TEKA Kom. Mot. Energ. Roln. – OL PAN, 2008, 8a, 112–117

TESTING OF A MATHEMATICAL MODEL OF GRAIN POROSITY

Elżbieta Kusińska*, Andrzej Kornacki**

* Department of Food Engineering and Machinery,
** Department of Applied Mathematics and Informatics, University of Life Sciences in Lublin, Poland

Summary. Numerous researchers from the field of agricultural engineering use mathematical models for the description of processes. Because of the existing many systems and processes, there is a possibility of making different types of models. Their systematic may be found in the Pabis [1985] and Powierża [1997] books. Once the form of a model is fixed, what remains is the testing of the correctness of the model based on experimental data. Regression models are frequently employed in engineering studies. Commonly used as a measure of efficient suitability of such a model to experimental data is the coefficient of determination R^2. This index is excellent in the case of linear models and of non-linear models that are intrinsically linear (i.e. can be reduced to linear through suitable transformations). However, for non-linear and intrinsically non-linear models [Dobosz 2001]) complications appear. For example, in a logistic model [Bochniak and Wesołowska-Janczarek 2006] the coefficient of determination R^2 may assume negative values. In such a case it loses its natural interpretation. This paper presents a review of several statistical tests available in literature, that can be applied for testing of the correctness of non-linear regression models on the basis of experimental data. The statistical methods under presentation are illustrated by means of results of experimental studies concerned with oat grain porosity [Kusińska 2007, Bowszys and Tomczykowski 2007].

Key words: kernels, porosity, statistical tests, efficiency of mofel's fitness to empirical data.

INTRODUCTION

Porosity of grain in bulk is of high importance in processes of drying, ventilation and cooling. Its value depends on the size, moisture, shape and elasticity of kernels, on the condition of their surface, on the amount and kind of contaminants in the grain in bulk, and also on the distribution of the grain deposit in the silo.

Grain porosity is one of the fundamental parameters that determine the value of air flow resistance through grain deposit. Apart from that, resistance to air flow is affected by such factors as density, airflow velocity, depth or thickness of the deposit, method of filling or charging of hopper or silo, time of storage, direction of airflow, and the degree of compaction of kernels in the deposit [Chung *et al.* 2001, Gunasekaran *et al.* 1988, Jayas *et al.* 1991]. Analysing the literature one can conclude that materials used in the studies were frequently characterised by modest differences in porosity and density, but even that was a sufficient cause for an increase in resistance to air flow. The increase was caused primarily by cultivar-related features of grain, its moisture, method of silo filling, and of grain mass compaction, e.g. with the method of vibration [Łukaszuk 2005, Łukaszuk *et al.* 2006, Molenda *et al.* 2005]. Notably greater changes in the porosity and airflow resistance in natural-sized silos are caused by gravity load acting on grain mass [Kusińska 2006, Kusińska et al. 2006].

The objective of the study presented herein was determination of oat grain porosity in bulk that has a decisive effect on resistance to air flow during ventilation. The density of oat grain varied under the effect of the value of load applied and of the duration of load application.

MATERIAL AND METHODS

The study was conducted on oat grain cv. Sławko, with moisture content of 15%. Measurement of porosity was made using a replaceable cylindrical container with inner diameter of 76 mm and height of 560 mm (filling with grain sample was possible up to the height of 400 mm). Tests were performed for grain without any preliminary loading (in freely heaped state) and after the application of vertical load on the upper layer of grain, with values of 17.5; 35; 52.5 and 70 kPa. The maximum value of the load corresponded to filling a silo with grain material to a height of ca. 10 m. The duration of the application of the load was 2, 4 and 24 hours.

Following each application of loading, the grain sample was weighed, its height was measured, and its porosity was determined from the formula (1):

$$\varepsilon = \frac{\rho_w - \rho}{\rho_w}, \tag{1}$$

where:
ε – grain porosity [%],
ρ_w – specific density of the grain [kg·m^{-3}],
ρ – grain density [kgdm^{-3}].

The true volume of the grain, for the calculation of specific density, was determined by means of an air pycnometer. All measurements were made in three replications.

STATISTICAL TESTS FOR DETERMINATION OF CORRECTNESS OF NON-LINEAR REGRESSION MODELS

Fig. 2 presents the results of measurements of oat grain porosity with relation to the value of vertical load and to the duration of its application.

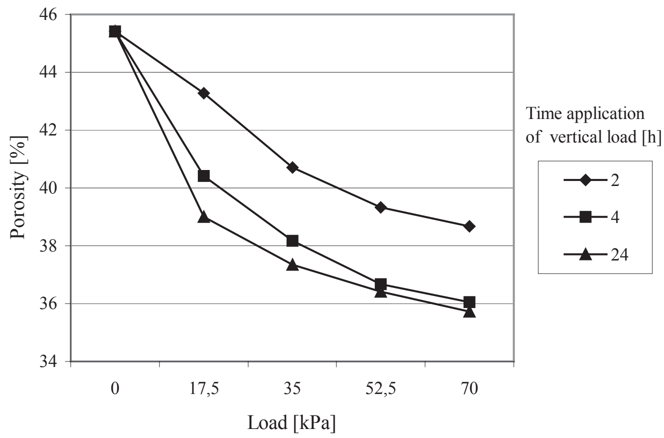

Fig. 2. Effect of vertical load and of duration of its application on oat grain porosity

Grain that was not subjected to loading was characterised by porosity of 45.1% (referred to as initial porosity ε_p). Application of vertical load to the upper layer of grain and extension of the time of its application caused a reduction in grain porosity. The lowest value of porosity (35.7%) was obtained for oat grain loaded with pressure of 70 kPa for 24 h. The observed decrease in porosity was by 21.38% (0.79-fold).

The results of that study were subjected to statistical analysis. The analysis showed that oat grain porosity depends significantly (at $\alpha<0.01$) on vertical load P and on the time of its duration . The relation is described by a non-linear regression model:

$$\varepsilon = 1{,}083 \cdot \varepsilon_p \cdot P^{-0{,}083} + \frac{2{,}947}{\ln \tau}, \tag{2}$$

where:
P – grain loading [kPa],
ε_p – initial porosity of the grain [%],
τ – duration of load application [h].

Equation (2) can be applied at ε_p = 45.1%, P = 0÷70 kPa, and τ = 0÷24 h.

The goodness of fitting of the obtained model to experimental data is commonly estimated by means of the coefficient of determination R^2. In our case the value of the coefficient is 96.6 %. It can be concluded, then, that the model under consideration describes well the experimental data.

If the sets of experimental data and of model predictions are considered as resulting from a probabilistic model, and thus being values of random variables, verification of the model should provide a conclusion as to the agreement of distributions of the features under study. We will now consider a few statistical tests that can be applied for that purpose.

Some authors, e.g. Powierża [1997], recommend verification of models through comparison of mean values by means of the t-student test. It is known that, for the difference of mean values, the t-student test requires the assumption of normality of features in both populations, and of the equality of variance. In our case we will test the normality of distribution using the Shapiro-Wilk test. The value of the test function is W=0.9062, while the critical values are $W_{0.01}$=0.929 and $W_{0.05}$=0.947. Therefore, regarding the relations $W< W_{0.01}$ and $W< W_{0.05}$ we can reject the zero hypothesis – the considered feature of porosity does not have a normal distribution. However, since the value of the Shapiro-Wilk test is very close to the critical values, and the test is fairly resistant to deviations from normality, in this case we can apply the t-student test:

$$t = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2 - 2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}. \tag{3}$$

In formula (3), $\bar{x}$, $\bar{y}$, S_1^2, S_2^2, n_1, n_2 denote the mean values, variance and populations of samples from both distributions, respectively. In the considered case of oat grain porosity, the value of t-student test function was t^0=0.0023. With the critical values $t_{0.05}$=1.9855 and $t_{0.01}$=2.629 of relation $t^0< t_{0.05}$, there are no grounds for rejection of the zero hypothesis of equality of mean values. Thus, it can be concluded that model (2) describes the experimental data well.

If the features under consideration do not have a normal distribution, non-parametric tests can be employed for comparison of their distributions. A test verifying the hypothesis of identity of distributions of a feature in two independent populations is the Kolmogorov-Smirnov λ-test [Domański 1990, Greń 1978]. The test requires samples that are fairly large numerically. In the case when certain parameters in the theoretical model are estimated on the basis of empirical data that are used at the same time for the verification of the model, we cannot adopt the assumption of independence of results in the two sets. Consideration of this condition indicates that for the verification of a model tests should be selected such that can verify the identical nature of distributions with no assumption of independence of variables.

For any distributions (not necessarily normal), the test of signs or the test of ranked signs, also known as the Wilcoxon test, can be applied for the purpose [Domański 1990, Greń 1978]. Both of these tests are included in the group on non-parametric tests.

Let us remember that the three above-discussed tests verify the zero hypothesis of equality of distribution functions for variables from two populations $H_0 : F_1(x) = F_2(y)$ against $H_1 : F_1(x) \neq F_2(y)$. Both samples must have the same numerical populations.

In the Kolmogorov-Smirnov test, applied to oat grain porosity data, the value of the test function:

$$\lambda = \sqrt{n}\sup_x \left|F_{n_1}(x) - F_{n_2}(x)\right|, \tag{4}$$

was: λ^0=0,5774, and the critical values were $\lambda_{0.01}$=1.627 and $\lambda_{0.05}$=1.358. In formula (4), F_{n_1} and F_{n_2} denote empirical distribution functions of samples from both distributions. In view of the inequality of $\lambda<\lambda_{0,05}$ and $\lambda<\lambda_{0,05}$, we find that there are no grounds for rejection of the zero hypothesis of the equality of the distributions. The model describes the experimental data well.

In the test of signs, we calculate differences $(x_i - y_i)$ for $i = 1,...,n$ and then count the number of positive differences, designated as r^+, and the number of negative differences . Symbol r is used to denote the lower of the values r^+ and r^-, i.e.:

$$r = \min(r^+, r^-). \tag{5}$$

Statistics of r is here the test function that, with true hypothesis H_0, has a binomial distribution. The left-hand side of the critical area is defined by the equality $P(r \leq r_\alpha) = \alpha$, and r_α is read from suitable Tables (see e.g. Domański [1990] p. 168). Returning to our example, we have: r^+=32, r^-=16, r=min(r^+, r)=16, $r_{0.01}$=15, and therefore $r > r_{0.01}$. Consequently, there are no grounds for rejection of the zero hypothesis of the equality of the distributions. The model describes the experimental data well.

In turn, in the test of ranked signs (Wilcoxon test) we calculate the values of differences $x_i - y_i$, but for further actions their absolute values $|x_i - y_i|$ should be taken. Ranks should be assigned to those absolute values, and then we should calculate separately the sum of ranks of positive differences, designating it as T^+, and the sum of ranks for the negative differences – T^-. The test function is equal to the lower value of those sums, i.e.:

$$T = \min(T^+, T^-). \tag{6}$$

The critical area of the test is determined by the relation: $P(T \leq T_\alpha) = \alpha$. Thre critical values can be read from Tables, e.g. those given in the book by Domański [1990], p. 269.

If the population of a sample is large ($n > 25$), with true H_0 hypothesis concerning the equality of distribution functions we can employ another form of test function that has an asymptotically normal distribution . This test function has the form:

$$u^o = \frac{T - E(T)}{D(T)}, \tag{7}$$

where: $E(T) = \frac{1}{4}n(n+1)$ and $D^2(T) = \frac{1}{24}n(n+1)(2n+1)$ and $D(T) = \sqrt{D^2(T)}$. (8)

If $\left|u^o\right| > u_\alpha$, then H_o is rejected as in the case of the other tests where the test function has normal distribution.

Taking into account the numerical values from the example, where n = 48: we have: T=min(1150.26)=26 ; ET=588; D2T=9506; DT=97.5 u^0=(T-ET)/DT=-0.37 ; $u_{0,01}$= 2,57 , $u_{0,05}$=1,96. With regard to the relation $\left|u^0\right| < u_{0,01}$ and $\left|u^0\right| < u_{0,05}$ there are no grounds for rejection of the zero hypothesis of the equality of the distributions. The model describes the experimental data well.

CONCLUSIONS

Summing up the considerations given in this report one can conclude that in the case of non-linear regression models the efficiency of a model's suitability to experimental data can be verified by means of numerous statistical tests. It can be the t-student test of equality of mean values, or tests of identical distributions of random variables, e.g. the Kolmogorov-Smirnov test, the test of signs, or the Wilcoxon test. The choice of tests depends on a specific situation and on meeting specific assumptions (e.g. the assumption of independence of features or of normal character of distributions, etc.). The conclusion as to the good suitability of a model to experimental data is: it is the more credible, the greater number of tests applied to verify it. In the example of oat grain porosity considered in the study, all the applied tests confirmed the good suitability of the non-linear regression model to experimental data.

REFERENCES

Bochniak A., Wesołowska-Janczarek M., 2006: Problem doboru wspólnej krzywej dla dwóch replikacji na przykładzie procesu kiełkowania ziaren zbóż stymulowanych polem magnetycznym. Inżynieria Rolnicza, 5(80), 39-47.

Bowszys J., Tomczykowski J., 2007: Self-segregation of maize kernels during gravitational discharge from a silo. Teka Komisji Motoryzacji i Energetyki Rolnictwa, V (VII), 38-42.

Dobosz M., 2001: Wspomagana komputerowo statystyczna analiza wyników badań, Akademicka Oficyna Wydawnicza EXIT, Warszawa.

Domański Cz., 1990: Testy statystyczne, Państwowe Wydawnictwo Ekonomiczne, Warszawa.

Chung D.S., Maghirang R.G., Kim Y.S., Kim M.S., 2001: Effects of moisture and fine material on static pressure drops in a bed of grain sorghum and rough rice. Trans. ASAE, 44(2), 331-336.

Greń J., 1978: Statystyka matematyczna. Modele i zadania, Państwowe Wydawnictwo Naukowe, Warszawa.

Gunasekaran S., Jackson C.Y., 1988: Resistance to airflow of grain sorghum. Trans. ASAE, 31(4), 1237-1240.

Jayas D.S., Sokhansanj S., Sosulski F.W., 1991: Resistance of bulk canola seed to airflow in the presence of foreign material. Canadian Agricultural Engineering, 33(1), 47-54.

Kusińska E., 2006: Ocena wpływu prędkości przepływu powietrza i gęstości upakowania ziarna żyta na opór hydrauliczny. Inżynieria Rolnicza, 11(86), 277-284.

Kusińska E., 2007: Wpływ porowatości ziarna owsa na opór przepływu powietrza. Inżynieria Rolnicza, 8(96), 149-155.

Kusińska E., Kizun V., 2006: Wpływ zagęszczenia ziarna owsa i prędkości przepływu powietrza na opór hydrauliczny. Inżynieria Rolnicza, 5(80), 403-410.
Łukaszuk J., 2005: Wstępna ocena wpływu sposobu formowania złoża ziarna pszenicy na opór przepływu powietrza. Acta Agrophysica, 6(3), 709-714.
Łukaszuk J., Molenda M., Horabik J., 2006: Wpływ sposobu formowania złoża pszenicy na opór przepływu powietrza. Acta Agrophysica, 8(4), 881-891.
Molenda M., Łukaszuk J., Horabik J., 2005: Airflow resistance of wheat as affected by grain density and moisture content. Electronic Journal of Polish Agricultural Universities, 8.
Pabis S., 1985: Metodologia i metody nauk empirycznych, PWN, Warszawa.
Powierża L., 1997: Elementy inżynierii systemów. Oficyna Wydawnicza Politechniki Warszawskiej, Warszawa.

TESTOWANIE MATEMATYCZNEGO MODELU POROWATOŚCI ZIARNA

Streszczenie. Wielu badaczy z dziedziny inżynierii rolniczej stosuje modele matematyczne do opisu przebiegu procesów. Po ustaleniu kształtu modelu niezbędne jest jeszcze sprawdzenie jego poprawności w oparciu o dane eksperymentalne [Pabis 1985, Powierża 1997]. W badaniach inżynierskich często stosuje się modele regresyjne. Jako miarę dobroci dopasowania takiego modelu do danych eksperymentalnych powszechnie używa się współczynnika determinacji R^2. Wskaźnik ten znakomicie sprawdza się w przypadku modeli liniowych i nieliniowych lecz wewnętrznie liniowych (tzn. sprowadzalnych do liniowego drogą odpowiedniej transformacji). Jednak dla modeli nieliniowych i wewnętrznie nieliniowych pojawiają się komplikacje [Dobosz 2001]. Na przykład w modelu logistycznym [Bochniak i Wesołowska-Janczarek 2006] współczynnik determinacji R^2 może przyjmować wartości ujemne. Traci on wtedy swoją naturalną interpretację. Niniejsza praca zawiera przegląd kilku występujących w literaturze testów statystycznych, które można zastosować do badania poprawności nieliniowych modeli regresyjnych w oparciu o dane eksperymentalne. Opisywane metody statystyczne zostały zilustrowane na wynikach badań eksperymentalnych dotyczących porowatości ziarna owsa [Kusińska 2007, Bowszys i Tomczykowski].

Słowa kluczowe: ziarno owsa, porowatość, testy statystyczne, dobroć dopasowania modelu do danych empirycznych.