Federation for state support to leading scientific schools of the Russian Federation (NSh-2502.2020.9), as well as in accordance with the R&D plan for 2019-2020 of the Federal Research Center for Biological Systems and Agricultural Technologies of the Russian Academy of Sciences (No. 0761-2019-0004).

References
[1] Strashnii V N 1977 Agrometeorological conditions and grain quality of winter wheat in the central black earth zone (Moscow) 25 p.
[2] Derevyanko A M 1989 Weather and quality of grain of winter crops (Leningrad: Gidrometeoizdat) 127 p.
[3] Khokonova M B, Adzhieva A A 2016 Patterns of loading and growth of dry biomass of spring wheat. *Bulletin of the Kabardino-Balkarian State Agrarian University named after V.M. Kokova* 2(12) 22–26
[4] Zhao J F, Guo J P, Wu D R, Fang S B, E Y H 2011 Climatic potential productivity of winter wheat and summer maize in Huanghuaihai Plain in 2011-2050. *Chinese J. of Applied Ecology* 22(12) 3189–3195
[5] Leshchenko M A 2015 The relationship of the SDS-sedimentation indicator with the main signs of the quality of grain of hard winter wheat. *Achievements of science and technology of the agro-industrial complex* 29(1) 20–23
[6] Sorokin A N, Morozova T M 2018 The effect of seed ozonation on productivity and grain quality indicators of spring wheat. *Vladimir farmer* 3(85) 32–35
[7] Neneidenko G N 2018 Fertilizer and improving the quality of wheat grain in the Upper Volga
region. *Vladimir farmer* 2(84) 20–27

[8] Sibirina T F, Melnikova E V, Mordvinova N M, Poluboyarinov N A, Belyakov A A 2020 Adaptive forecast of spring wheat grain quality in the forest-steppe zone with differentiation according to temperature and precipitation modes. *The era of science*. 21 19–35

[9] Sandakova G N, Besaliev I N, Panfilov A L, Karavaitsev A L, Kiyaeva E V, Akimov S S 2019 Influence of agrometeorological factors on wheat yields. IOP Conference Series: Earth and Environmental Science The proceedings of the conference AgroCON-2019 012022

[10] Sandakova G N Kryuchkov A G 2013 *Orenburg: Press Agency LLC* 101