Method for prediction of raw gluten content in wheat grain

A V Pasynkov¹ and E N Pasynkova¹, ²

¹ Leningrad Research Agriculture Institute Branch of Russian Potato Research Centre, Russia
² E-mail: pasynkova.elena@gmail.com

Abstract. The possibility of using the developed equations reflecting the dependence of the raw gluten content (Y, %) on the protein content (X₁, %) and the 1000-grain weight (X₂, g) to predict its (gluten) content in wheat grain is discussed. The developed equations allow to determine the main trends in the change in raw gluten content of wheat grains, occurring when the protein content and 1000-grain weight vary under the influence of various factors, and explain the contradictory nature of these dependencies available in the scientific literature. The equations can be used to control the quality of the manual or mechanical determination of the raw gluten content of winter and spring soft and hard wheat grains, as well as to test the operation and reliability of the various analyzers used in protein and gluten analysis. Key words: wheat, protein, raw gluten, 1000-kernel weight, multiple regression analysis

1. Introduction

Statistical processing using multiple regression analysis made it possible to obtain second-order equations reflecting the dependence of raw gluten content (dependent variable - Y) on protein content and 1000-grain weight (independent variables: X₁ and X₂, respectively) [1]. Further research allows concluding that the developed equations can be used to predict the gluten content of wheat grains after determining the protein content and kernel weight. To test this assumption, an algorithm was developed to test their prognostic capabilities [2].

Testing the prognostic possibilities of the equations showed that, as accurately as possible, the dependence of gluten content (Y, %) in wheat grain on protein content (X₁ = Ntot • 5.7, %) and 1000-grain weight (X₂, g) was reflected with the second order equation: Y (1) = - 41.928 + 0.081X₁² + 2.548X₂ – 0.028X₂², in which all variables (protein content, gluten and 1000-grain weight) are given by 12% humidity or air-dry matter (a.d.m.) [3].

Figure 1 show a graphical representation of equation (1) or a response surface of a function within the obtained experimental data. Analysis of the developed equation (1) and its graphic image showed that the dependence of the raw gluten content in wheat grain on the protein content and 1000-grain weight is non-linear. Each subsequent increase in protein (per unit) leads to a greater increase in gluten content compared to the previous (+ X₁²). Regardless of the protein content, as the 1000-grain weight increases from its minimum values, the gluten content increases too. Each subsequent increase in the 1000-grain weight (per unit) slows down the growth rate of the accumulation of raw gluten in the grain (+ X₂ - X₂²). After the 1000-grain weight reaches the extreme point (45.5 g), each subsequent increase results in a greater reduction in the gluten content of the wheat grain compared to the previous one [1-3].
2. Research methodology

If the protein content and/or the 1000-grain weight is determined on a base of an absolutely dry matter (d.m.), then using equation (1) to predict the raw gluten content in wheat grain, they are recalculated using coefficients of 0.88 and 1.136 respectively [4]. In order to exclude the recalculation procedure in cases where the raw gluten content and the 1000-grain weight are determined without taking into account the grain humidity (or at air-d.m.) matter, but the protein content is determined at d.m. matter, the second equation was calculated: \( Y(2) = -41.928 + 0.063X_1^2 + 2.548X_2 - 0.028X_2^2 \) [5].

![Figure 1](image.png)

Figure 1. Dependence of raw gluten content on protein content and 1000-grain weight

The next step is to check the accuracy of the forecast of raw gluten content in wheat grain. For this purpose, data of more than two hundred literary sources of Soviet, Russian and foreign authors were used, and subsequently summarized [5]. That is, independent experimental data were used on the content of protein, raw gluten and 1000-grain weight obtained by other authors during field experiments with wheat cultivars and in time frames and soil-climatic conditions differed from those used by the authors of the article. Substituting the experimental data obtained by other authors on protein content and 1000-grain weight into the developed equations (1) or (2), and using simple mathematical actions, the theoretical (\( Y_t \)) (predicted) raw gluten content in wheat grain is calculated. The next step is to compare the theoretical values of the protein content in wheat grain calculated by equations (1) or (2) with the experimental data (\( Y_e \)) or (\( Y_e - Y_t \)).

The criterion for assessing the accuracy of the developed equations is the permissible deviation regulated by the State standard of Russia (GOST R 54478-2011. Grain. Methods for determining the amount and quality of gluten in wheat): "Both results are recognized as acceptable if the critical difference in the results of the definitions by the amount of gluten ... does not exceed 2%" in absolute expression. The second criterion: the accuracy (justifiability) of the forecast or the ratio of the number of values, when the deviations of the experimental values of the raw gluten content in wheat grain from the theoretical (\( Y_e - Y_t \)) did not exceed the deviation to the total number of observations (n) expressed in %, allowed by State standard of Russia [5].

3. Results and discussion

Tables 1 and 2 show the algorithm and results of checking the accuracy of predicting the raw gluten content of wheat grains from independent data. Given that the scope of publication is limited and does not allow to give all cases of deviations (n = 4630 [5]), Tables 1 and 2 provide only some typical examples found in the comparative evaluation of the predictive capabilities of equations. Table 1 shows an algorithm for checking the accuracy of the prediction of the gluten content of grains of
different wheat cultivars according to equation (1). That is, for those cases where all variables (quality indicators: protein content, gluten and 1000-grain weight) are determined without taking into account the grain moisture content or at air-dry substance (air-d.m.) matter; and in Table 2 - according to equation (2), when the protein content is determined at a.d.m. matter and gluten content and 1000-grain weight - at air-d.m. matter.

The previously published data [2-6] and data presented in Tables 1 and 2 on the checking of effectiveness of predicting raw gluten content on the basis of the developed equations make it possible to conclude their relatively high predictive capabilities. The forecast of raw gluten content in wheat grain is confirmed at modification (fertilizers, sowing rates, sowing time, etc.) and genotypic (cultivar) differences [4].

**Table 1.** Algorithm for checking the accuracy of prediction of raw gluten content in wheat grains: \( Y(1) = -41.928 + 0.081X_1^2 + 2.548X_1 - 0.028X_2^2 \)

| \( X_1 \) | \( X_2 \) | \( Y_t \) | \( Ye \) | \( (Ye-Y_t) \) | \( X_1 \) | \( X_2 \) | \( Y_t \) | \( Ye \) | \( (Ye-Y_t) \) |
|---|---|---|---|---|---|---|---|---|---|
| Data on \( X_1, X_2 \) and Ye from article [7] (Poland) | Data on \( X_1, X_2 \) and Ye from article [8] (Pakistan) |
| 13.5 | 34.9 | 27.7 | 28.4 | 0.7 | 12.34 | 37.0 | 26.35 | 29.15 | 2.8* |
| 13.6 | 34.8 | 27.8 | 28.5 | 0.7 | 11.82 | 39.0 | 26.17 | 27.37 | 1.2 |
| 13.3 | 34.7 | 27.1 | 27.8 | 0.7 | 12.13 | 37.5 | 26.17 | 27.68 | 1.5 |
| 13.6 | 34.3 | 27.5 | 28.7 | 1.2 | 11.53 | 41.5 | 26.36 | 26.73 | 0.4 |
| 13.6 | 35.0 | 27.9 | 28.7 | 0.8 | 12.78 | 38.0 | 27.69 | 29.59 | 1.9 |
| 13.5 | 34.3 | 27.3 | 28.3 | 1.0 | 11.97 | 38.3 | 26.19 | 27.10 | 0.9 |
| 13.6 | 35.0 | 27.9 | 28.5 | 0.6 | 11.19 | 40.5 | 25.48 | 26.06 | 0.6 |
| 13.6 | 34.9 | 27.9 | 28.5 | 0.6 | n = 7 | NV = 1 | P = 83.7% |

| \( X_1 \) | \( X_2 \) | \( Y_t \) | \( Ye \) | \( (Ye-Y_t) \) | \( X_1 \) | \( X_2 \) | \( Y_t \) | \( Ye \) | \( (Ye-Y_t) \) |
|---|---|---|---|---|---|---|---|---|---|
| Data on \( X_1, X_2 \) and Ye from article [9] (Egypt) | Data on \( X_1, X_2 \) and Ye from article [10] (Russia) |
| 13.4 | 47.45 | 30.5 | 30.2 | -0.3 | 15.00 | 33.6 | 30.3 | 30.0 | -0.3 |
| 13.7 | 49.05 | 30.9 | 28.4 | -2.5* | 14.74 | 33.6 | 29.7 | 29.3 | -0.4 |
| 12.3 | 51.20 | 27.4 | 27.3 | -0.1 | 13.19 | 33.9 | 26.4 | 27.3 | 0.9 |
| 11.2 | 51.62 | 25.2 | 25.5 | 0.3 | 13.31 | 30.8 | 24.3 | 26.5 | 2.2* |
| 14.1 | 47.82 | 32.0 | 30.7 | -1.3 | 14.38 | 32.2 | 27.8 | 28.2 | 0.4 |
| 13.9 | 49.11 | 31.3 | 29.9 | -1.4 | 13.70 | 32.1 | 26.2 | 27.0 | 0.8 |
| 13.3 | 51.37 | 29.4 | 28.9 | -0.5 | 12.90 | 34.6 | 26.2 | 25.9 | -0.3 |
| 13.0 | 51.95 | 28.6 | 28.9 | 0.3 | 14.41 | 33.6 | 28.9 | 28.5 | -0.4 |
| n = 8 | NV = 1 | P = 87.5% | n = 9 | NV = 1 | P = 88.9% |

Notes: \( X_1 \) is the protein content of the grain, \%; \( X_2 \) - 1000-grain weight, g; \( Y_t \) - is the theoretical content of raw gluten (calculated by equation (1) or (2)), \%; \( Ye \) - experimental gluten content, \%; \( (Ye - Y_t) \) - deviations of experimental values from theoretical, ± 2.8* - values beyond ± 2%; \( n \) is the total number of observations; \( NV \) is the number of values that go beyond ± 2%; \( P \) - predictability, % (the same in text and Table 2)

Summary of data from 265 literary sources of domestic and foreign authors with a total number of observations \( n = 4630 \) on more than three hundred cultivars of winter and spring soft and hard wheat grown from 1959 to 2019 in various soil and climatic zones of Russia and abroad (Austria, Belarus, Bulgaria, Egypt, Kazakhstan, Pakistan, Poland, Slovakia, Ukraine, Croatia, and Czech Republic) showed that the number of values exceeding the limits regulated by the State standard of Russia (2%) was 844 or 18.2%. Thus, the justification of the forecast reached 81.8% [5].
Table 2. Algorithm for checking the accuracy of prediction of raw gluten content in wheat grains: \( Y(2) = -41.928 + 0.063X_1^2 + 2.548X_2 - 0.028X_2^2 \)

| \( X_1 \) | \( X_2 \) | \( Y_F \) | \( Y_3 \) | \( Y_{3-Y_F} \) | \( X_1 \) | \( X_2 \) | \( Y_F \) | \( Y_3 \) | \( Y_{3-Y_F} \) |
|--------|--------|--------|--------|------------|--------|--------|--------|--------|------------|
| Data on \( X_1, X_2 \) and \( Y_e \) from article [11] (Russia) | | | | | | | | | |
| 14.98  | 34.5   | 26.8   | 28.8   | 2.0      | 12.8   | 37.2   | 24.4   | 24.5   | 0.1       |
| 14.52  | 35.6   | 26.6   | 27.9   | 1.3      | 12.7   | 37.2   | 24.3   | 24.4   | 0.1       |
| 15.30  | 34.4   | 27.3   | 29.0   | 1.7      | 12.3   | 36.3   | 23.2   | 23.2   | 0.0       |
| 14.55  | 37.1   | 27.4   | 28.3   | 0.9      | 13.3   | 38.9   | 26.0   | 25.6   | -0.4      |
| 13.88  | 36.1   | 25.7   | 26.7   | 1.0      | 13.2   | 38.8   | 25.8   | 25.4   | -0.4      |
| 13.96  | 38.0   | 26.7   | 26.9   | 0.2      | 12.7   | 37.9   | 24.6   | 24.2   | -0.4      |
| 14.25  | 37.4   | 27.0   | 27.8   | 0.8      | 14.4   | 40.2   | 28.3   | 28.7   | 0.4       |
| 14.00  | 38.0   | 26.8   | 27.5   | 0.7      | 14.3   | 40.4   | 28.2   | 28.3   | 0.1       |
| 13.72  | 40.2   | 27.1   | 26.6   | -0.5     | 13.8   | 39.4   | 27.0   | 26.9   | -0.1      |
| 13.89  | 39.3   | 27.1   | 26.6   | -0.5     | 14.8   | 40.3   | 29.1   | 30.6   | 1.5       |
| 13.24  | 39.7   | 26.1   | 23.4   | -2.7*    | 14.6   | 40.5   | 28.8   | 30.0   | 1.2       |
| n = 11  | NV = 1 | P = 90.9% | | | n = 12  | NV = 1 | P = 91.7% | | | |
| Data on \( X_1, X_2 \) and \( Y_e \) from article [13] (Slovakia) | | | | | | | | | |
| 13.4   | 42.4   | 27.1   | 28.5   | 1.4      | 14.7   | 45.6   | 29.7   | 30.3   | 0.6       |
| 11.8   | 47.6   | 24.7   | 24.9   | 0.2      | 14.6   | 46.5   | 29.5   | 29.8   | 0.3       |
| 12.2   | 44.6   | 25.4   | 25.9   | 0.5      | 14.6   | 45.6   | 29.5   | 29.8   | 0.3       |
| 13.4   | 46.0   | 27.3   | 28.4   | 1.1      | 14.8   | 45.6   | 29.8   | 30.5   | 0.7       |
| 12.7   | 43.7   | 26.1   | 26.9   | 0.8      | 14.2   | 46.6   | 28.7   | 29.1   | 0.4       |
| 12.8   | 48.4   | 26.1   | 26.8   | 0.7      | 13.7   | 47.1   | 27.8   | 26.7   | -1.1      |
| 12.4   | 42.1   | 25.4   | 26.6   | 1.2      | 12.4   | 48.0   | 25.6   | 23.1   | -2.5*     |
| 13.0   | 54.4   | 24.5   | 28.1   | 3.6*     | 13.9   | 52.9   | 26.7   | 28.0   | 1.3       |
| 13.0   | 57.2   | 22.9   | 26.9   | 4.0*     | 14.5   | 47.0   | 29.2   | 29.6   | 0.4       |
| 14.2   | 41.8   | 28.4   | 29.7   | 1.3      | 13.7   | 40.7   | 27.2   | 25.9   | -1.3      |
| 11.7   | 42.8   | 24.5   | 25.0   | 0.5      | 14.6   | 37.4   | 27.6   | 30.2   | 2.6*      |
| n = 11  | NV = 2 | P = 81.8% | | | n = 10  | NV = 2 | P = 80.0% | | | |

Further studies have shown that the developed equations can be used to control the quality of manual or mechanical determination of the raw gluten content in wheat grains in almost all cases when the results of the analysis of the protein and gluten content in wheat grains, as well as the 1000-grain weight are given at actual or fixed humidity, or on a dry matter basis. In the case of determining protein and/or gluten content on various types of analyzers (Infralum FT-10 or FT-12, Spektra Star 2400, Infraneo, DA 7200 NIR analyzer, Inframatic 9200 and Infratec 1241), the developed equations (after determining the 1000-grain weight) can be used to indirectly check the accuracy of determining the raw gluten content in wheat grain [6].

4. Conclusion

The developed multiple nonlinear regression equations, reflecting the dependence of the raw gluten content of wheat grains on the protein content and the 1000-grain weight, can be used to predict the raw gluten content of winter and spring soft and hard wheat grains in all cases when the listed technological indicators are determined by conventional analytical methods. The equations can also be used to indirectly check the accuracy of determining the raw gluten content of wheat grain when the protein and/or gluten content is determined using various instruments.
References

[1] Pasynkov A V and Pasynkova E N 2011 Statistical dependencies of the main indicators of grain crops quality *Agrohimiya* 2: 24-40 (in Russian)

[2] Pasynkov A V, Dubovik D V and Pasynkova E N 2017 Raw gluten content prediction in wheat kernels based on the equations of multiple regression *Vestnik Kurskoj GSBA* 4: 8-14 (in Russian)

[3] Pasynkov A V and Pasynkova E N 2018 Features of using multiple regression equations to predict raw gluten content in wheat kernels. *Grain Econ. Russia* 4(64): 19-26. DOI: 10.31367/2079-8725-2019-64-4-19-26 (in Russian)

[4] Pasynkov A V and Pasynkova E N 2020. Improving the forecast of raw gluten content in wheat grain *Russian Agricul. Sci*. 46 (3): 207-12. DOI: 10.3103/S1068367420030143

[5] Kraska P, Andruszczak S, Kwiecińska-Poppe E and Pałys E 2014. The effect of tillage systems and catch crops on the yield, grain quality and health of spring wheat *Acta Sci. Pol. Agricult. Sustain.* 13(1): 21-38

[6] Muhammad N S, Khalid N, Nouman S, Muhammad A, Tabassum H and Samina K 2009. Quality evaluation of different wheat varieties for the production of unleavened flat bread (Chapatti) *Pakistan J. Nutri.*, 8(11): 1773-8

[7] Mahmoud S, Sayed A, Mahmoud I and Gamal Z 2011. Productivity, grain and dough quality of bread wheat grown with different water regimes. *J. AcroCrop Sci.* 2(1): 11-7

[8] Kolmakov Y V, Chibis V V, Pakhotina I V and Pristalyuk A L 2015 *Quality of wheat grain grown after predecessors in the southern forest-steppes* Izvestia of the Orenburg State Agrarian University 4(54): 18-20 (in Russian)

[9] Grigoriev Y P, Belan I A and Kolmakov Y V 2015 Competitive variety testing spring soft wheat in sub-taiga zone of Omsk Region. *Vestnik Kur'skoj gosudarstvennoj sel'skhozoyastvennoj akademii*, 7: 119-21 (in Russian)

[10] Centilo L V 2019 Formation of quality of winter wheat grain depending on the system of fertilization and tillage *Myronivskyi Visnyk* 8: 152-61 (in Ukraine)

[11] Žofajová A, Havrlentová M, Ondrejovič M et al. 2017 Variability of quantitative and qualitative traits of coloured winter wheat *Agriculture* 63(3): 102-11

[12] Jiří H, Leona L S, Jana B and Klára K et al. 2019 Genetic characterization and evaluation of twenty Chinese winter wheat cultivars as potential sources of new diversity for breeding *Czech J Genetics Plant Breeding* 55(1): 8-14