Regression model of valerian root yields in the Forest-Steppe zone of Ukraine depending on fertilization rates and water use of the crop

Pavlo Lykhovyd[1*], Nataliya Lavrenko[2], Iryna Biliaieva[1], Olena Piliarska[1]and Valerii Piliarskyi[1]

[1]Institute of Irrigated Agriculture of NAAS, Naddniprianske, Kherson, 73483, **Ukraine**
[2]Kherson state agrarian university, Stritenska Str. 23, Kherson, 73020, **Ukraine**

The study is devoted to the creation and description of the mathematical model of valerian root productivity formation in the Forest-Steppe zone of Ukraine. The study is based on the perennial (2013-2015) research results of Ukrainian scientists in the field of medicinal plant cultivation in the drip-irrigated conditions. The generalization and statistical analysis of the data was performed, and the multiple linear regression model of valerian root yields was proposed. The results of modeling testify that the effect of mineral fertilizers is greater than the latter of water supply for the crop in the above-mentioned conditions of growing. At the same time, it was proved that the value of irrigation is sufficient, and the crop requires adequate irrigation to form high yields of root mass. The figures of the coefficient of determination for all the models, engaged in the study, fluctuate within 0.96-0.99, testifying about their high descriptive value. The mean absolute percentage errors of the models are within 10-18%, therefore, their performance is good enough for operational planning and practical use in the crop yield prediction depending on the studied agro-technological factors.

Keywords: *Valeriana officinalis* **L., irrigation, multiple regression, mineral fertilizers, yield model.**

INTRODUCTION

Valerian (*Valeriana officinalis* L.) is one of the most known medicinal plants. Its use has begun many centuries ago, and the evidence of its efficient application in medicine could be traced back to the ancient times. Extracts from valerian underground bodies – roots – are widely used in the traditional medicine of many cultures as a sedative, calming, and sleep-promoting agents (Houghton 1999). The popularity of valerian root extract is stable and even increasing in the recent years (especially in European countries and in the USA) due to the understanding that synthetic drugs are usually more harmful than natural aids, so people are looking upon herbal medications as

good alternative for pharmacological products. The World Health Organization reported that nearly 80% of global population uses herbal medicine, while valerian is in the list of the most popular remedies in mild forms of mental and sleep disorders (Cass 2004). Therefore, the demand for valerian root in the nearest future seems not to decline, especially considering significant change in mental health in the recent years and aggravation of mental disorders even in primarily healthy population due to the COVID-19 pandemic (Pfefferbaum and North 2020). Satisfactory response of agricultural sector of the global economy to the demand could be offered only through the implementation of modern

agrotechnologies in medicinal plant growing, including climate-smart (which is extremely relevant in the context of global warming affection on medicinal plants) and precision agriculture achievements (Cavaliere 2009, Campbell et al. 2014, Talebpour et al. 2015). The goal of our study is to offer a simple mathematical tool for the estimation and planning the yield of valerian root in the zone of Left-Bank Forest-Steppe of Ukraine depending on the agro-technological parameters, namely: NPK fertilizers application rates and crop water use (irrigation water, rainfall, and soil moisture reserves).These two factors are both incredibly important in cultivation technology formation, especially for rational use of water resources under the conditions of their gradual depletion (Arnell 1999). The study is based on the results of perennial research of domestic Ukrainian scientists (Pryvedeniuk 2016), who were the first to develop mathematical models for the crop yield estimation by the duration of vegetation period under different irrigation schedules (Pryvedeniuk 2016). Another goal of our research is to provide mathematical proof for the importance of each of the studied agro-technological factors as preconditions for the formation of high sustainable yields of valerian roots.

MATERIALS AND METHODS

Statistical yielding models are one of the prospective ways for the improvement of agricultural planning and better understanding of the underlying processes in crops formation (Vozhehova et al. 2020). The data on the crop yields and water use peculiarities were taken from the results of experimental field studies of Pryvedeniuk (2016), which were conducted in the time span of 2013-2015 years. We offer new insight on these data through thorough statistical processing in order to determine the regularities of valerian productivity formation and provide convenient and simple tool for the crop yield prediction using the studied agro-technological factors as the inputs. Statistical (multiple linear regression) analysis was performed using BioStat v7 add-in for Microsoft Excel 365. The calculations were performed by the standard processing. Every regression model is supported by the most valuable statistical indices such as Pearson's linear correlation coefficient R, coefficient of determination R^2 (simple, adjusted, and predicted), mean absolute percentage error MAPE, standard deviation S, mean square of the error MSE, root-mean-square error RMSE, etc.

Analysis of variance ANOVA was also performed to prove that the studied factors really have significant impact on the yield of the crop (Anscombe 1973, Cook and Weisberg 1982, Neter et al. 1996, Pedhazur 1997, Stevens and Pituch 2016, Huber 2004, Belsley et al. 2005). The developed models were approximated and tested

RESULTS

As a result of the data generalization, we formed two descriptive tables for further statistical processing (Table 1, 2)

Regression analysis of the yielding data allowed us to create two equations, which describe the peculiarities of the yield formation, and ANOVA results testified that the established regularities are significant (zero hypothesis was denied). The results of regression and ANOVA are represented in the Tables 3-4.

The Table 5 represents the equations for valerian roots yield depending on the studied factors (according to the results of the regression analysis).

Thus, according to the results of ANOVA, all the studied factors significantly affect the crop's productivity, zero hypotheses were denied in all cases. Predicted R^2 values within the range of 0.94-0.98 (reaching to 1.0) testify about high performance of both models in valerian root yield prediction, that is additionally supported by the values of MAPE, which are in the range of 10-18% (good forecasting according to Moreno et al. 2013).

The graphical approximation of the models is given in the Figures 1-2.

Thus, according to the results of ANOVA, all the studied factors significantly affect the crop's productivity, zero hypotheses were denied in all cases. Predicted R^2 values within the range of 0.94-0.98 (reaching to 1.0) testify about high performance of both models in valerian root yield prediction, that is additionally supported by the values of MAPE, which are in the range of 10-18% (good forecasting according to Moreno et al. 2013).

The graphical approximation of the models is given in the Figures 1-2.

Models' approximation testifies about slight discrepancies between the modeled and true yields of valerian roots. The better fit is in the Model No. 2. However, both models are suitable for dynamic assessment and prediction of the crop raw material yields in practical and theoretical fields of application.

Table 1: Water use structure and valerian roots yield

Soil moisture reserves + Rainfall, m³/ha (SR)	Irrigation Rates, m³/ha (IR)	Water Use, m³/ha (WU)	Yield, t/ha
2550	0	2550	1.9
3189	0	3189	2.3
2340	2090	4430	3.7
2836	2470	5306	3.6
3898	2850	6748	3.4
2300	2470	4770	4.0
2635	3120	5755	4.1
3801	3380	7181	4.3
2230	2860	5090	5.2
2518	3315	5833	5.0
3718	3965	7683	5.2

Table 2: Valerian roots yield depending on NPK fertilization rates and water use

NPK, kg/ha	WU, m³/ha	Yield, t/ha
0	4770	3.6
0	5755	3.5
0	7181	3.7
90	4770	4.0
90	5755	4.1
90	7181	4.3
90	4770	4.2
90	5755	4.4
90	7181	4.1
0	4770	3.8
0	5755	3.7
0	7181	4.5
90	4770	4.1
90	5755	4.5
90	7181	5.0
90	4770	4.5
90	5755	4.8
90	7181	5.2

Table 3: Regression analysis results of valerian root yields depending on NPK fertilization rates and water use peculiarities

Regression statistics	Regression pairs	
	SR + IR = Yield	NPK + WU = Yield
R	0.9827	0.9913
R^2	0.9658	0.9826
Adjusted R^2	0.9619	0.9815
Predicted R^2	0.9478	0.9779
S	0.8218	0.5872
MSE	0.6754	0.3449
MAPE (%)	17.24	10.87
RMSE	0.9181	0.6222

Table 4: Results of ANOVA statistics of valerian root yields depending on NPK fertilization rates and water use peculiarities

Regression statistics	Regression pairs	
	SR + IR = Yield	**NPK + WU = Yield**
t-statistics	2.9665 – SR; 4.3531 – IR	2.9422 – NPK; 15.5244 – WU
p	0.0158 – SR; 0.0018 – IR	0.0101 – NPK; $1.194{\times}10^{-10}$ – WU
H0 (zero hypothesis)	SR – denied; IR – denied	NPK – denied; WU – denied

Table 5: Mathematical equations of valerian root yields depending on NPK fertilization rates and water use peculiarities

Pair / Model No. & Code	Equation
1) Yield = SR + IR	$0.0006{\times}SR + 0.0009{\times}IR$
2) Yield = NPK + WU	$0.0094{\times}NPK + 0.0006{\times}WU$
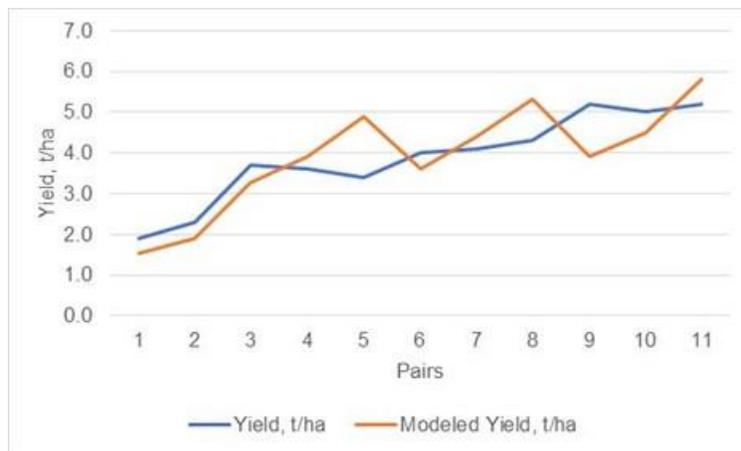
Figure 1: Approximation of the modeled yields of valerian roots (Model No. 1 Yield = SR + IR)

Figure 2: Approximation of the modeled yields of valerian roots (Model No. 2 Yield = NPK + WU)

DISCUSSION

The models of valerian roots yields provided us with an opportunity of deep studying of the yield formation processes under the impact of agro-technological factors. Thus, now we can state that irrigation is of great importance for high valerian root yields in the conditions of the Forest-Steppe zone of Ukraine because soil moisture reserves and rainfall in the zone fail to satisfy the demands of the crop in water (see Model No. 1 coefficients for SR and IR). However, on the other hand, there is an evidence for greater efficiency of NPK fertilization for the crop yield increase (greater crop response per 1 kg/ha of the complex NPK fertilizer than per 1 m^3/ha of the irrigation water supply). Both fertilization and irrigation favor valerian productivity.

Our study is the first one in the field of valerian yield modeling using agro technological factors as the model inputs. However, we must admit that our approach is not something new. Regression models are implemented in agricultural yielding models since the 60s' of the XX century (Miller 1960). Nowadays, agro technological approaches are not the only ones for yield modeling. Sometimes, life factor approach shows better performance (Lykhovyd 2019). Besides, regression models sometimes are outperformed by another method of statistical processing, engaging artificial neural networks (Vozhehova et al. 2019). However, the latter is inferior to regression models in case if we need to obtain a concrete mathematical expression of the model because artificial neural networks provide us just with a pure result and no explanation for how this result was achieved (Vozhehova et al. 2019).

All in all, our study is limited, providing the results, which are relevant just for a particular zone of the crop cultivation, and embracing very few factors, effecting the crop productivity. Therefore, additional field experiments have to be performed and additional modeling studies should be performed to offer the practitioners reliable and comprehensive model of valerian root yields.

CONCLUSION

Regression models of valerian root yield prediction were developed and proved to be reliable for preliminary planning, theoretical research, and practical use. The models are limited for the use in the zone of Ukrainian Forest-Steppe and embrace just two agro-technological inputs – NPK fertilization rates and water use. The models' performance tests testify about their high

accuracy and reliability (predicted R^2 within 0.94-0.98, MAPE within 10-18%). Further research work and modeling studies are required to offer comprehensive valerian root productivity model.

CONFLICT OF INTEREST

The authors declared that present study was performed in absence of any conflict of interest.

AUTHOR CONTRIBUTIONS

PL and IB designed and performleed the study. PL and NL wrote the manuscript. OP performed literature search. VP reviewed and formatted the manuscript. All authors read and approved the final version.

REFERENCES

Anscombe FJ, 1973. Graphs in statistical analysis. The American Statistician 27(1): 17-21.

Arnell NW, 1999. Climate change and global water resources. Global Environmental Change 9: S31-S49.

Campbell BM, Thornton P, Zougmoré R, Van Asten P, Lipper L, 2014. Sustainable intensification: What is its role in climate smart agriculture?. Current Opinion in Environmental Sustainability 8: 39-43.

Cass H, 2004. Herbs for the nervous system: Ginkgo, kava, valerian, passionflower. Seminars in Integrative Medicine 2(2): 82-88.

Cavaliere C, 2009. The effects of climate change on medicinal and aromatic plants. Herbal Gram 81: 44-57.

Cook RD, Weisberg S, 1982. Residuals and Influence in Regression. Chapman and Hall.

New York. USA.

Houghton PJ, 1999. The scientific basis for the reputed activity of Valerian. Journal of Pharmacy and Pharmacology 51(5): 505-512.

Huber PJ, 2004. Robust Statistics. John Wiley & Sons. 523 pp.

Lykhovyd P, 2019. A life factor approach to the yield prediction: a comparison with a technological approach in reliability and accuracy. Journal of Ecological Engineering, 20(6): 177-183.

Miller RJ, 1960. Soybean yield responses and plant composition as affected by phosphorus and potassium fertilizers.

Moreno JJM, Pol AP, Abad AS, Blasco BC, 2013. Using the R-MAPE index as a resistant measure of forecast accuracy. Psicothema 25(4): 500-506.

Neter J, Wasserman W, Kutner MH, 1996. Applied Linear Statistical Models. Irwin, Chicago, US, pp. 241-407.

Pedhazur EJ, 1997. Multiple Regression in Behavioral Research, Ed 3. Harcourt Brace, Orlando, US, pp. 1058.

Pfefferbaum B, North CS, 2020. Mental health and the Covid-19 pandemic. New England Journal of Medicine 383(6): 510-512.

Pryvedeniuk NV, 2016. Valeriana Officinalis yield with a drip irrigation system under conditions of Left-Bank Forest-Steppe of Ukraine. Land Reclamation and Water Economy 104: 72-76.

Pryvedeniuk NV, 2016. Water use of valeriana officinalis under drip irrigation in the conditions of Left-Bank Forest-Steppe of Ukraine. Irrigated Agriculture 65: 113-116.

Stevens JP, Pituch KA, 2016. Applied Multivariate Statistics for the Social Sciences, Ed 6. Routledge, New York, US, pp. 65-129.

Talebpour B, Türker U, Yegül U, 2015. The role of precision agriculture in the promotion of food security. International Journal of Agricultural and Food Research 4(1): 1-23.

Vozhehova R, Kokovikhin S, Lykhovyd P, Balashova H, Lavrynenko Yu, Biliaieva I, Markovska O, 2020. Statistical yielding models of some irrigated vegetable crops in dependence on water use and heat supply. Journal of Water and Land Development 45: 190-197.

Vozhehova RA, Lykhovyd PV, Kokovikhin SV, Biliaieva IM, Markovska OYe, Lavrenko SO, Rudik OL, 2019. Artificial neural networks and their implementation in agricultural science and practice. Warszawa: Diamond trading Tour, pp. 108.

Vozhehova RA, Lykhovyd PV, Lavrenko SO, Kokovikhin SV, Lavrenko NM, Marchenko TYu, Sydyakina OV, Hlushko TV, Nesterchuk VV, 2019. Artificial neural network use for sweet corn water consumption prediction depending on cultivation technology peculiarities. Research Journal of Pharmaceutical, Biological and Chemical Sciences 10(1): 354-358.