

Modeling of the phenolic compounds and antioxidant activity of blueberries by artificial neural networks for data mining

Raquel P. F. Guiné - raquelguine@esav.ipv.pt

CI&DETS/ESAV, Instituto Politécnico de Viseu, Portugal.

Susana Matos - su_matos_9352@hotmail.com

Dep. Indústrias Alimentares, ESAV, Instituto Politécnico de Viseu, Portugal.

Fernando Gonçalves - goncalves7ster@gmail.com

CI&DETS/ESAV, Instituto Politécnico de Viseu, Portugal.

Daniela V. T. A. Costa - daniela@esav.ipv.pt

Dep. Ecologia e Agricultura Sustentável, ESAV, Instituto Politécnico de Viseu

Mateus Mendes - mmendes@isr.uc.pt

Instituto Politécnico de Coimbra, ESTGOH, Oliveira do Hospital, Portugal

Abstract

The present work's goal was to evaluate the effect of different production and conservation conditions, as well as extraction procedures on the phenolic compounds and antioxidant activity of blueberries from cultivar Bluecrop. The production factors considered were origin, altitude of the farm location and age of the bushes, and the conservation was under freezing as opposed to the fresh product. The extraction procedures included two different solvents and different orders of the extraction. Data analysis was performed by training artificial neural networks to model the data and extract information from the model. The results showed that the type of extract and the order of extraction influence the concentrations of phenolic compounds as well as the antioxidant activity of the different samples studied. As to the origin of the farms from where the blueberries were collected, it was found to significantly influence these properties, so that the blueberries from Oliveira do Hospital showed less phenolic compounds and lower antioxidant activity. Older bushes at higher altitudes seem to produce richer berries. Regarding conservation, no influence was observed for phenols but a slight influence could be detected for antioxidant activity.

Keywords

Antioxidant activity, ANN, extraction, production conditions, storage, total phenols.

Modeling of the phenolic compounds and antioxidant activity of blueberries by artificial neural networks for data mining

Conference Topic - TS1 - Food Engineering and Properties

1. Introduction

Blueberries (*Vaccinium* spp) are nowadays widely consumed around the world due to their high amounts of phenolic compounds, among which stand anthocyanins, flavonols, chlorogenic acid and procyanidins [1]. Blueberries have demonstrated a diversity of bioactivities such as antidiabetic, antimicrobial, antiproliferative, memory loss preventive, urinary disease protecting, apoptotic, liver protective, lowering blood cholesterol, lifespan-prolonging, anti-inflammatory, preventing vision problems, delaying ageing, cancer preventive, atherosclerosis preventive and cardioprotective properties [2-7]. Besides these biological activities, it is known that the strong antioxidant activity is one of the characteristics that have contributed for the popularity of this small berry in the past decades [8-10]. Blueberries have in fact been widely used as sources of natural antioxidants and in a variety of applications for preventing diseases of the nervous system and chronic disorders [11-12]. Because of the various health benefits associated to blueberry consumption, besides their unique taste, and nutritional value, the worldwide production and consumption have increased rapidly in the past years, so that they have become the second most important soft fruit species after strawberry [13].

However, blueberries are highly perishable and susceptible to rapid spoilage [14]. Their postharvest life is very short when they are kept at ambient temperature due to the susceptibility to microbial spoilage, mechanical damage or loss of moisture and nutrients [15]. According to Duan et al. [16] fresh blueberries have a shelf life between 1 and 8 weeks, which is greatly dependent on fruit ripeness stage, method of harvest, presence of fruit disease, and storage conditions. The transfer of the fruits from producer to final consumer generally requires 5 to 10 days, and, hence, cold storage must be used to maintain quality [10].

Zhou et al. [17] demonstrated that low-temperature storage (including freezing) can delay senescence as well as help to preserve fruit quality. Hence this procedure may be used to successfully extend the postharvest life of blueberry. Nevertheless, the authors notice that fruit quality can gradually decline during long term low-temperature storage. Furthermore, the quality can be greatly compromised when the stored fruits are subsequently moved to shelf life conditions at ambient temperature.

Many factors affect the blueberries' production and yield, as well as the fruits' characteristics. Irrigation, for example, is a determinant factor in the growth and production of blueberry. This is because the root system of this plant is superficial, quite confined, with a small amount of root hairs, and all these contribute for to the restriction in the water uptake capacity [18]. The blueberry is very dependent on a low temperature environment, this being a critical factor that limits its growth and survival and, hence, conditioning its geographical distribution [19-20].

Artificial neural networks (ANN) have been used in the past for modelling many processes in food engineering and food chemistry. Guiné et al. [21] produced an ANN model to predict the antioxidant activity and the content of phenolic compounds in bananas submitted to different drying treatments. Cabrera and Prieto [22] applied ANN modeling to describe the antioxidant activity in essential oils. Xi et al. [23] used ANN to model the ultrahigh pressure extraction of green tea polyphenols. Vats and Negi [24] optimized by ANN modeling the release of polyphenols and sugars from pine fallen foliage, in the development of bioprocesses. Gutés et al. [25] applied artificial neural network analysis for the determination of phenolic compounds.

The present work's goal was to model by artificial neural networks the effect of different production and conservation conditions, as well as extraction procedures, on the phenolic compounds and antioxidant activity of blueberries from cultivar Bluecrop. The production factors considered were origin, altitude of the farm location and age of the bushes, and the conservation was under freezing as opposed to the fresh product. The extraction procedures included two different solvents and different orders of the extraction.

2. Experimental Procedure

2.1. Harvesting and conservation

To carry out this study were used blueberries *Vaccinium corymbosum* L., from cultivar Bluecrop grown in conventional production mode in five different farms in Portugal, as shown in Table 1. The fruits were manually harvested in a maturation state corresponding to commercialization. The analyses were made right after harvest in the fresh fruits and also after 6 months of storage under freezing at a temperature of -20 °C.

Table 1 - Production conditions of the blueberries.

Location of the farm	Altitude of the farm (m)	Age of the bushes (years)
Sever do Vouga	463	20
Estarreja	45	3
Oliveira do Hospital	510	21
Vouzela	450	3
Braga	125	28

2.2. Chemical analysis

The phenolic compounds were extracted with methanol and with acetone solutions, and in each case the sample was left for 1 hour in an ultrasonic bath at room temperature. The extracts obtained were then used to quantify the phenolic composition and the antioxidant activity. The total phenolic content in the fruit extracts was determined by the Folin-Ciocalteu method according to Gonçalves et al. [26]. The antioxidant activity was determined by DPPH [27] method.

2.3. Modeling by ANN

ANNs are an interconnected association of artificial neurons. The input for each neuron can be one or more variables with their corresponding weights, and the output is a function of the weighted inputs. Neurons learn by adjusting the weights of the input variables, in such a way that the error between the neuron's expected output and the measured output value is minimized. The weights learnt by the neural network during training are an indication of the relevance of each variable for the whole system. The analysis of the weights gives a wealth of information about the contribution of each input to the final output.

2.3.1 Data encoding and modeling

The ANN used was a feed-forward model, created using Matlab™¹ fitnet function, which is designed for data fitting of datasets such as the ones used in the present work. Sigmoidal transfer function, which is a universal approximator, was used in the hidden layer, with a linear transfer function in the output layer. Training is based on Levenberg-Marquardt method

¹Matlab is a registered trademark of Mathworks. www.mathworks.com.

and the Mean Squared Error (MSE) method for performance assessment, as in fitnet function default configuration.

Since neural networks use inductive reasoning, in general more samples produce better models and more confidence in the result. In the present work, the dataset contained 121 experimental results. Table 2 summarises the input variables, which were split into a total of 12 variables for better neural network fitting and analysis. Age and altitude of the farm were floating point numbers normalised in the interval [0, 1]. L_i , S_i and E_i were Boolean values.

Table 2. Encoding of the input variables.

Input #	Input Variable	Description
1-5	L1-L5	Location (Sever do Vouga, Estarreja, Oliveira do Hospital, Vouzela, Braga)
6	A	Age of the bushes [0, 1]
7	H	Altitude of the farm [0, 1]
8, 9	S1, S2	State (Fresh, Frozen)
10, 11	E1, E2	Type of extract used (Methanol, Acetone)
12	O	Order of the extract (First, Second)

To facilitate the process of training the neural networks and analysis of the results, each output variable was processed separately. Each network had just one hidden layer with one input neuron and one output neuron. For each run, the Matlab script randomly selected approximately 70 % of the samples for the train subset, 15 % for the validation subset and the remainder samples were used for the test subset.

3. Results and Discussion

3.1. Total phenolic compounds

Figure 1 shows the total phenolic compounds evaluated in the blueberries from Sever do Vouga, according to type of solvent used for the extraction, order of extraction and state (fresh as opposed to frozen for a period of 6 months). The results show that the first extraction with methanol is by far the most effective in recovering phenolic compounds from the sample, accounting for 59% of all the phenols extracted. The following extraction with methanol is not so effective, allowing obtaining just 11% of the total phenols extracted. The first extraction with acetone is again more effective than the second, but comparing to methanol it is much lower, 21% and 9%, respectively for the first and second extractions. This reveals that the phenolic compounds present in the blueberries were mainly polar, and hence their higher affinity for the polar solvent (methanol).

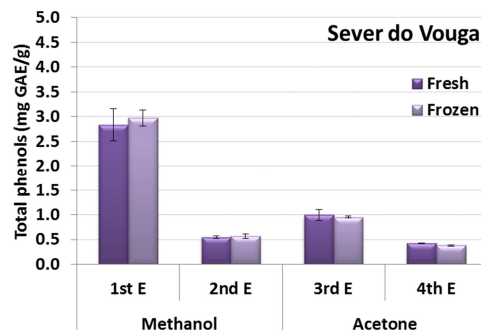


Figure 1 - Phenolic compounds by extract in the blueberries from Sever do Vouga.

Figure 2 shows the same results for the samples collected in Oliveira do Hospital, Vouzela, Estarreja and Braga. The first extraction with methanol always accounts for the highest percentage of phenolic compounds, for all sample origins and both in the fresh and frozen

blueberries. Following comes the 3rd extract, which is the first with the solvent acetone. The relative importance of the 2nd and 4th extracts is low, for these extract a very low amount of phenolic compounds. As to the different origins, it stands out the sample from Oliveira do hospital, with the lowest quantity of phenols, as comparing to all others.

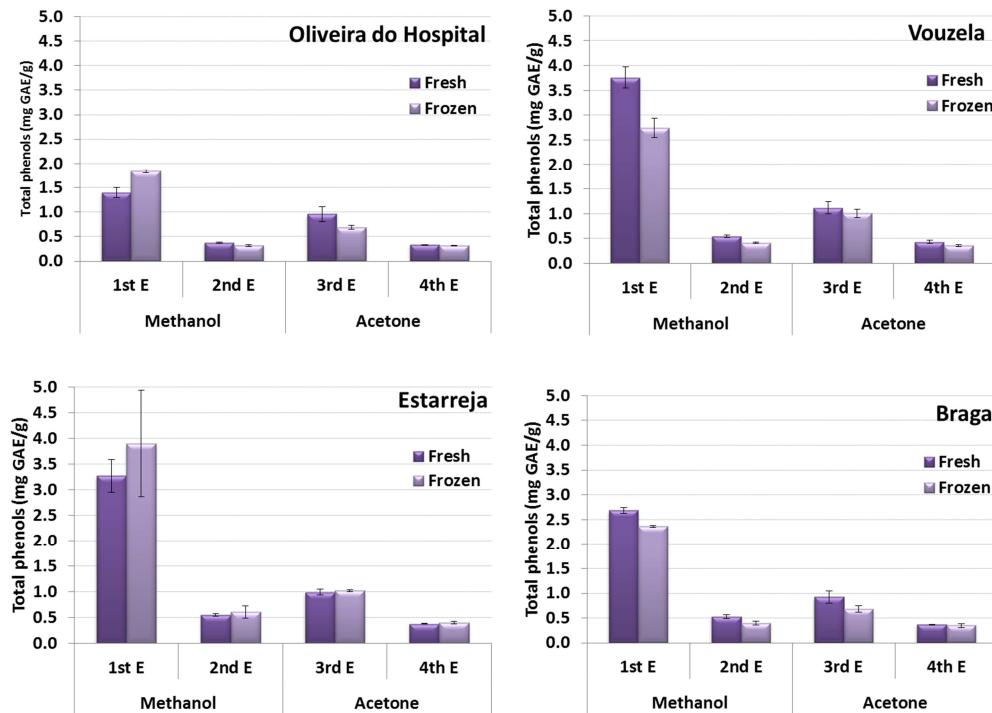


Figure 2 - Phenolic compounds by extract in the blueberries from Oliveira do Hospital, Vouzela, Estarreja and Braga.

Figure 3 presents for all samples the sum of the phenolic compounds extracted with methanol or with acetone, and then the global value (which includes both solvents). The graphs show that methanol is more efficient in extracting the phenols from the blueberries, accounting for more than twice those extracted with acetone, and this trend is valid for both samples, fresh or frozen.

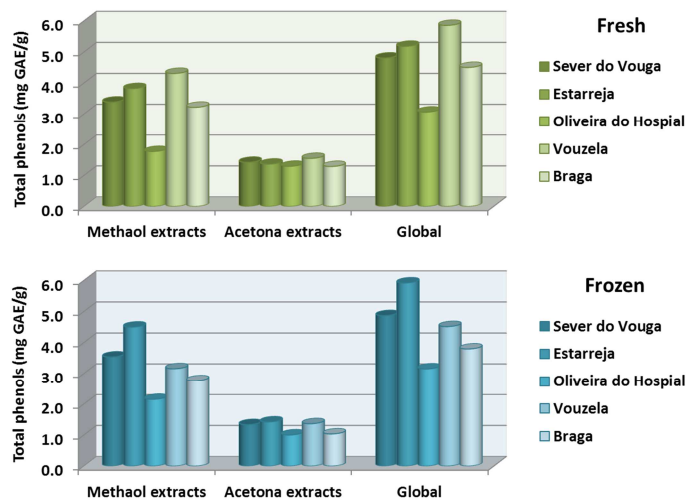


Figure 3 - Total Phenolic compounds in the blueberries for all origins.

3.2. Antioxidant activity

The graphs in Figure 4 present the antioxidant activity determined by DPPH method. The results show that the antioxidant activity is higher in the first extract with methanol for all sample origins, followed by that in the 3rd extract (the first with acetone). This result is in accordance with that of the phenolic compounds, since it is expected that phenols have a large contribution for the antioxidant activity. Again the blueberries from Oliveira do Hospital show the lowest antioxidant activity quantified from all origins.

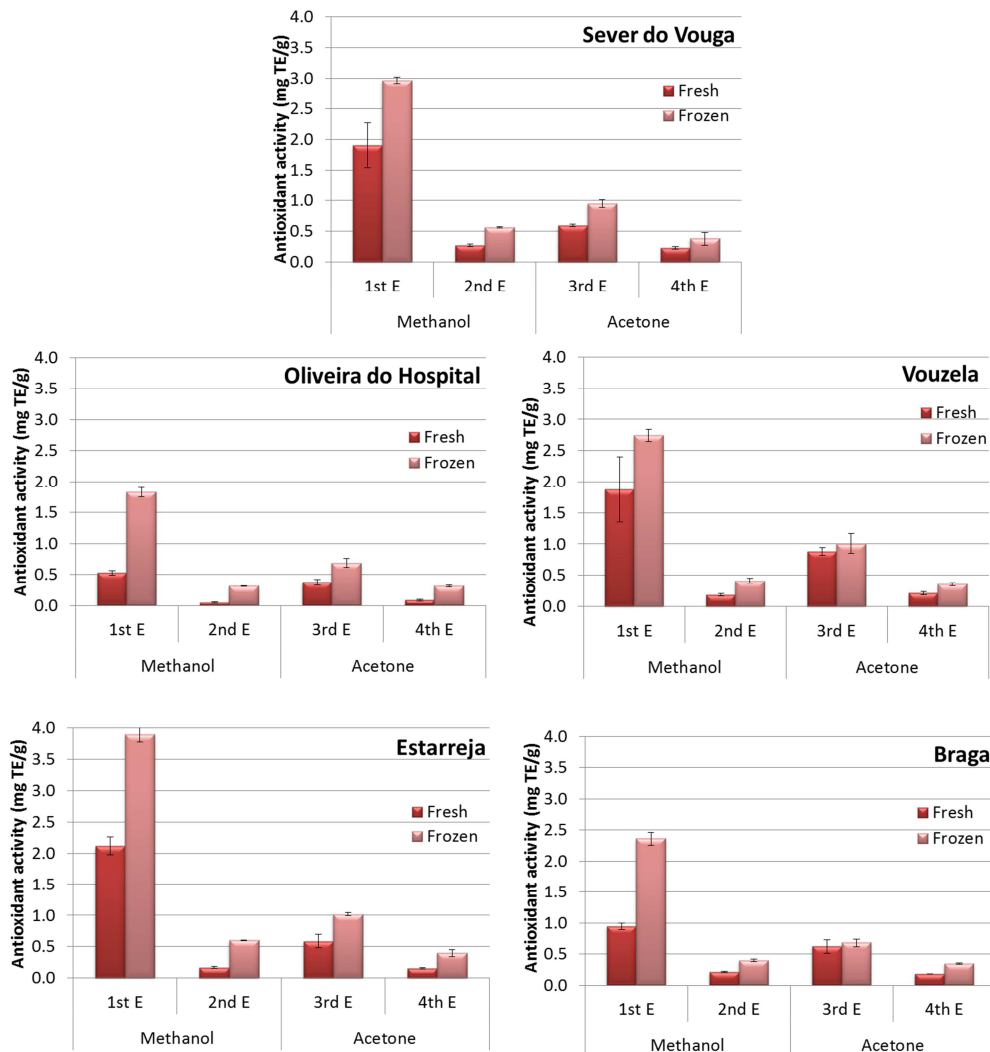


Figure 4 - Antioxidant activity (DPPH) in the blueberries quantified in the different extracts.

The results for the antioxidant activity in Figure 5 confirm that the methanol extracts exhibit higher antioxidant capacity as compared to the acetone extracts, both in fresh and after freezing. Furthermore, a consistent trend was observed for the antioxidant activity to be higher when the blueberries were kept frozen for a period of 6 months than the fresh sample.

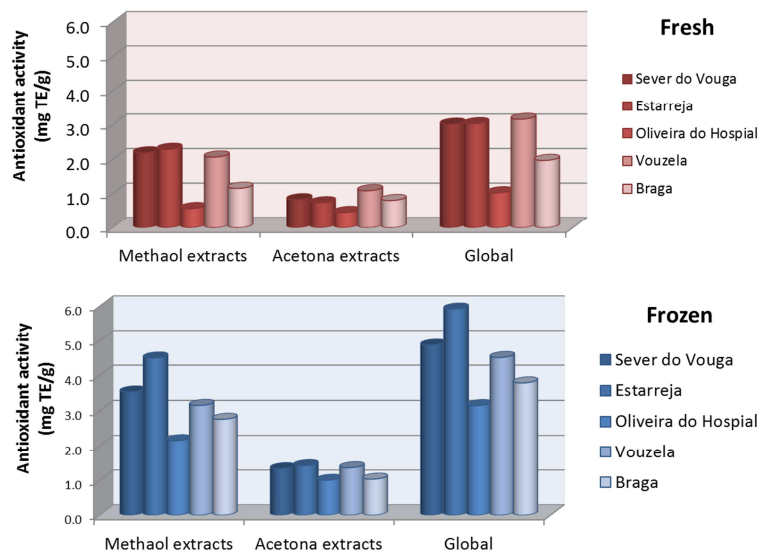


Figure 5 - Antioxidant activity (DPPH) in the blueberries from different origins.

3.3. ANN modeling

The neural networks used learn by adjusting weights and bias of connections, so those weights and bias reflect the contribution of the inputs to predict the output. The bias, being constants, are not relevant to the present analysis.

3.3.1 Experimental results

Many different experiments were performed for each output variable. The result of the learning process depends on some random values, such as the initial weights and bias of each neural connection and the samples selected for the test and validation sets. Therefore, the final results obtained can be different between experiments. The best neural network models will produce the best correlations between values predicted by the neural network and the experimental values measured in the laboratory. For better confidence in the analysis, different experiments were performed in the present study, and the results shown are the average of five selected experiments. Training each model in Matlab took, on average, less than 16 epochs, and required, on average, less than 0.3 seconds, as measured using Matlab tic - toc commands.

Table 3 summarizes some performance parameters of the models. The correlation factors R are greater than 0.98 for the whole dataset, and the Mean Squared Errors (MSE) are very low. Therefore, the models are in general very reliable for the dataset.

Table 3 - Performance of the ANN models used to predict FT and DPPH, assessed as correlation factor R and MSE.

		<i>FT</i>	<i>DPPH</i>	
Model performance measures	R	Train subset	0.989	0.991
		Test subset	0.987	0.935
		All dataset	0.988	0.995
	MSE	0.060	0.023	

Table 4 shows the weights measured for each input variable, normalized in the interval [-1, 1] to facilitate the analysis. The sign of a weight is important to show whether the corresponding variable influences the result positively or negatively. The absolute value

predicts how important the variable is in the context. Variables with low absolute values will be almost irrelevant to the output.

Table 4 - Input weights for each variable.

		<i>FT</i>	<i>DPPH</i>	
Input weight for each input (average of 5)	Location	1 - Sever do Vouga	0.060	0.067
		2 - Estarreja	0.256	0.255
		3 - Oliveira Hospital	0.099	0.121
		4 - Vouzela	0.371	0.403
		5 - Braga	0.333	0.312
	Origin	Bushes' age	0.244	0.244
		Altitude	0.244	0.246
	State	Fresh	0.253	0.203
		Frozen	0.286	0.343
	Extraction	Methanol	0.494	0.451
		Acetone	0.059	0.069
		Extract order	-1.000	-0.926

3.3.2 Analysis of the input neurons' weights

Figure 6 charts the input weights as shown in Table 4. It is clear from the chart that both FT and DPPH behave in a similar way. The inputs that are important for one variable affect the other in a very similar proportion.

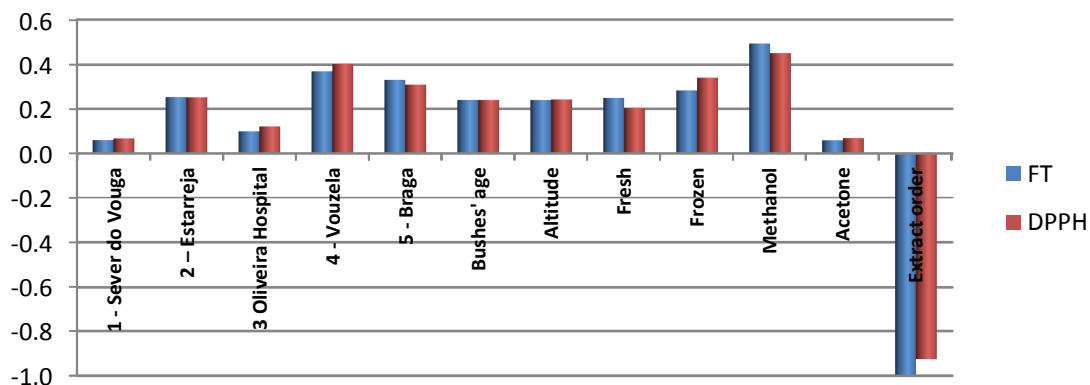


Figure 6 - Input weights for each variable.

Not surprisingly, the order of the extract is the most important predictor. The second extraction always retrieves much less phenolic contents or antioxidants than the first extraction, and the neural weights reflect that by showing a large negative weight to the order of the extract. The type of solvent used is the second best predictor. Methanol has a normalized weight above 0.45 for both variables, showing that when methanol is used the amount of phenolic contents and antioxidants extracted will be higher, while the use of acetone shows an almost negligible contribution.

The chart also illustrates very clearly the importance of the farm location. Vouzela and Braga, followed by Estarreja, were given significant weights, always above 0.25. On the other hand, the models are almost totally indifferent to Oliveira do Hospital and Sever do Vouga, which receive weights less than 0.13 for both variables.

The role of altitude of the farm is also illustrated, as well as the bushes' ages. Both variables seem to play marginal roles, but they have an equal contribution to the amounts of FT and DPPH measured. Since altitude, in meters, and age, in years, are mapped in the interval [0, 1] in the input, the models point that older bushes at higher altitudes are most likely to produce berries richer in phenolic compounds and antioxidants.

As for the conservation methods, the models concur that freezing should preserve at least more DPPH - the latter is more likely to be found in frozen samples than in fresh samples. The state Frozen is actually the fourth most important predictor for DPPH, and fifth for FT.

4. Conclusions

The amount of phenolic contents and antioxidants found in different samples of blueberries was measured, using different methods, and the results were studied using regular statistics methods and neural network models.

The type of extract and the order of extraction influence the phenolic compounds and the antioxidant activity. Also the cultivation place influences these properties, so that the blueberries from Oliveira do Hospital showed lower phenolic contents and antioxidant activity. There also seems to be a small correlation showing that older bushes at higher altitudes produce berries richer in phenols and antioxidants. With regard to conservation, no influence was observed for phenols but a slight influence could be detected for antioxidant activity.

References

- [1] Koca, I., Karadeniz, B.: "Antioxidant properties of blackberry and blueberry fruits grown in the Black Sea Region of Turkey" *Scientia Horticulturae*, Vol. 121 (2009), pp. 447-450.
- [2] Chen, H., Cao, S., Fang, X., Mu, H., Yang H., Wang, X., Xu, Q., Ga, H.: "Changes in fruit firmness, cell wall composition and cell wall degrading enzymes in postharvest blueberries during storage" *Scientia Horticulturae*, Vol. 188 (2015), pp. 44-48.
- [3] Prior, R.L., Wu, X., Gu, L., Hager, T., Hager, A., Wilkes, S., Howard, L.: "Purified berry anthocyanins but not whole berries normalize lipid parameters in mice fed an obesogenic high fat diet" *Molecular Nutrition & Food Research*, Vol. 53 (2009), pp. 1406-1418.
- [4] Adams, L. S., Phung, S., Yee, N., Seeram, N. P., Li, L., Chen, S.: "Blueberry phytochemicals inhibit growth and metastatic potential of MDA-MB-231 breast cancer cells through modulation of the phosphatidylinositol 3-kinase pathway" *Cancer Research*, Vol. 70 (2010), pp. 3594-3605.
- [5] Wu, X., Kang, J., Xie, C., Burris, R., Ferguson, M. E., Badger, T. M., Nagarajan, S.: "Dietary blueberries attenuate atherosclerosis in apolipoprotein E-deficient mice by upregulating antioxidant enzyme expression" *Journal of Nutrition*, Vol. 140 (2010), pp. 1628-1632.
- [6] Shi, J., Pan, Z., McHugh, T. H., Wood, D., Hirschberg, E., Olson, D.: "Drying and quality characteristics of fresh and sugar-infused blueberries dried with infrared radiation heating" *LWT - Food Science and Technology*, Vol. 41 (2008), pp. 1962-1972
- [7] Zielinska, M., Sadowski, P., Błaszczak, W.: "Freezing/thawing and microwave-assisted drying of blueberries (*Vaccinium corymbosum* L.)" *LWT - Food Science and Technology*, Vol. (2015), pp. 555-563.
- [8] Wu, X., Beecher, G. R., Holden, J. M., Haytowitz, D. B., Gebhardt, S. E., Prior, R.L.: "Lipophilic and hydrophilic antioxidant capacities of common foods in the United States" *Journal of Agricultural and Food Chemistry*, Vol. 52 (2004), pp. 4026-4037.
- [9] Wolfe, K. L., Kang, X. M., He, X. J., Dong, M., Zhang, Q. Y., Liu, R. H.: "Cellular antioxidant activity of common fruits" *Journal of Agriculture and Food Chemistry*, Vol. 56 (2008), pp. 8418-8426.
- [10] Zhou, Q., Zhang, C., Cheng, S., Wei, B., Liu, X., Ji, S.: "Changes in energy metabolism accompanying pitting in blueberries stored at low temperature" *Food Chemistry*, Vol. 164 (2014), pp. 493-501.
- [11] Youdim, K. A., Shukitt-Hale, B., Martin, A., Wang, H., Denisova, N., Bickford, P.C.: "Short-term dietary supplementation of blueberry polyphenolics: Beneficial effects on aging brain performance and peripheral tissue function" *Nutritional Neuroscience*, 3 (2000), pp. 383-397.

- [12] Stoner, G. D., Wang, L. S., Casto, B. C.: "Laboratory and clinical studies of cancer chemoprevention by antioxidants in berries" *Carcinogenesis*, 29 (2008), pp. 1665-1674.
- [13] Giongo, L., Poncetta, P., Loretto, P., Costa, F.: "Texture profiling of blueberries (*Vaccinium* spp.) during fruit development, ripening and storage" *Postharvest Biology and Technology*, 76 (2013), pp. 34-39.
- [14] Cantina, C. M., Minas, I. S., Goulas, V., Jiménez, M., Manganaris, G. A., Michailides, T. J., Crisosto, C. H.: "Sulfur dioxide fumigation alone or in combination with CO₂-enriched atmosphere extends the market life of highbush blueberry fruit" *Postharvest Biology and Technology*, 67 (2012), pp. 84-91.
- [15] Hancock, J., Callow, P., Serce, S., Hanson, E., Beaudry, R.: "Effect of cultivar, controlled atmosphere storage, and fruit ripeness on the long-term storage of highbush blueberries" *Horticulture Technology*, 18 (2008), pp. 199-205.
- [16] Duan, J., Wu, R., Strik, B. C., Zhao, Y.: "Effect of edible coatings on the quality of fresh blueberries (Duke and Elliott) under commercial storage conditions" *Postharvest Biology and Technology*, 59 (2011), pp. 71-79.
- [17] Zhou, Q., Ma, C., Cheng, S. C., Wei, B. D., Liu, X. Y., Ji, S. J.: "Changes in antioxidative metabolism accompanying pitting development in stored blueberry fruit" *Postharvest Biology and Technology*, 88 (2014), pp. 88-95.
- [18] Holzapfel, E. A., Hepp, R. F., Mariño, M. A.: "Effect of irrigation on fruit production in blueberry" *Agricultural Water Management*, Vol. 67 (2004), pp. 173-184.
- [19] Rowland, L. J., Hancock, J. F., Bassil, N. V.: "Blueberry" in K.M. Folta, C. Kole (Eds.), *Genetics, Genomics and Breeding of Berries*, pp. 1-40, Science Publishers, Enfield, NH, USA, 2011.
- [20] Die, J. V., Rowland, L. J.: "Elucidating cold acclimation pathway in blueberry by transcriptome profiling" *Environmental and Experimental Botany*, Vol. 106 (2014), pp. 87-98.
- [21] Guiné, R. P. F., Barroca, M. J., Gonçalves, F. J., Alves, M., Oliveira, S., Mendes, M.: "Artificial neural network modelling of the antioxidant activity and phenolic compounds of bananas submitted to different drying treatments" *Food Chemistry*, Vol. 168 (2015), pp. 454-459.
- [22] Cabrera, A. C., Prieto, J. M.: "Application of artificial neural networks to the prediction of the antioxidant activity of essential oils in two experimental in vitro models" *Food Chemistry*, Vol. 118 (2010), pp. 141-146.
- [23] Xi, J., Xue, Y., Xu, Y., Shen, Y.: "Artificial neural network modeling and optimization of ultrahigh pressure extraction of green tea polyphenols" *Food Chemistry*, Vol. 141 (2013), pp. 320-326.
- [24] Vats, S., Negi, S.: "Use of artificial neural network (ANN) for the development of bioprocess using *Pinus roxburghii* fallen foliage for the release of polyphenols and reducing sugars" *Bioresource Technology*, Vol. 140 (2013), pp. 392-398.
- [25] Gutes, A., Céspedes, F., Alegret, S., del Valle, M.: "Determination of phenolic compounds by a polyphenol oxidase amperometric biosensor and artificial neural network analysis" *Biosensors and Bioelectronics*, Vol. 20 (2005), pp. 1668-1673.
- [26] Gonçalves, F. J., Rocha, S. M. Coimbra, M. A.: "Study of the retention capacity of anthocyanins by wine polymeric material" *Food Chemistry*, Vol. 134 (2012), pp. 957-963.
- [27] Brand-Williams, W., Cuvelier, M. E., Berset, C.: "Use of a free radical method to evaluate antioxidant activity" *LWT - Food Science and Technology*, Vol. 28 (1995), pp. 25-30.