Simultaneous determination of hydroquinone and catechol in compost bioremediation using a tyrosinase biosensor and artificial neural networks

ZHOU Yaoyu[1,2], **TANG Lin**[1,2,*], **ZENG Guangming**[1,2,*], **ZHANG Yi**[1,2], **LI Zhen**[3], **LIU Yuanyuan**[1,2],

CHEN Jun[1,2], **YANG Guide**[1,2], **ZHOU Lu**[1,2], **ZHANG Sheng**[1,2]

1 College of Environmental Science and Engineering, Hunan University, Changsha, 410082, China

2 Key Laboratory of Environmental Biology and Pollution Control, Hunan University, Ministry of

Education, Changsha 410082, China

3 The Institute of Packaging and Materials Engineering of Hunan University of Technology, Zhuzhou

412007, China

---

[*] Corresponding author: Tel.: +86-731-88822778; Fax.: +86-731-88822778

E-mail: tanglin@hnu.edu.cn(L. Tang), zgming@hnu.edu.cn(G.M. Zeng).

1

1 **Abstract**

2   A biosensor based on tyrosinase immobilized with ordered mesoporous carbon-Au (OMC-Au),

3 L-Lysine membrane and Au nanoparticles (tyrosinase/OMC-Au/L-Lysine/Au) was combined with

4 artificial neural networks (ANNs) for the simultaneous determination of catechol (CC) and

5 hydroquinone (HQ) in compost bioremediation of municipal solid waste. The good performance of

6 biosensor provided the potential applicability for the simultaneous identification and quantification of

7 catechol and hydroquinone in real samples, and the combination with ANNs offered a good

8 chemometric tool for data analysis in respect to the dynamic, nonlinear, and uncertain characteristics of

9 the complex composting system. Good prediction ability was attained after the ANNs model

10 optimization, and the direct detection range for catechol and hydroquinone were directly analyzed by

11 the ANNs model varied between 1.0×10^{-7} and 1.1×10^{-4} M, significantly extended than the linear model

12 $(4.0 \times 10^{-7}$ to 8.0×10^{-5} M). Finally, the performance of the ANNs model was compared with the linear

13 regression model. The results demonstrated that the prediction results by the ANNs model were more

14 precise than those by the linear regression, and the latter was far from accurate at high levels of

15 catechol and hydroquinone beyond the linear range. All the results showed that the combination of the

16 biosensor and ANNs was a rapid and sensitive method in the quantitative study of composting system.

17 **Keywords:** biosensor; DPV signals; catechol; hydroquinone; artificial neural networks; compost

18 bioremediation.

19

20   **Introduction**

21   Phenolic compounds are widely distributed as environmental pollutants because many of them are

22   resistant to biotic and abiotic degradation. They are mostly derived from various agricultural and

23   industrial activities, including waste discharge from wood preservatives, coking, textiles, plastics, dyes,

24   paper, herbicides industries and the partial degradation of phenoxy contaminants in remediation

25   processes[1,2]. The toxicity of phenols generated from bioremediation processes, such as composting, can

26   bring on undesirable ecological effects and seriously reduce removal efficiencies[3]. Catechol (CC) and

27   hydroquinone (HQ) are two isomers of phenolic compounds which are harmful to human health and

28   ecological environment. During the application of composting technology in disposal of municipal

29   solid waste, CC and HQ are generally direct pollutant or by-product of the aromatic pollutant[4].

30   Therefore, detection and quantification of the toxicity of these phenolic compounds from compost

31   bioremediation of municipal solid waste is a critical issue. Up to now, a great number of analytical

32   methods have been established to determine dihydroxybenzene isomers in compost systems. On the

33   one hand, there are techniques such as high-performance liquid chromatography (HPLC)[5],

34   spectrophotometry[6] and gas chromatography[7], which allow individual identification of phenols, but

35   these procedures usually require specific equipment, laboratory conditions, and are not suitable for

36   on-site analyses. On the other hand, electrochemical methods are applied to detect the hydroquinone,

37   catechol, phenol, resorcinol, cresol. These methods have the advantages of fast response, cheap

38   instrument, low cost, simple operation, timesaving, but the key point lies in improving the sensitivity,

39   selectivity and the potential applicability for the quantification of phenols in real samples. In an attempt

40   to overcome the deficiencies of these analytical methods, the applications of enzyme sensors to specific

41   pollutant detection have been increasingly reported to exhibit superior sensitivity, stability, reusability,

3

42   selectivity, and portability[8]. Especially the biosensor provided the potential to quantify the pollutant

43   levels in real environmental. The operation efficiency of compost systems will be much improved if

44   enzyme sensors are applied in pollutant detection.

45   In our previous works, a tyrosinase biosensor was developed for linear calibration and

46   simultaneous determination of hydroquinone and catechol[9]. The biosensor was evaluated by differential

47   pulse voltammetry (DPV) measurements, which is used to make electrochemical measurements, and

48   the DPV peak currents increased linearly with concentration over the range of 4.0×10^{-7} to 8.0×10^{-5} M,

49   the detection limits of HQ and CCwere 5×10^{-8} and 2.5×10^{-8} M (S/N=3), respectively. The sensitivities

50   in the linear calibration regions for low concentration show the following order: 0.4511 A/M (catechol,

51   n=4) > 0.338 A/M (hydroquinone, n=13). And the electrode showed a rapid and sensitive

52   bioelectrocatalytic response of 65 and 89 s after addition of catechol and hydroquinone, respectively.

53   Using the differential pulse voltammetry (DPV), the wide peak separation and low peak potential

54   ensured the avoidance of interferences, making this biosensor a potential device for real sample

55   applications. However, the detection procedures are still susceptible to the complex samples containing

56   heterogeneous organic components and certain functional groups, such as phenolic OH and carboxyl,

57   especially in compost system which a variety of organic compounds coexisting, owing to both the

58   redox and sorption of the interfering matrix constituents on the electrode surface[8]. As a result, an

59   unstable baseline or even the overlapped differential pulse voltammetry signal will be obtained with a

60   carbon electrode when it was applied to large quantities of compost samples. Although the data

61   generated by simultaneous determination of phenols compounds from compost bioremediation can be

62   analyzed using the linear regression model, nonlinearities and uncertainties also occur in the process as

63   mentioned above, which restrict the biosensor in practical application. Thus, the quantification

4

64   capability of the linear model will be limited by the dynamic, nonlinear, and uncertain characteristics of

65   the complex composting system, and will give erroneous results if the linear range is exceeded.

66   Artificial neural network (ANN) are computational models inspired by animal central nervous systems

67   (in particular the brain) that are capable of machine learning and pattern recognition. They are usually

68   presented as systems of interconnected "neurons" that can compute values from inputs by feeding

69   information through the network. It have found extensive utilization in solving many complex

70   real-world problems. ANNs could be deemed as advanced signal processing variants allowing the

71   interpretation, modelling and calibration of complex analytical signals for they can process very

72   nonlinear and complex problems even if the data are imprecise and noisy[8,10,11]. The combination of the

73   tyrosinase biosensor with ANNs modelling may represent an alternative to classical methods. This

74   approach has already been introduced towards the analysis of phenols. For example, the group of

75   Xavier Cetó and Francisco Céspedes has used this method to manage the sensor signal, and established

76   electronic tongue and Bio-Electronic Tongue (BioET) based on voltammograms correlated to phenol

77   contents in wines[12-16]. In addition, Tang group has used this method to handle the biosensor signal,

78   processing the amperometric signals correlated to enzyme activities or phenol contents in compost

79   system[8,26].

80   In this work, the application of ANN technique for evaluation of the DPV signals of

81   multi-component analysis generated by the tyrosinase biosensor for the simultaneous determination of

82   CC and HQ in compost extract samples was explored, which has not been reported. This method

83   combining the advantages of both parts, calibrated the complex overlapping analytical signals and

84   imprecise data from composting samples. The aim of the study was to extend the limited measuring

85   range of the biosensor to a useful and wider working band. This assay provided the potential

5

86　applicability of the biosensor for the quantification of CC and HQ in compost system, and the

87　development of fast and inexpensive on-line monitoring systems in municipal solid waste compost

88　bioremediation.

89　**Experimental**

90　**Apparatus and reagents**

91　　Cyclic voltammetric (CV) measurement and differential pulse voltammetry (DPV) measurement

92　were carried out on CHI660B electrochemistry system (Chenhua Instrument, Shanghai, China). Model

93　PHSJ-3F laboratory pH meter (Leici Instrument, Shanghai, China) was used to test pH value. A Sigma

94　4K15 laboratory centrifuge, a vacuum freezing dryer and a mechanical vibrator were used in the assay.

95　The three-electrode system used in this work consisted of a tyrosinase/OMC-Au/L-Lysine/Au/glassy

96　carbon electrode (GCE) as working electrode, a saturated calomel electrode (SCE) as reference

97　electrode and a Pt foil auxiliary electrode. All the work was conducted at room temperature (25 $^{\circ}$C)

98　unless otherwise mentioned.

99　　Tyrosinase (EC 1.14.18.1, from mushroom as lyophilized powder), catechol and hydroquinone

100　were purchased from Sigma-Aldrich (USA). Tetraethoxysilane (TEOS), L-Lysine, Gold(III) chloride

101　trihydrate (HAuCl$_4$·4H$_2$O, 99.9%) and all other chemicals were of analytical grade and used as received.

102　Phosphate buffer solutions (1/15 M PBS) with different pH 6.98 were prepared by mixing the stock

103　solutions of KH$_2$PO$_4$ and Na$_2$HPO$_4$·12H$_2$O. All solutions were prepared with double-distilled water.

104　The synthesis of OMCs-Au nanoparticles and the immobilization of tyrosinase on the surface of

105　nanoparticles were achieved according to the procedure introduced by Tang et al[9].

106　**Procedures**

6

107   The preparation of tyrosinase/OMC-Au/L-Lysine/Au/GCE and the measurements of CC and HQ

108   were carried out as described in our previous work[9]. Briefly, the AuNPs and L-Lysine were

109   immobilized on glassy carbon electrode by electrochemical method. OMC-Au/L-Lysine/Au/GCE was

110   prepared by casting 5.0 μL of the OMC-Au suspension onto the surface of the L-Lysine/Au/GCE,

111   Finally, tyrosinase was immobilized on the electrode surface, as presented in Scheme S1. Au

112   nanoparticles (AuNPs) modified glassy carbon electrode (GCE) due to their high effective surface area,

113   nano-scaled dimension effects, and most importantly, binding affinity with L-Lysine. In addition,

114   L-Lysine provided amino and became the cross-linking agent between AuNPs film and OMC-Au film,

115   and OMC-Au could not only unite with L-Lysine, but also combined with tyrosinase. This makes the

116   enzyme more fixed on the biosensor, accelarates the electron transfer from the enzyme-catalysed redox

117   reaction to electrode surface, and extend its using life as well[9]. Under the optimized condition, 10 mL

118   compost extract samples containing different concentrations of CC and HQ were added into an

119   electrochemical cell, and then the three-electrode system was installed on it. The DPV was recorded

120   from +0.6 to -0.2 V with pulse amplitude of 0.05 V, pulse width of 0.05 s, and pulse period of 0.2 s.

121   The CV was performed between −0.6 and +0.8 V at scan rate of 50 mV·s^{-1}, sample interval of 0.0001

122   V and quiet time of 2 s.

123   **Preparation of compost extracts**

124   The biosensor simultaneous determination of the CC and HQ concentration was applied in

125   compost bioremediation. The composting process has been introduced previously[17]. The components of

126   compost were soil, straw, restaurant leftover, and bran, and the water ratio was 51%. The soil was

127   collected from 100 cm underground on the unfrequented hillside of Yuelu Mountain (Changsha, China),

128   from which large organic scraps were removed. Then aerobic compost was managed 40 days under the

7

129   condition of 30 °C temperature and $0.033m^3 \cdot h^{-1}$ ventilation. 10 g of compost sample was placed in a

130   flask and 200 mL water was added in. The suspension was agitated on a mechanical vibrator at 200

131   rpm for 2 h. The supernatant was centrifuged at 10000 rpm for 5 min, and then filtered to get the filtrate

132   as the compost extract. All the work was done at room temperature unless otherwise mentioned. The

133   dosage of CC and HQ into each compost extract was controlled using certain volumes of CC and HQ

134   stock solutions.

135   **Data processing**

136   Chemometric processing was done by specific routines in MATLAB 7.0 (Math Works, Natick,

137   MA) written by the authors, using Neural Network Toolboxes to develop the ANN models. SigmaPlot

138   12.0 (Systat Software Inc, California, USA) was used for graphic representations of data and results.

139   The measured data of a total set of 38 samples using the biosensor were divided into three datasets.

140   22 samples for the training set were used to build the proper modeling of the response, 8 samples

141   randomly distributed for the testing set were used to estimate the modeling performances, and another

142   8 extract samples were used to validate the ANN model application. The biosensor DPV responses of

143   compost samples with corresponding CC and HQ concentrations were analyzed using a feed-forward

144   back propagation (BP-ANN). This artificial neural network model for variable selection aims to find an

145   optimal set of inputs that can quickly and successfully classify or predict the desired outputs. It was a

146   feed-forward network combining a back propagation algorithm which was used to train the network

147   according to a learning rule[18]. For each sample, a complete DPV was recorded for forming the array

148   and data. In order to reduce the high dimensionality of the recorded signals, to prevent larger numbers

149   from overriding smaller ones, and to prevent premature saturation of hidden nodes, which impedes the

150   learning process, a pre-processing stage was required. There is no one standard procedure for

8

151   normalizing inputs and outputs[19]. But it is recommended that the data be normalized between slightly

152   offset values such as 0.1-0.9 and to avoid saturation of the sigmoid function leading to slow or no

153   learning[20,21]. For this, the input values of both the training and the test subsets were kept in interval

154   [0.1,0.9] corresponding to the range of the normalized function:

155   $$X_i = 0.1 + 0.8 \left(\frac{Z_i - Z_i \min}{Z_i \max - Z_i \min} \right) \qquad (1)$$

156   where X_i is the normalized value of input variable, Z_i is the original value, and $Z_i \max$, $Z_i \min$ are

157   the maximum and minimum original values of primitive data, respectively. After simulation of the

158   networks, the estimated results were reconverted by inverse function of Eq. (1) to be compared with the

159   target values.

160   For complete assessment of model performances, the root mean square error (RMSE) was used,

161   which was calculated between expected and predicted concentration values for each sample (i) and for

162   each of the two analytes (j) considered:

163   $$RMSE = \sqrt{\frac{\sum_{ij} (Z_{ij} - \hat{Z}_{ij})^2}{3n - 1}} \qquad (2)$$

164   **Results and discussion**

165   **Artificial neural network architecture**

166   In present study, examples of the different curves of current versus time were obtained

167   corresponding to the mixed CC and HQ concentration in spiked compost extract samples. Fig. 1 shows

168   the current response curves for 22 compost extract samples in the training set. The concentrations of

169   CC and HQ in the filtrates both varied from 0.10 to 110 μM. In addition, Fig. 1 presents that a

170   maximum and a minimum signal (any of the 38 currents) of the target concentration were included in

171   the training set, avoiding the need for extrapolation when test the model with the external dataset. It

172   will not give precise results to assign a specific reduction peak potential to each phenolic compound

9

173   using the statistics of the fitted regression linear model[12], due to some signal overlapping (as shown in

174   Fig. 1). Therefore, BP-ANN method was used to deconvolve the strong overlapping signal and to

175   quantify the concentrations of two phenolic compounds separately, because ANN modelling was

176   considered to be an appropriate chemometric tool for solving overlapping and nonlinear problems,

177   whose structure was designed to imitate the organization of human brain[22].

178   "Here Fig. 1"

179   Generally, a BP-ANN comprises three parts: an input layer, an output layer and in between the

180   two layers, there are one or more hidden layers[23]. Each layer is formed by a series of interconnected

181   neurons, and the value at each neuron is weighted and transformed by a transfer function[24]. A

182   simplified scheme of the procedure followed for the measurement and data processing could be seen in

183   Fig. 2. The architecture of the ANN used was defined by these data: the response curves of 22 samples

184   for the training set, the response curves of 8 samples to evaluate model's response, and another 8

185   extract samples to validate the BP-ANN model application compared with regression liner model. The

186   input layer consisted of a certain number of individual data points of each DPV curve and the output

187   layer consisted of two neurons, namely the two concentrations sought. We used a single intermediate

188   layer, known as the hidden layer, since it was stated that an appropriate level of modelling could be

189   achieved with a single hidden layer in the electrochemical signal resolving process in the relative

190   literature[25]. So did our experience in previous work also show[8,26]. Thus, Networks with more than one

191   hidden layer were not considered.

192   "Here Fig. 2"

193   **Network optimization**

194   A study of the BP-ANN architecture was performed in order to optimize the separate

10

195   quantifications of the two phenols considered. 22 current intensities at specific potentials for the array

196   of DPV were selected as input vector in the BP-ANN, the corresponding concentrations being the

197   targets that the modelling should reach. The learning accomplished (the degree of modelling) was

198   estimated by the root mean square error (RMSE, equation 2). The training process was continued until

199   a preset fitness degree was achieved (RMSE value). Fig. 2 shows the BP-ANN architecture and scheme

200   of this BP-ANN based approach. There are four elements that comprise the ANNs architecture: (a)The

201   number of layers, (b)The number of neurons in each layer, (c)The activation function of each layer,

202   (d)The training algorithm (because this determines the finalvalue of the weights and biases). The

203   number of neurons in each of these two layers is specified by the number of input and output

204   parameters that are used to model each problem so it is readily determined. Therefore, the objective is

205   to find the number of neurons in hidden layer firstly[24]. Besides, the effects of different transfer function

206   combinations and hidden neuron numbers on the network performance were studied synchronously.

207   Combinations of tan-sigmoidal (Tansig), sat-lineal(Satlin), pure-lineal (Purelin) and log-sigmoidal

208   (Logsig) transfer functions and the hidden neuron numbers (changed from 2 to 16) were tested, as seen

209   on Fig. 3A, with the optimum results of 27 as input neuron number and Levenberg-Marquardt

210   backpropagation (trainlm) as optimization algorithm. Each architecture was retrained five times to get

211   the average RMSEs for the external test set to result in a accurate measure of performance. According

212   to Fig. 3A, the lowest RMSE value was obtained with 10 hidden neurons and Logsig-Purelin as

213   transfer function.

214   Afterwards, the next step was to determine the importance of network inputs and different

215   optimization algorithms. Similarly, the effects of the input neuron and different optimization algorithms

216   on the model performance were evaluated and optimized in parallel. Fig. 3B shows the RMSEs for

11

217   different input neuron numbers and optimization algorithms with the optimal transfer function

218   combination of Logsig-Purelin and hidden neuron number of 10. According to Fig. 3B, the BP-ANN

219   models with trainbr (Bayesian regularization backpropagation), trainbfg (BFGS quasi-Newton method),

220   traingdm (momentum backpropagation), traincgb (Powell-Beale restarts), traingd (gradient descent

221   backpropagation) and traingdx (backpropagation) as optimization algorithms, respectively, could not

222   meet the performance goal and lowest RMSE. So those algorithms were not taken into account.

223   Trainlm (Levenberg-Marquardt backpropagation) was chosen as the one for the best performance.

224   Once the BP-ANN model was trained, inputs that made relatively small contributions to the variance in

225   our experiment, and it was reasonable that the accuracy of the simulation of the ANN model might

226   increase with more input current values, but the training time was prolonged remarkably with no

227   obvious decrease of RMSE. Therefore, the value number of 9 was selected as the input neuron number

228   with adequate accuracy of simulation.

229   "Here Fig. 3"

230   For all these reasons, the best model was obtained by using a 9×10×2 network that used a Logsig

231   transfer function in the hidden layer and a Purelin function in the output layer with

232   Levenberg-Marquardt backpropagation (trainlm) as optimization algorithm (shown in Table 1).

233   "Here Table 1"

234   **Performance of the best ANN**

235   Fig. 4 presented the training performances for the two analytes with the optimal BP-ANN

236   configuration, where the predicted concentrations of the two considered phenols were compared with

237   their expected concentrations. The concentrations of CC and HQ added in compost extract in the

238   experiment both varied between 0.10 and 110 μM. Error bars were plotted by five different retrainings

12

239   with random reinitialization of weights for the best architecture, giving information about the

240   reproducibility of model. According to Fig. 4, an excellent ability to represent the information on the

241   learning process was obtained with BP-ANN. More valuable was the modelling and prediction

242   capability working with dataset not included in the learning process. Fig. 5 showed the performance of

243   the best ANN on the external testing subset, with data not included in the learning process. Prediction

244   capability of the model could be considered satisfactory due to the very good correlations were

245   obtained in all cases.

246   "Here Fig. 4"

247   "Here Fig. 5"

248   **Comparison of prediction results between regression model and ANN model in composting**

249   **system**

250   In order to compare the performance of the BP-ANN model with the linear regression model in

251   respect to correlation coefficient, adaptability to uncertainty, etc., some compost extract samples were

252   spiked with various amounts of the two phenolic compounds distributed in the range of the

253   experimental design. These were prepared and analyzed employing the BP-ANN model and linear

254   regression model. Both the linear model composed of Eqs. (3) and Eqs. (4) obtained in our previous

255   work[9] and the BP-ANN model established here were applied into composting system to predict CC and

256   HQ concentrations in eight compost extract samples.

257   P_{HQ}= -66.954-9.5357lg[HQ] (P_{HQ}: μA, [HQ]: M); (R = 0.9565)   (3)

258   P_{CC}= -88.394-13.081lg[CC] (P_{CC}: μA, [CC]: M); (R=0.9771)   (4)

259   Practically, there exist a variety of organic compounds in compost extract, such as aromatic,

260   aliphatic, phenolic and quinolic derivatives with varying molecular sizes and properties. It was a

13

261 complex mixture with diversity, nonlinearity, and uncertain characteristics. In this case, although high

262 specificity and selectivity of biosensor were obtained, when linear model is applied to determine the

263 real samples, the overlapped differential pulse voltammetry signal and the concentration of analyte

264 often exceeds the linear detection range of biosensor, which will affect the accuracy of determination.

265 Therefore, for the sake of obtaining a more applicable and convenient detection method, the

266 combination of biosensors with BP-ANN modelling may turn out to be an alternative tool to classical

267 methods, taking benefit of the advantages of both parts. On one hand, the selectivity, reproducibility

268 and stability of biosensor confirmed the potential applicability for the simultaneous determination of

269 CC and HQ in real environmental samples[9]. On the other hand, the use of ANNs modelling to

270 deconvolve complex signals can enlarge the detection range, and then make the quantification and the

271 result analysis more efficient[25].

272   In this study, The DPV peak currents of HQ and CC were linear with correlative concentrations

273 over the range from 4.0×10^{-7} to 8.0×10^{-5} M[9], while the BP-ANN model can directly analyze CC and

274 HQ concentrations varying between 1.0×10^{-7} and 1.1×10^{-4} M. Each of the calibration was done five

275 times with the relative standard deviations (RSD) not more than 5%. Also in this case, the recovery

276 yield for the two phenolic compounds was calculated, which is summarized in Table 2. As can be seen,

277 the recovery yield of CC obtained by linear regression model ranges from 73.9% to 115.2%, while that

278 obtained by BP-ANN model ranges from 96.0% to 115.3%. It is also observed that the recovery yield

279 of HQ calculated by linear regression model ranges from 74.6% to 119.0% , while that calculated by

280 BP-ANN model ranges from 88.15% to 112.0%. As seen on Table S1, the RSD in the linear regression

281 model for CC and in the ANN for CC were 7.73% and 3.7781%, respectively. Although the RSD of

282 linear regression model in the compost extract sample of 4 is lower than the RSD of ANN, the RSD of

14

283   the rest samples are lower when analyzed by the ANNs model. In addition, the average (RSD) of ANN

284   is lower than the RSD of linear regression model. What's more, the RSD in the linear regression model

285   for CC (21.2004%) is significantly higher than the RSD for the ANN (2.1151%) when sample

286   concentration exceeded the linear range of the biosensor. Correspondingly, as seen on Table S2, the

287   RSD in the linear regression model for HQ and in the ANN for HQ were 10.9592% and 4.8468%,

288   respectively. Obviously, the average (RSD) of ANN is lower than the RSD of linear regression model.

289   The results demonstrated that the prediction results by the ANN model were more precise than the

290   linear regression. The prediction result by linear regression was far from accurate at high levels of CC

291   and HQ beyond the linear range, while the fitting degree of experimental and predicted value using the

292   ANN model were satisfactory (see table 2), thus confirming the BP-ANN model was superior to the

293   linear regression especially for the determination of high levels of CC and HQ in the compost system.

294   Furthermore, the results also showed that the correlation coefficient, adaptability to uncertainty, etc.,

295   obtained after combining the biosensor with BP-ANN were superior to direct linear determination of

296   the CC concentration by the biosensor in the compost system. Obviously, combined with the BP-ANN

297   model, the direct detection range for CC and HQ in the compost system of the biosensor were widened,

298   and the satisfactory results confirmed the potential applicability of the biosensor for quantification of

299   CC and HQ in real compost extract sample determination.

300   "Here Table 2"

301   **Conclusions**

302   In summary, a very good quantification of the two phenolic compounds has been achieved by

303   using the tyrosinase biosensor to get specific signal and BP-ANN as the tool for building the response

304   model. From all the results shown above, it is demonstrated that the combination of tyrosinase

15

305   biosensor and BP-ANN can give satisfactory quantifications of the CC and HQ concentration

306   simultaneously in composting system with good rapidity and sensitivity. Besides, the direct detection

307   range for CC and HQ of the biosensor was extended to 1.0×10^{-7}-1.1×10^{-4} M, which was superior to the

308   direct determination by the biosensor with linear data analysis. What's more, this assay provided the

309   potential applicability of the biosensor for the quantification of CC and HQ in composting system

310   though with plenty of interfering substances. In future work, this biosensor combined with artificial

311   neural networks model may be alternatively applied for the quantification of different phenolic

312   mixtures in real contaminated compost samples or other complex environments samples.

313   **Acknowledgements**

16

References

320   1 D. Shan, J. Zhang, H. G. Xue, Y. C. Zhang, S. Cosnier and S. N. Ding, *Biosens. Bioelectron.*, 2009,

321   **24**, 3671-3676.

322   2 L. Lu, L. X. Zhang, S. Zhang, S. Huan, G. Shen and R. Yu, *Anal. Chim. Acta.*, 2010, **665**, 146-151.

323   3 A. Lante, A. Crapisi, Krastanov A and P. Spettoli, *Process Biochem.,* 2000, **36**, 51-58.

324   4 S. Canofeni, S. D. Sario, J. Mela and R. Pilloton, *Anal Lett.,* 1994, **27**, 1659-1662.

325   5 A. Escarpa, M. D. Morales and M. C. Gonzalez, *Anal. Chim. Acta.,* 2002, **460**, 61-72.

326   6 P. Nagaraja, R.A. Vasantha and K.R. Sunitha, *J. Pharm, Biomed. Anal.,* 2001, **25**, 417-424.

327   7 B. Levaj, V. D. Uzelac, K. Delonga, K. K. Ganic, M. Banovic and D. B. Kovacevic, *Food. Technol.*

328   *Biotechnol.,* 2010, **48**, 538-547.

329   8 L.Tang, G. M. Zeng, J. X. Liu, X. M. Xu, Y. Zhang, G. L. Shen, Y. P. Li and C. Liu, *Anal. Bioanal*

330   *Chem.,* 2008, **391**, 679-685.

331   9 T. Lin, Y. Y. Zhou, G. M. Zeng, Z. Li. Y. Y. Liu, Y. Zhang, G. Q. Chen, G. D. Yang, X. X. Lei and M. S.

332   Wu, *Analyst,* 2013, **138**, 3552-3560.

333   10 F. Despagne and D. L. Massart, *Analyst,* 1998, **123**, 157R-178R.

334   11 Y. N. Ni and S. Kokot, *Anal. Chim. Acta.,* 2008, **626**, 130-146.

335   12 A. Gutés, F. Céspedes, S. Alegret and M. delValle, *Biosens. Bioelectron.*, 2005, **20**, 1668-1673.

336   13 X. Cetó, F. Céspedes, M. I. Pividori, J. M. Gutiérrez and M. delValle, *Analyst*, 2012, **137**, 349-356.

337   14 A. Gutés, A. B. Iba´nez, F. Céspedes, S. Alegret and M. delValle, *Anal. Bioanal.Chem.,* 2005, **382**,

338   471-476.

339   15 X. Cetóa, J. M. Gutiérrezb, M. Gutiérrezc, F. Céspedesa, J. Capdevilad, S. Mínguezd, C. J.

340   Jorquerac and M. Vallea, *Anal. Chim. Acta., 2012,* **732**, 172-179.

341   16 X. Ceto, F. Céspedes and M. Valle, *Talanta*, 2012, **99**, 544-551.

342   17 G. M. Zeng, L. Tang, G. L. Shen, G. H. Huang and C. G. Niu, *Int. J. Environ. Anal. Chem.,* 2004, **84**,

343   761-774.

344   18 J. R. M. Smits, W. J. Melssen, L. M. C. Buydens and G. K. Chemom. *Intell. Lab. Syst.,* 1994, **22**,

345   165-189.

346   19 I. A. Basheera and M. Hajmeer, *J. Basic. Microbiol.,* 2000, **43**, 3-31.

347   20 M. H. Hassoun, MIT Press, Cambridge, MA. 1995.

348   21 T. Masters, Academic Press, Boston, MA. 1994.

349   22 F. Rosenblatt, *Psychol. Rev.,* 1958, **65**, 386-408.

350   23 C. G. Looner, Oxford University Press, New York, MA. 1997.

351   24 P. G. Benardos and G. C. Vosniakos, *Engineering Applications of Ar.,* 2007, **20**, 365-382.

352   25 A. Gutés, A. B. Ibáńez, F. Céspedes, S. Alegret and M. delValle, *Anal. Bioanal. Chem.,* 2005, **382**,

353   471-476.

354   26 L. Tang, G. M. Zeng, G. L. Shen, Y. Zhang, G. H. Huang and J. B. Li, *Anal. Chim. Acta.*, 2006, **579**,

355   109-116.

1
2
3
4   **Table captions**
5
6   **Table 1.** Optimal results of ANN architecture and training parameters.
7
8
9   **Table 2.** Detailed results obtained for the spiked compost extract samples against added concentrations
10
11   of the two phenolic compounds considered. Recovery yield was also expressed for each compost
12
13
14   extract sample.
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59   19
60

Figure captions

Fig. 1 Measured signals were obtained from 22 compost extract samples using in the training set.

Fig. 2 Example of the ANN architecture used to interpret DPV signals. The input vector comprises 9 to 27 individual data points in the DPV curve. The number of hidden neurons ranges from 2 to 16 (for clarity, only 10 are shown here).

Fig. 3 Obtained RMSEs in: (A) prediction for different transfer function combinations and neuron numbers in the hidden layer with input neuron number of 27 and Levenberg-Marquardt backpropagation (trainlm) as optimization algorithm. (B) prediction for different input neuron numbers and optimization algorithms with the optimal transfer function combination of Logsig–Purelin and hidden neuron number of 10.

Fig. 4 Modeling performance achieved for the optimized BP-ANN with 22 samples from the training set. Error bars correspond to 5 different retrainings with random reinitialization of weights for the final architecture. Expected concentrations are plotted against those obtained from the BP-ANN, good correlations were obtained for catechol and hydroquinone.

Fig. 5 Modelling performance of the optimised BP-ANN for the external test set. Expected concentrations are plotted against those obtained by BP-AN. Good correlations were obtained for catechol and hydroquinone.

20

Table 1. Optimal results of ANN architecture and training parameters.

Architecture / parameter	Value
Input neuron number	9
Hidden neuron number	10
Output neuron number	2
Transfer function in the hidden layer	Logsig
Transfer function in the output layer	Purelin
Optimization algorithm	Levenberg-Marquardt backpropagation (trainlm)

Table 2. Detailed results obtained for the spiked compost extract samples against added concentrations

of the two phenolic compounds considered. Recovery yield was also expressed for each compost

extract sample.

Compost extract sample	CC concentration /μM					HQ concentration /μM				
	Added	[L] predicted	[B]predicted	[L]Recovery	[B]Recovery	Added	[L] predicted	[B]predicted	[L]Recovery	[B]Recovery
1	1.3	1.1±0.37	1.5±0.33	84.6%	115.3%	2.5	2.0±0.46	2.8±0.36	80.0%	112.0%
2	4.6	5.3±0.41	5.0±0.18	115.2%	108.7%	15.5	14.3±0.39	14.9±0.28	92.3%	96.1%
3	17.8	17.4±0.23	17.9±0.11	97.8%	100.6%	20.5	20.0±0.44	20.6±0.37	97.6%	100.5%
4	25.6	27.5±0.44	28.0±0.17	107.4%	109.4%	36.3	31.2±0.40	32.0±0.16	86.0%	88.15%
5	32.3	30.5±0.39	31±0.29	94.4%	96.0%	10.5	12.5±0.29	8.9±0.19	119.0%	88.6%
6	39.5	37.7±0.35	40.3±0.30	95.4%	102.0%	60.5	57.1±0.36	58.6±0.23	94.4%	96.9%
7	59.3	63.5±0.42	60.2±0.15	107.1%	101.5%	83.6	65.8±0.32	85.9±0.29	78.7%	102.8%
8	95.5	70.6±0.47	98.4±0.26	73.9%	103.0%	105.4	78.6±0.38	109.8±0.21	74.6%	104.2%

[B] BP-ANN model
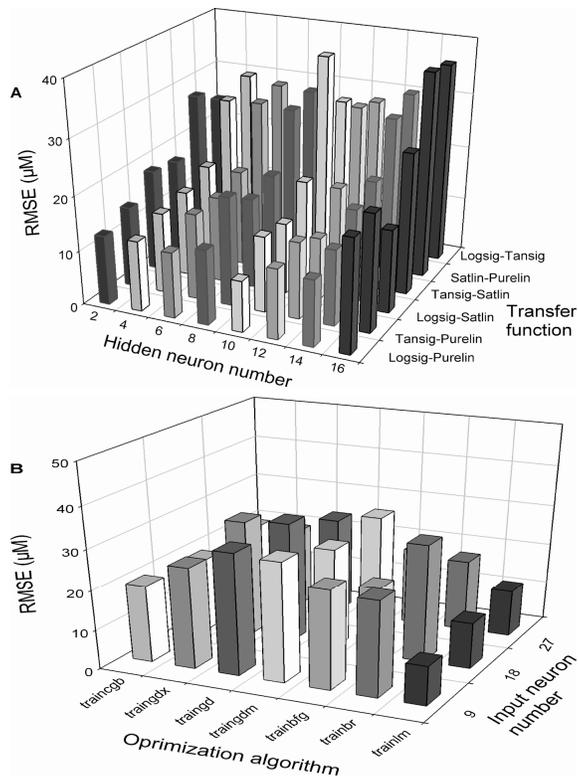
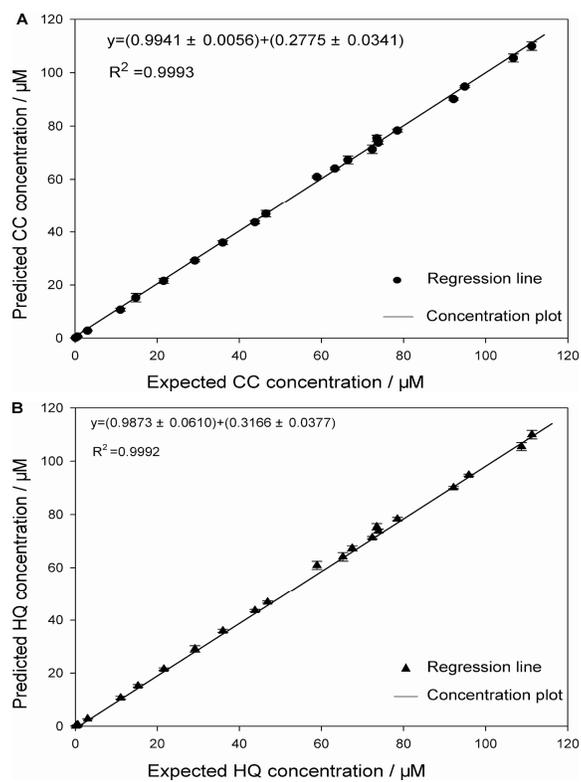[L] linear model
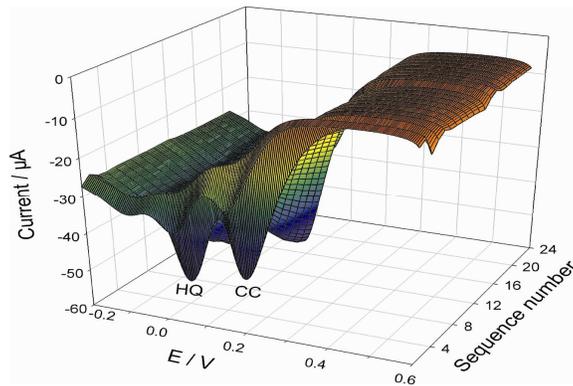
Fig. 1
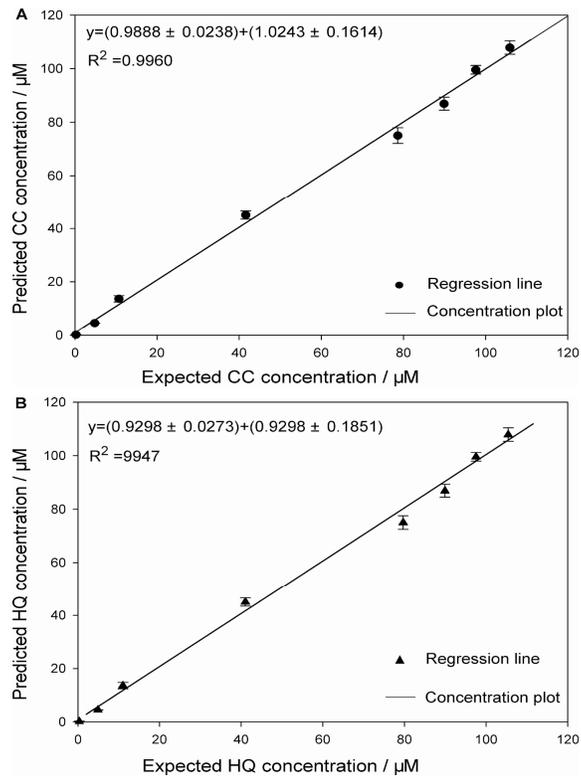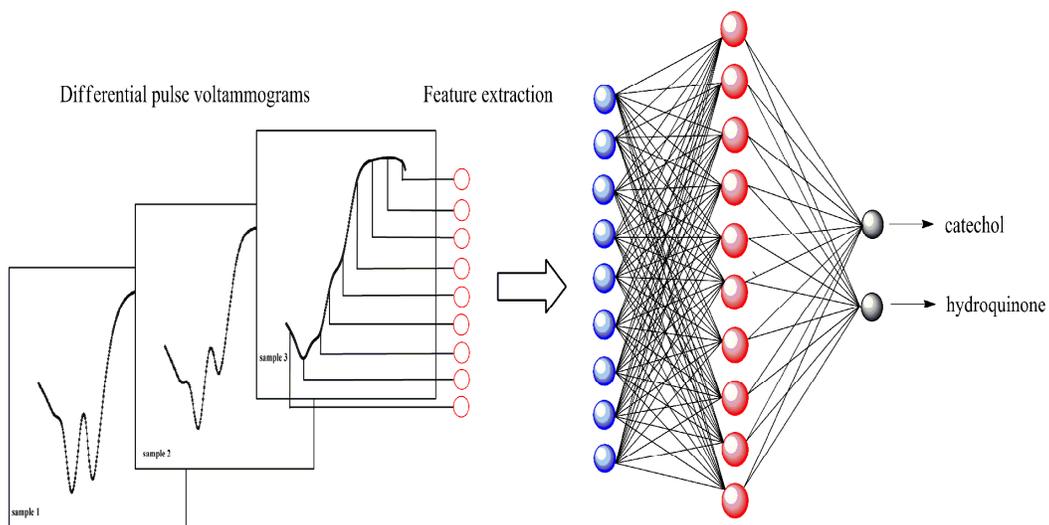
23

Fig. 2

24

Fig. 3

Fig. 4

26

Fig. 5

Contents entry

Selected current intensities of the DPV from biosensor are taken as the input in the ANN. Appropriate weights and biases are applied by the learning algorithm until the targets are satisfied.