

*Marina I. Putnik-Delić¹, Ivana V. Maksimović¹,
Emilija B. Nikolić-Đorić¹, Nevena M. Nagl²*

¹ Faculty of Agriculture

Trg Dositeja Obradovića 8, 21000 Novi Sad, Serbia

² Institute of Field and Vegetable Crops

Maksima Gorkog 30, 21000 Novi Sad, Serbia

ANALYSES OF STATISTICAL TRANSFORMATIONS OF ROW DATA DESCRIBING FREE PROLINE CONCENTRATION IN SUGAR BEET EXPOSED TO DROUGHT

ABSTRACT: Eleven sugar beet genotypes were tested for their capacity to tolerate drought. Plants were grown in semi-controlled conditions, in the greenhouse, and watered daily. After 90 days, water deficit was imposed by the cessation of watering, while the control plants continued to be watered up to 80% of FWC. Five days later concentration of free proline in leaves was determined. Analysis was done in three replications. Statistical analysis was performed using STATISTICA 9.0, Minitab 15, and R2.11.1. Differences between genotypes were statistically processed by Duncan test. Because of nonnormality of the data distribution and heterogeneity of variances in different groups, two types of transformations of row data were applied. For this type of data more appropriate in eliminating nonnormality was Johnson transformation, as opposed to Box-Cox. Based on the both transformations it may be concluded that in all genotypes except for 10, concentration of free proline differs significantly between treatment (drought) and the control.

KEY WORDS: tolerance, water deficiency, sugar beet, proline, statistical transformation, Box-Cox transformation, Johnson transformation

INTRODUCTION

According to most authors, for the successful production of sugar beet total annual amount of precipitation should be around 600 mm (S p a s i ć, 1989). Approximately 10—20% of the total water requirements of sugar beet are coming from reserves of water in the soil and the rest is provided by rainfall and irrigation. The amount of evaporated water is 392 mm on average and ranges from 198 mm in dry years to 542 mm in rainy years (M a k s i m o v i ć and D r a g o v i ć, 2001).

Lack of water has a very complex effect on physiological processes in plants. Research on the physiological basis of drought tolerance in many cases was the first step in the selection for tolerance to drought. It was found that there is variability in responses to water stress in sugar beet (Sadeghian et al., 2000, Ober and Luterbacher, 2002, Maksimović et al., 2004 and 2006, Perčić et al., 2005, 2006, 2007) and in some putative structural and morphological adaptive features of sugar beet to the lack of water (Luković et al., 2009). If stress occurs during the early stages of growth and development, it slows down the growth of roots, which can reduce yield by 46% (Noghabi and Williams, 2000).

The first effects of water stress are expressed in leaves, one of which is accumulation of osmoregulation substances in sugar beet such as glycine betaine, proline, and fructans. This phenomenon is correlated with decreasing of water potential of cells. It is not clear if plants, which in terms of lack of water build up these substances, better tolerate the lack of water or not (Ghoulam et al., 2002; Maksimović et al., 2004 and 2006; Perčić et al., 2005).

When inhomogeneous results are obtained during the measurement of certain parameters, it is difficult to draw reliable conclusions about the effect of applied treatments without additional statistical analyses of row data. Such example of statistical transformation of core values using Box-Cox and Johnson transformations will be shown in this paper. The effect of those transformations on the final conclusion on the effect of water deficiency on free proline concentration in sugar beet leaves is discussed.

MATERIAL AND METHODS

Eleven sugar beet genotypes (1–11), pre-selected for their different ability to maintain turgor in the field, were grown in semi-controlled conditions in the greenhouse. Substrate was a mixture of soil and sand, and plants were watered daily. After 90 days, water deficit was imposed by the cessation of watering, while the control plants continued to be watered up to 80% of FWC. Five days later the concentration of free proline in leaves was determined following the procedure of Bates (1973). Analyses were carried out in three replications.

In order to examine distribution of experimental results basic statistical exploratory methods box-whisker diagram and plot (Q-Q plot) were applied. Box-whisker diagram was obtained on the base of median (M_c) upper (Q_3) and lower quartiles (Q_1) and interquartile range that is a measure of dispersion of the central portion of a distribution ($I_Q = Q_3 - Q_1$). This diagram is very useful in establishing skewness of the distribution and presence of outlier and extreme values. The outlier is defined as the value outside the range of ($Q_1 - 1.5 \cdot I_Q$, $Q_3 + 1.5 \cdot I_Q$), and the extreme value is the value outside the range of ($Q_1 - 3 \cdot I_Q$, $Q_3 + 3 \cdot I_Q$).

In the Q-Q plot, the observed values of the variable are ordered ($x_1 < \dots < x_n$), and then these values (x_i) are plotted against the inverse probability distribution function (theoretical quantiles). If the observed values fall on the re-

gression line, then it can be concluded that the observed values follow the specified distribution. In our case, specified distribution is normal. The lack of normality was established by means of exploratory analysis and several normality tests based on different approaches. Anderson-Darling and Lilliefors tests are empirical density function (EDF) omnibus tests. Shapiro-Wilk test is based on the squared correlation between the ordered sample values and the (approximated) expected ordered quantiles from the standard normal distribution. Jarque-Bera and D'Agostino omnibus tests measure deviation of empirical skewness and kurtosis from zero.

Applied transformations are preprocessing techniques used to stabilize variance and make the data more normal distribution like. Parametric analysis of transformed data is considered a better strategy than non-parametric analysis because the former appears to be more powerful than the latter (Rasmussen & Dunlap, 1991).

The Box-Cox transformation, introduced by statisticians George E. P. Box and David Cox in 1964, is a family of a power transformations that might be used to convert a general set of n observation observations into a set of n independent observations from a normal distribution with constant variance. The transformation involves a parameter that can be estimated from the data using the method of maximum likelihood:

$$Y = \frac{X - 1}{\ln X}, \quad 0$$

Special cases of Box-Cox transformation are log, square root and inverse transformation. The Johnson system (Johnson, 1949) is a very flexible system for describing statistical distributions as it includes four parameters. It is defined by,

$$z = \gamma + \delta \log[f(u)], \text{ with } u = \frac{x - \bar{x}_i}{s_i}$$

and where function $f()$ has four possible forms depending on original data distribution:

- SL: $f(u) = u$ for the log normal distribution,
- SU: $f(u) = u + \sqrt{u^2 - 1}$ for an unbounded distribution,
- SB: $f(u) = u/(1-u)$ for a bounded distribution,
- SN: $f(u) =$ for e^u the normal distribution.

Three programs Statistica 9, Minitab 15 and R2.11.1 were used to perform statistical analysis.

Exploratory data analysis and normality tests were done by R2.11.1. Johnson transformation was done by Minitab 15 program, and Box-Cox trans-

formation and ANOVA and Duncan test using Statistica 9. Duncan test was applied on original data, and on the both sets of transformed results.

RESULTS AND DISCUSSION

The concentration of free proline showed heterogeneity of variances in genotype and treatment groups according to the tests of homogeneity of variances (Hartley, Cochran, Bartlett, Levene).

On the base of the histogram of row data (mg proline/g DW), it can be concluded that distribution is highly skewed. Normal Q-Q plot and box-plot show deviation from normality, as the right tail of the empirical distribution is heavier than the right tail of normal distribution (first line of Fig. 1). Highly significant deviation from normality was also confirmed by statistical tests (Table 1).

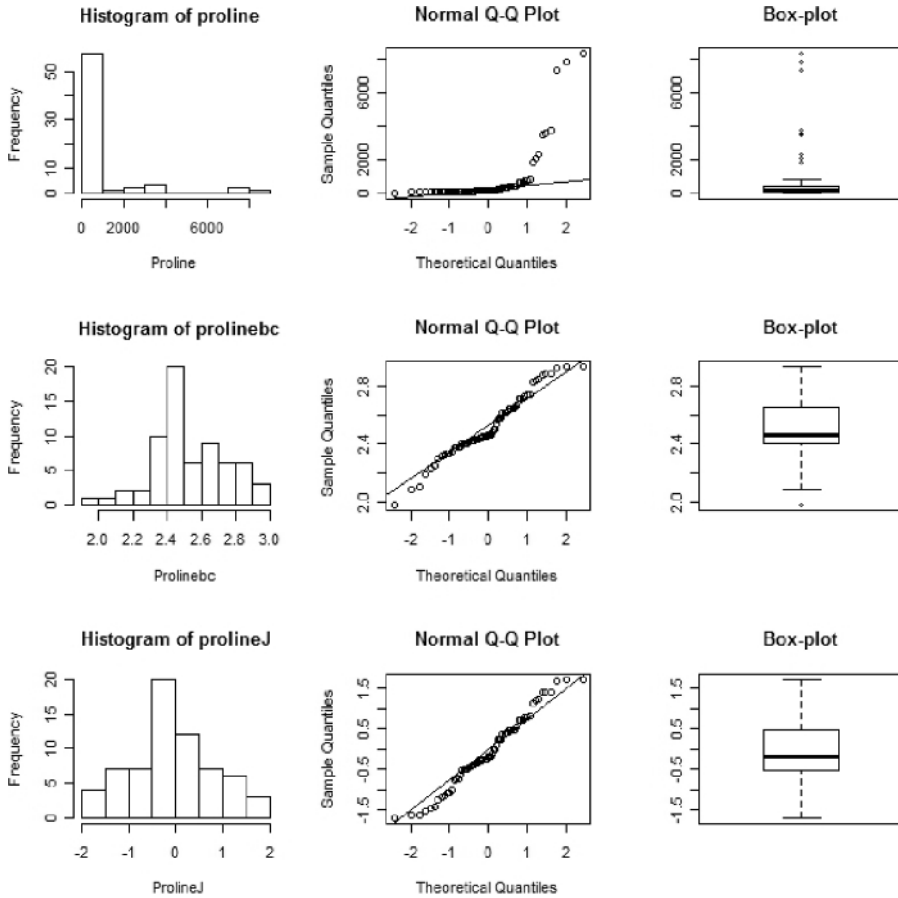


Fig. 1. — Histogram, normal Q-Q plot and box-whisker diagram for the original and transformed data

Tab. 1. — Results of normality tests applied on the raw and transformed data

Normality test	Original data	Box-Cox transformation	Johnson transformation
Anderson-Darling p-value	14.2153 < 2.2e-16	0.8372 0.02932	0.3697 0.4164
Lilliefors (Kolmogorov-Smirnov) p-value	0.3634 < 2.2e-16	0.1349 0.004512	0.0882 0.2292
Shapiro-Wilk p-value	0.4555 3.444e-12	0.9689 0.09585	0.976 0.2290
Jarque-Bera p-value	374.3926 < 2.2e-16	0.2723 0.8727	1.2122 0.5455
D'Agostino p-value	68.0175 1.665e-15	0.1243 0.9397	1.5037 0.4715

In order to apply parametric analysis of variance and treatment comparison, transformation of data was necessary.

The distribution of Box-Cox transformed data with maximum likelihood estimate $\lambda = -0.321864$, is much closer to normal distribution. It may be noticed that its left tail is heavier comparing with normal distribution because of presence of several outliers (second line of Fig. 2). The deviation from normality was confirmed by results of Anderson-Darling and Lilliefors test (Table 1). The transformation eliminated heterogeneity of variances in treatment groups and decreased it in genotype groups.

The best Johnson transformation was SU type and transformation function equals

$$f(u) = 0.911 - 0.3754 \sinh^{-1} \frac{x - 82.2085}{15.1082}, \text{ where } \sinh^{-1}(u) = \log(u + \sqrt{1 + u^2}).$$

The Johnson transformation was effective in eliminating nonnormality and variance stabilization (third line of Fig. 3, Tab. 1).

Data transformation did not affect the overall F test and F test for interaction genotype x treatment (imposed drought) but it affected pairwise comparisons (Fig. 2, 3, 4).

Figures were all made for $\alpha = 0.05$, because the result was exactly the same at $\alpha = 0.01$. Graphical presentation of the sample mean values of free proline in sugar beet leaves and corresponding 95% intervals suggest that in the case of transformed data, there is greater number of significant comparisons.

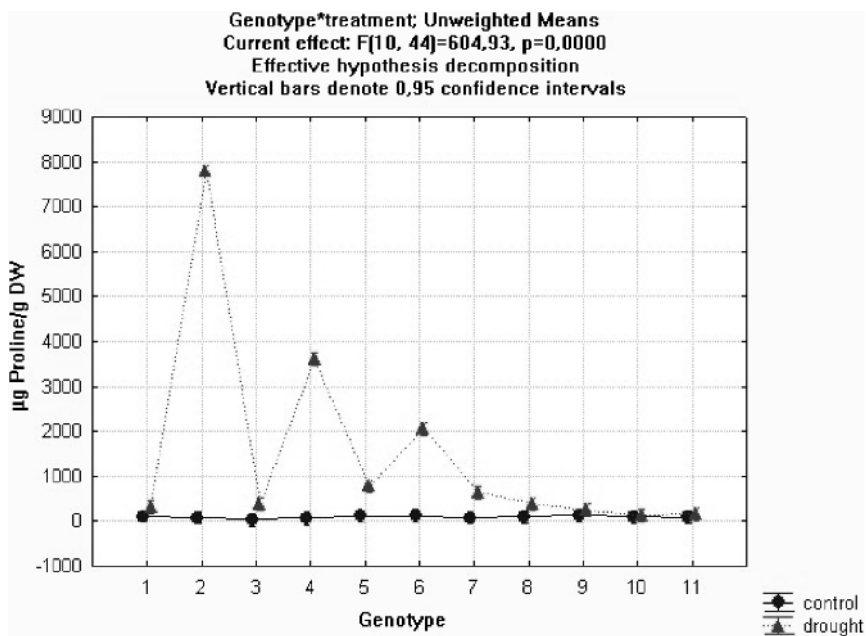


Fig. 2. — The result of analysis of variance for the concentration of free proline in eleven genotypes of sugar beet — the row data

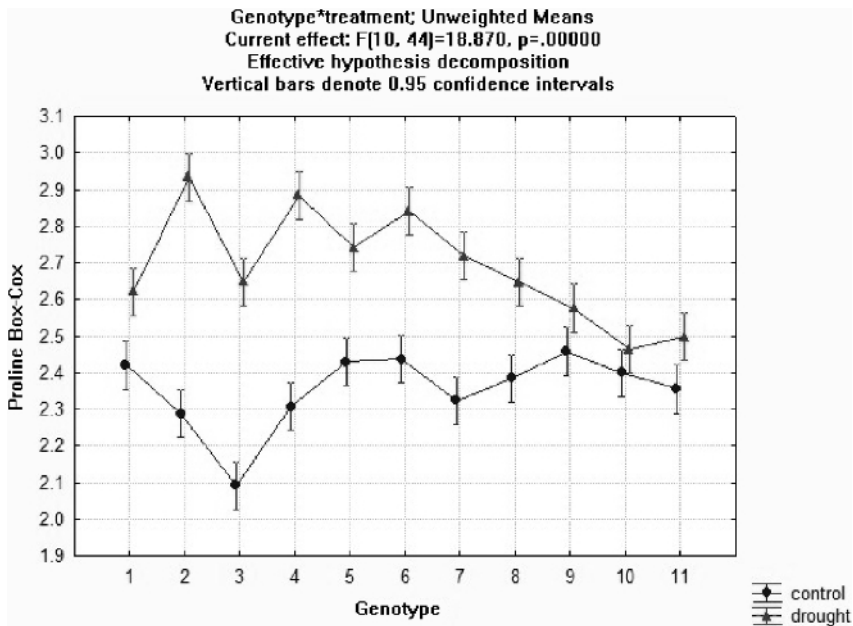


Fig. 3. — The result of analysis of variance for the concentration of free proline in eleven genotypes of sugar beet after Box-Cox transformation of row data

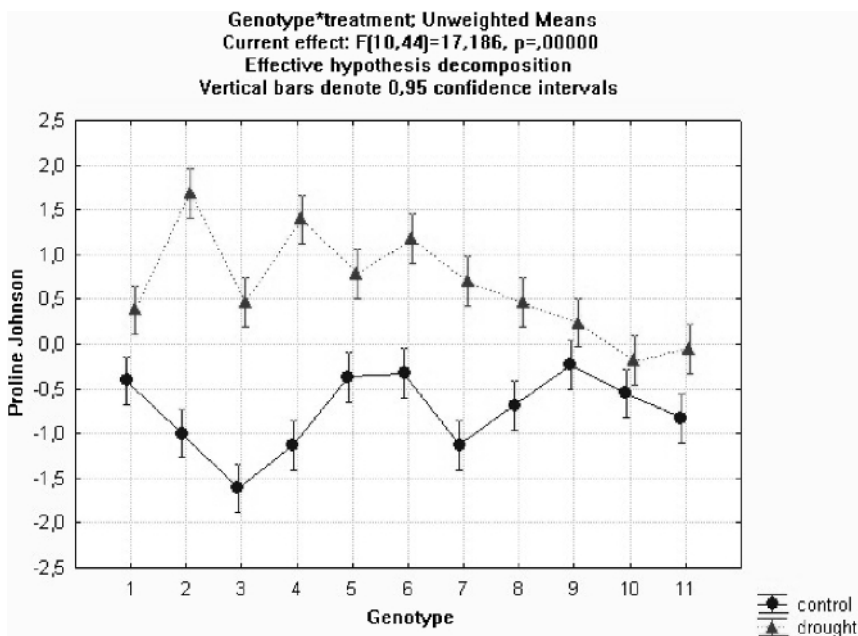


Fig. 4. — The result of analysis of variance for the concentration of free proline in eleven genotypes of sugar beet after Johnson transformation of row data

Tab. 2. — Results of Duncan’s test for free proline concentration in sugar beet leaves, row and transformed data. “c” stands for control, “d” for drought (water deficit in semi control conditions)

Duncan test		1c	1d	2c	2d	3c	3d	4c	4d	5c	5d	6c	6d	7c	7d	8c	8d	9c	9d	10c	10d	11c	
1c																							
1d	●△																						
2c	●△	□●△																					
2d	□●△	□●△	□●△																				
3c	●△	□●△	●△	□●△																			
3d	□●△		□●△	□●△	□●△																		
4c	●△	□●△		□●△	●△	□●△																	
4d	□●△	□●△	□●△	□●△	□●△	□●△	□●△																
5c		●△	●△	□●△	●△	□●△	●△	□●△															
5d	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△														
6c		●△	●△	□●△	●△	□●△	●△	□●△	□●△	□●△													
6d	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△												
7c	△	□●△		□●△	●△	□●△		□●△	●△	□●△	●△	□●△											
7d	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△										
8c		●△		□●△	●△	□●△	△	□●△	□●△	□●△	□●△	△	□●△										
8d	□●△		□●△	□●△	□●△		□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△							
9c		●△	●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△					
9d	●△		●△	□●△	□●△		●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△				
10c		●△	●△	□●△	□●△	□●△	△	□●△		□●△		□●△	△	□●△		□●△		□●△	●△				
10d		●△	●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△			
11c		□●△	□●△	□●△	□●△	□●△	□●△	△	□●△	△	□●△	△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△		
11d		●△	●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△	□●△

□ — row data
 — Box-Cox transformation
 — Johnson transformation

The statistically significant pairwise comparisons based on post-hoc Duncan's test, done on row and on transformed data, are presented in Table 2. The results indicate that in all genotypes except 1, 9, 10 and 11 there was statistically significant differences in free proline concentration between treatment (drought) and control if the analysis is done on the original data. In the case of transformed data, the difference is significant in all genotypes except for 10. If we consider all 231 comparisons presented in Table 2, the 123 comparisons based on original data are significant. For transformed data, there is a greater number of significant comparisons: 173 in the case of Box-Cox and 182 for Johnson transformation.

CONCLUSIONS

The Box-Cox transformation is usually applied in biometrics to eliminate deviation from normality and to stabilize variance. Although this transformation is easy to understand and apply, it often does not find a suitable solution. The Johnson transformation, implemented in statistical software (Minitab, R), may be a good alternative. In this research, results obtained with both, transformations concerning comparison of free proline concentration between treatment (drought) and control are consistent.

REFERENCES

- Bates, L. S. (1973): *Rapid determination of free proline for water-stress studies*. Plant Soil 39: 205—207.
- Box, G. E. P., Cox, D. R. (1964): *An analysis of transformation*. Journal of the Royal Statistical Society. Series B 26: 211—243.
- Ghoulam, C. A., Foursy, A., Fares, K. (2002): *Effects of salt stress on growth, inorganic ions and praline accumulation in relation to osmotic adjustment in five sugar beet cultivars*. Environ. Exp. Bot. 47: 39—50.
- Johnson, N. L. (1949): *Systems of frequency curves generated by methods of translation*. Biometrika 36: 149—176.
- Luković, J., Maksimović, I., Zorić, L., Nagl, N., Perčić, M., Polić, D., Putnik-Delić, M. (2009): *Histological characteristics of sugar beet leaves potentially linked to drought tolerance*. Industrial Crops and Products 30: 281—286.
- Maksimović, I., Čurčić, Ž., Nagl, N., Kovačev, L. (2006): *Changes in petioles and leaf blades in response to progressive drought in young sugar beet plants*. XV FESPB Congress, Federation of European Societies of Plant Biology, Lyon, France, p. 184.
- Maksimović, I., Nagl, N., Malenčić, Đ., Kovačev, L., Petrović, N., Perčić, M., Milić, S. (2004): *Effect of short-term water deficiency on sugar beet genotypes*. The 14th FESPB Congress Book of Abstracts, Acta Physiologiae Plantarum, Vol. 26, No. 3 Supplement, August 23—27, Cracow, Poland, p. 85—86.

- Maksimović, L., Dragović, S. (2001): *Potrebe za vodom i efekat navodnjavanja ratarskih biljaka*. Zbornik radova X Kongresa JDPZ, Beograd, CD.
- Noghabi, A. M., Williams, B. F. (2000): *Drought stress and weed competition in sugar beet*. British Sugar Beet Rev. 1: 47—49.
- Ober, E. S., Luterbacher, M. C. (2002): *Genotypic variation for drought tolerance in Beta vulgaris*. Ann. Bot. 89: 917—924.
- Perčić, M., Luković, J., Maksimović, I. (2006): *Strukturni odgovor tkiva liške šećerne repe na uslove kratkotrajnog vodnog deficita*. Treći simpozijum Sekcije za oplemenjivanje organizama Društva Genetičara Srbije i Četvrti naučno-stručni simpozijum iz selekcije i semenarstva Društva selekcionara i semenara Srbije, Hotel „Palisad” (Zlatibor), 16—20. maja 2006.
- Perčić, M., Maksimović, I., Krstić, L., Luković, J. (2008): *Effect of short-term water deficiency on vascular tissue of petiole and leaf lamina of sugar beet*. In: Mihailovic, Dragutin; Miloradov, Mirjana Vojinovic (eds) *Environmental, Health And Humanity Issues In The Down Danubian Region: Multidisciplinary Approaches*, Proceedings of the 9th International Symposium on Interdisciplinary Regional Research ISBN10: 9812834397, ISBN13: 9789812834393 Pub. Date: 11/9/2008, World Scientific Pub Co Inc.
- Perčić, M., Maksimović, I., Putnik-Delić, M., Nagl, N., Kovačev, L. (2005): *The Effect of Water Deficiency on Some Characteristics of Sugar Beet Leaves*. 7th International Symposium “Young people and multidisciplinary research”, 22—23 September, Resita, Rumania, Zbornik radova objavljen na CD-u, rad br B 33.
- Rasmussen, J. L., Dunlap, W. P. (1991): *Dealing with nonnormal data: Parametric analysis of transformed data vs nonparametric analysis*. Educational & Psychological Measurement, 51: 809—820.
- Sadeghian, S. Y., Fazli, H., Mohammadian, R., Taleghani, D. F., Mesbah, M. (2000): *Genetic variation for drought stress in sugarbeet*. J. Sugar Beet Res. 37: 55—77.
- Spasić, P. (1989): *Planiranje prinosa i kvaliteta šećerne repe pri određenim ekološkim uslovima proizvodnje*. Zbornik radova, sv. 16. Poljoprivredni fakultet, Institut za ratarstvo i povrtarstvo Novi Sad.

ПРИМЕНА СТАТИСТИЧКИХ ТРАНСФОРМАЦИЈА
НА ИЗМЕРЕНЕ КОНЦЕНТРАЦИЈЕ ПРОЛИНА У ЛИСТУ
ШЕЋЕРНЕ РЕПЕ У УСЛОВИМА СУШЕ

Марина И. Путник-Делић¹, Ивана В. Максимовић¹,
Емилија Б. Николић-Ђорић¹, Невена М. Нагл²

- ¹ Пољоприведни факултет, Департман за ратарство и повртарство,
Трг Доситеја Обрадовића 8, 21000 Нови Сад, Србија
² Институт за ратарство и повртарство Нови Сад, Максима Горког 30,
21000 Нови Сад, Србија

Резиме

Анализирана је толерантност 11 генотипова шећерне репе према недостатку воде. Биљке су гајене у полуконтролисаним условима, у стакленику, на супстрату који је био мешавина земље и песка, при чему су биљке свакодневно заливане. После 90 дана, водни дефицит је изазван престанком заливања, док су биљке контролне групе и даље заливане, до 80% ПВК. Пет дана касније утврђена је, у три понављања, концентрација слободног пролина у листовима. Статистичка анализа је извршена коришћењем програма Statistica 9.0, Minitab 15 и R2.11.1. Због велике варијабилности података и одступања од нормалне расподеле анализиран је утицај различитих трансформација експерименталних података на коначан закључак. Разлике између аритметичких средина поређене су применом Данкановог теста. На основу обе примењене трансформације може се закључити да се код свих генотипова, изузев генотипа 10, концентрација слободног пролина у листовима значајно разликује између биљака изложених суши и контроле.