## ARTIFICIAL NEURAL NETWORK MODELLING FOR PREDICTION



## SAMPLES BASED ON NIR SPECTRA



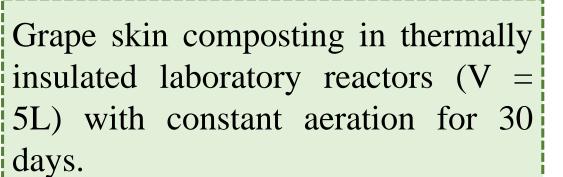


Tea Sokač Cvetnić<sup>1</sup>, Korina Krog<sup>2</sup>, Davor Valinger<sup>2</sup>, Jasenka Gajdoš Kljusurić<sup>2</sup>, Tamara Jurina<sup>2</sup>, Maja Benković<sup>2</sup>, Ivana Radojčić Redovniković<sup>2</sup>, Ana Jurinjak Tušek<sup>2</sup> <sup>1</sup>Institute of Public Health of Zagreb County, Mokrička 54, 10290 Zaprešić, Croatia

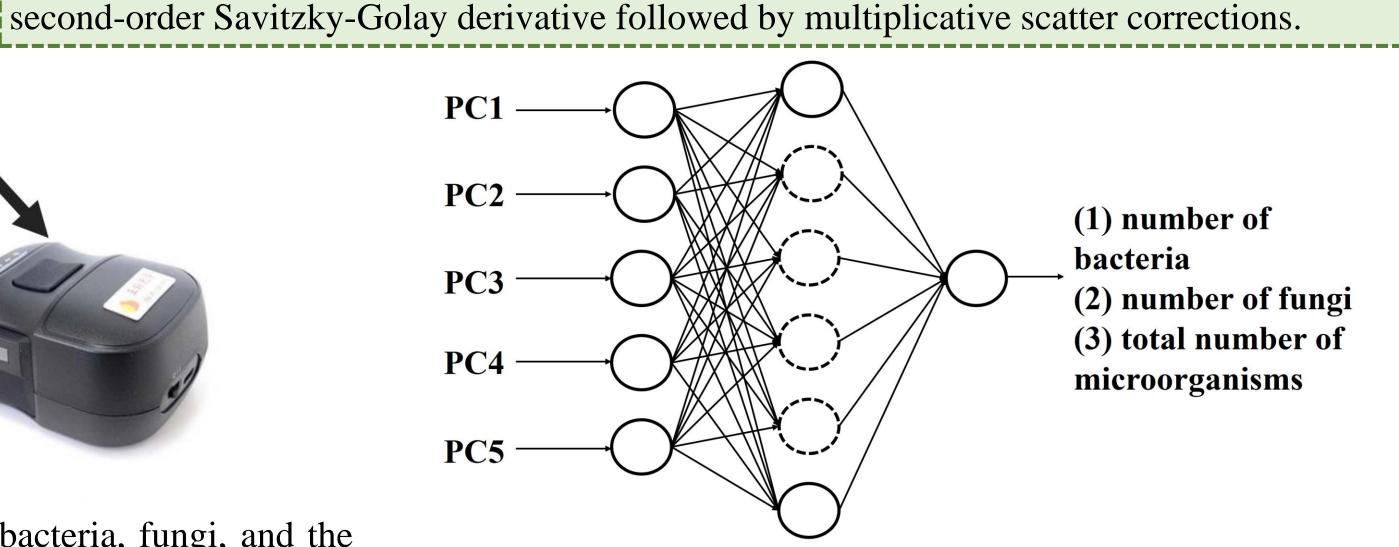
<sup>2</sup>University of Zagreb, Faculty of Food Technology and Biotechnology, Pierottijeva 6, HR-10000 Zagreb, Croatia

**Abstract:** In this work, we coupled near-infrared (NIR) spectra with artificial neural network modeling to predict the number of microorganisms in grape skin compost during 30 days of composting in five independent composting experiments. NIR spectra were gathered in the wavelength range of  $\lambda = 900-1700$  nm using a portable NIR instrument. ANNs were developed using first five factors obtained from PCA analysis of NIR spectra after different processing methods as model inputs, and, the number of bacteria, fungi, and the total number of microorganisms as model outputs. Considering the prediction of the number of bacteria, the acceptable pre-processing method for NIR spectra was RER smoothing (RER=10.083). For the number of fungi and the total number of microorganisms, the pre-processing method SG2D+MSC (second-order Savitzky-Golay derivative followed by multiplicative scatter corrections) was the most acceptable and the developed ANN model resulted in the highest RER values of 15.075 and 12.040, respectively.

NIR spectra preprocessing methods: smoothing, first-order Savitzky-Golay derivative, second-order Savitzky-Golay derivative, standard normal variate, multiplicative scatter **Experimental set-up:** corrections, smoothing followed by standard normal variate, smoothing followed by multiplicative scatter corrections, first-order Savitzky-Golay derivative followed by standard normal variate, first-order Savitzky-Golay derivative followed by multiplicative scatter corrections, second-order Savitzky-Golay derivative followed by standard normal variate, and



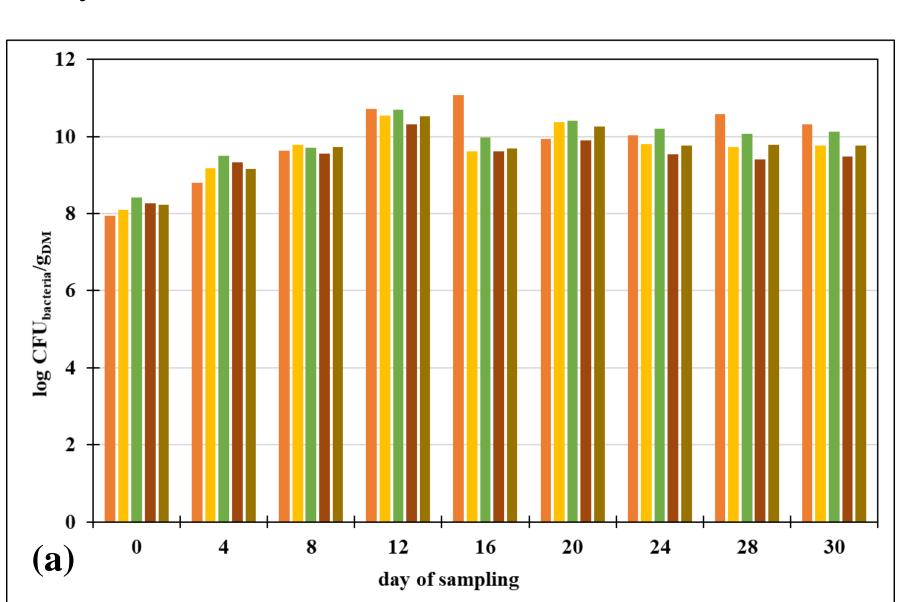
- Measurement of the number of bacteria, fungi, and the total number of microorganisms of the compost samples.
- (ii) NIR spectra were gathered in the wavelength range of  $\lambda$ = 900–1700 nm using a portable NIR instrument.



ANNs were developed using first five factors obtained from PCA analysis of NIR spectra after different processing methods as model inputs, and, the number of bacteria, fungi, and the total number of microorganisms as model outputs

## **Results:**

The results of microbiological analysis during composting processes at different conditions are shown in Figure 1. The microbial growth was monitored every 96 hours.



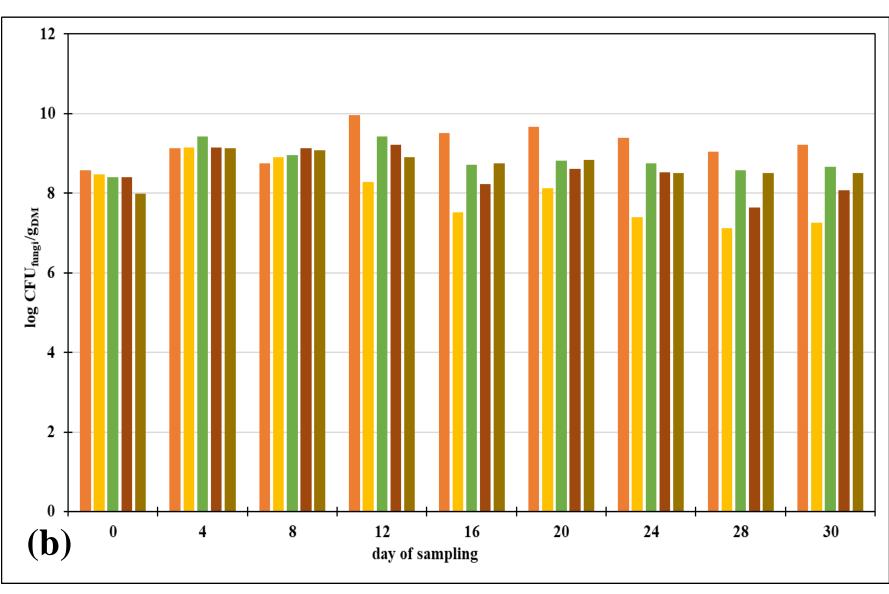


Figure 1. Microbial growth during the 30 days of the grape skin composting process: (a) Bacterial growth and (b) Growth of fungi

The NIR spectra of compost samples are given in Figure 2. The NIR spectra show the changes in during the absorbance composting maximum processes, which indicates the chemical changes of the grape skin that occurred in the process.

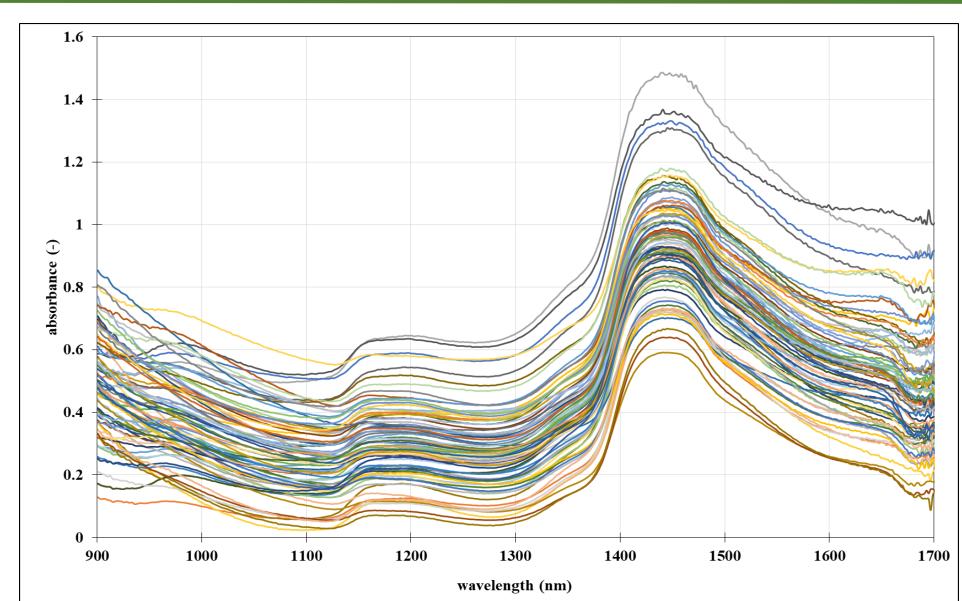


Figure 2. Average NIR spectra of compost samples gathered with portable NIR spectrometer (NIR-S-G1, InnoSpectra, Taiwan)

Due to complexity of the NIR spectra, the preprocessing of spectra has been applied in order to obtain important information about microbiology of compost. The artificial neural network (ANN) models were developed based on the recorded NIR spectra for the compost samples to predict the number of microorganism. Results are given in Table 1.

Considering the prediction of the number of bacteria, the acceptable pre-processing method for NIR spectra was RER smoothing (RER=9.2376). For the number of fungi and the total number of microorganisms, was the most acceptable and the developed ANN model resulted in the highest RER values of 15.075 and 12.040, respectively.

Table 1. Artificial neural network (ANN) models for prediction of number of bacteria during the composting process based on the NIR spectra gathered using portable NIR spectrometer (NIR-S-G1, InnoSpectra, Taiwan)  $(R_{pred}^2)$  = coefficient of determination for prediction;  $R_{\text{pred}}^2$  adjusted coefficient of determination for prediction; RMSEP=root mean square of prediction; SEP=standard error of prediction: RPD=ratio of prediction to deviation: RER=ratio of the error range)

error of prediction; RPD=ratio of prediction to deviation; RER=ratio of the error range)												
NIR spectra pretreatment	Network name	Calibration					Prediction					
		Training perf. Training	Test perf. Test error	Validation perf. Validation	Hidden activation	Output activation	$R_{ m pred}^{2}$	$R_{ m pred}^{-2}_{ m adj}$	RMSEP	SEP	RPD	RER
		error	0.44.	error								
raw	MLP 5-4-1	0.6761 0.3316	0.6157 0.3644	0.6032 0.3865	Exponential	Exponential	0.6789	0.6290	0.7677	0.1637	1.7882	6.6129
smoothing	MLP 5-9-1	0.7165 0.2988	0.7233 0.3798	0.7026 0.3624	Logistic	Identity	0.7705	0.7399	0.5473	0.1167	2.2861	9.2376
SG1D	MLP 5-11-1	0.7406 0.1988	0.6868 0.2678	0.6796 0.2988	Logistic	Identity	0.4405	0.3403	1.1603	0.2474	1.2022	4.3571
SG2D	MLP 5-7-1	0.7321 0.1999	0.6541 0.3344	0.6223 0.2874	Exponential	Identity	0.5692	0.4962	0.9633	0.2054	1.4781	5.2745
SNV	MLP 5-11-1	0.7450 0.1855	0.7302 0.2125	0.7246 0.2574	Exponential	Exponential	0.7510	0.7162	0.5655	0.1206	2.1381	8.9775
MSC	MLP 5-7-1	0.7321 0.1999	0.6541 0.3344	0.6223 0.2874	Exponential	Identity	0.5867	0.5173	0.9629	0.2053	1.4941	5.2722
smoothing+SNV	MLP 5-11-1	0.7450 0.1855	0.7302 0.2125	0.7246 0.2574	Exponential	Exponential	0.6788	0.6288	0.7499	0.1599	1.5242	6.3739
smoothing+MSC	MLP 5-7-1	0.6133 0.3111	0.6206 0.3552	0.6266 0.3445	Logistic	Tanh	0.4727	0.3794	0.9289	0.1980	1.1163	4.3740
SG1D+SNV	MLP 5-7-1	0.8865 0.1888	0.6469 0.3665	0.5936 0.4912	Exponential	Identity	0.5498	0.4727	0.7454	0.1589	1.3418	6.7434
SG1D+MSC	MLP 5-10-1	0.7855 0.2125	0.7642 0.2478	0.7004 0.2999	Tanh	Tanh	0.7515	0.7168	0.5452	0.1162	2.1160	8.7250
SG2D+SNV	MLP 5-3-1	0.7477 0.2575	0.7364 0.2322	0.7333 0.2741	Tanh	Identity	0.7397	0.7026	0.6465	0.1378	2.0829	7.4441
SG2D+MSC	MLP 5-4-1	0.6725 0.2144	0.6938 0.2347	0.6156 0.2777	Exponential	Exponential	0.5332	0.4525	0.6954	0.1483	1.3382	6.6151