Int. J. Electrochem. Sci., 10 **(2015)** 10119 - 10131

International Journal of
ELECTROCHEMICAL
SCIENCE
www.electrochemsci.org

Short Communication

Discrimination of Milk Adulterated with Urea Using Voltammetric Electronic Tongue coupled with PCA-LSSVM

Yaping Yu[1,2], *Hui Zhao*[1,2,*], *Guimei Dong*[2], *Renjie Yang*[2], *Liuan Li*[3], *Yuan Liu*[2], *Haiyun Wu*[2], *Weiyu Zhang*[2]

[1]School of Electrical Engineering & Automation, Tianjin University, No.92 Weijin Road, Tianjin 300072, China;
[2] College of Engineering and Technology, Tianjin Agricultural University, No.22 Jinjing Road, Tianjin 300384, China;
[3]College of Animal Science and Veterinary Medicine, Tianjin Agricultural University, No.22 Jinjing Road, Tianjin 300384, China;
[*]E-mail: zhaohui3376@163.com

Received: 26 August 2015 / *Accepted:* 23 September 2015 / *Published:* 4 November 2015

This paper proposed a novel method for qualitative analysis of urea-tainted milk samples based on voltammetric electronic tongue coupled with multivariable analysis method. The electronic tongue adopted three kinds of metals, such as gold, platinum, and palladium, as working electrodes. 26 pure milk samples and 26 urea-tainted milk samples with different concentration values were prepared. The measurement principles involved the extraction of information from cyclic voltammograms recorded in unadulterated and adulterated milk samples. The response current for different working electrodes has been considered for data analysis. A mean-centering and normalization method was employed for the raw current data. The classification algorithm of this system is separated into two phases: (1) the feature reduction by principal component analysis (PCA); (2) the classification by least square support vector machine (LSSVM). When principal component factors are equal to 10, the optimal PCA-LSSVM model was obtained and discrimination rate of the LSSVM model are 100% and 94.12% in the calibration set and the prediction set, respectively. The results demonstrated that the voltammetric electronic tongue technology with PCA-LSSVM multivariable data analysis method can be successfully applied to the classification between unadulterated milk samples and adulterated milk samples.

Keywords: Voltammetry; milk adulteration; Urea; PCA; LSSVM

1. INTRODUCTION

Milk is one of the most nutritious foods consumed around the world. But it is alarming that many unscrupulous vendors are adulterating it with hydrogen peroxide, melamine, and urea which have harmful effect on human health. The adulteration of urea (CN2H4O) in milk was greater than that of other adulterants [1], due to high nitrogen content of 46%, which can be added in raw milk to enhance the nitrogen content resulting in a false appearance of a higher level of protein using the Kjeldahl method [2]. In addition, a small amount of urea was added into pure milk in order to increase shelf life of milk [3]. As we know, the urea above 700mg/L in milk may cause indigestion, malfunctions of kidney, even cancers, etc [4]. Hence, it is essential to develop a rapid, available, and cost-effective method to identify the milk adulterated with urea.

There are several different methods reported for the detection of urea-tainted milk, such as chemical method, impedance biosensor[5], pH electrodes, ammonium ion-selective electrodes[6], optical, amperometric and conductometric transducers, Chromatography [7] with UV or mass spectroscopy (MS) detector [8], two-dimensional infrared correlation spectroscopy [9,10], and electrochemistry[11-13]. The voltammetric electronic tongue, as a kind of measurement system, which has been performed in discrimination, classification, quality control and process monitoring fields of foods, such as tea[14,15], water[16], beers[17], wine[18], honey [19], beverage[20] and milk[21] etc. Voltammetry technique has a lot of advantages, such as high sensitivity, versatility, simplicity, and good signal to noise ratio [22]. When a potential is applied between the working electrode and the reference electrode, the transient current is obtained from the working electrode surface in voltammetry. Also voltammetry includes versatile techniques like cyclic voltammetry(CV), SWV(Small amplitude pulse voltammetry) and LAPV(Large amplitude pulse voltammetry), etc. In this work, cyclic voltammetry is chosen in our study. The current response data of each working electrode were collected, and the results showed non-adulterated milk samples have bigger current value than the ones found for the milk adulterated with urea. There are two main reason, one reason is passivation of the electrode surface by urea adulterant[12,13]; The other is that addition of urea make milk sample more acidic, conductivity change of the samples measured varies from urea concentration [1]. In this paper, the electronic tongue was used to classify urea-tainted milk samples according to those reasons. But voltammetric sensors generated numerous data, which make their treatment extremely difficult. To avoid redundancy in input data and to obtain LSSVM model with better discrimination ability, PCA was employed for data reduction and the extracted significant information, and as the input data of LSSVM to analyze and predict urea-tainted milk.

2. EXPERIMENTAL SECTION

2.1 Materials and Samples

A total 52 samples were analyzed, including 26 pure milk samples and 26 urea-tainted samples. Pure milk samples（Yili dairy company, China）were obtained from a local supermarket and mixed with different amounts of urea. The concentration of the adulterant were 200mgL^{-1}, 300mg L^{-1}, 400mg L^{-1}, 500mg L^{-1}, 600mg L^{-1}, 700mg L^{-1}, 800mg L^{-1}, 900mg L^{-1}, 1gL^{-1},1.3 gL^{-1}, 1.5 gL^{-1}, 2 gL^{-1}, and 2.5gL^{-1}, the two times samples were prepared for the same concentration, in total 26 urea-tainted milk samples. Urea reagent was of analytical grade; prior to the experiment, milk samples adulterated with different concentrations of urea were manually stirred to make them homogeneous. All milk samples were not subjected to pre-treatment and dilution.

2.2 Instrumentation and Electrodes

Electrochemical station CHI660E, which is provided by Shanghai Chenhua Company, was used to collect response current. The electronic tongue was developed using an array of electrodes , including three kinds of metals - gold (Au), platinum (Pt) , and palladium (Pd), a platinum counter electrode, and a reference electrode (Ag/Agcl , 3M, Chenhua instrument inc. Shanghai, China). Working electrodes and the counter electrode（Tianjin Aida hengsheng inc. Tianjin, China）adopt 2mm diameter of disk electrodes and have a purity of 99.99% with the exception of gold electrode with a purity of 99.999%.

2.3 Voltammetry measurement

In this paper, the cyclic voltammetric waveform was adopted as excitation signal, the voltage range of 0-1.6V for the all working electrodes. All current response plots were obtained by using a scan rate of 50mV/s, sample interval 0.005V, and sample quiet time was 2s. The choice of the potential windows was based on gaining a large quantity of information related to samples measured and preventing irrelative electrochemical process to the measured samples [12].

After each measurement, reference electrode was placed in a 3M KCl aqueous solution, and the disk surfaces of working electrodes and the counter electrode were mechanically polished for 15 seconds with aluminum oxide of 1 and 0.3μm on a microcloth polishing pad consecutively. Thereafter, working electrodes and counter electrode were sonicated (Ultrasonic cleaner KQ5200DE) in ethanol: water (1:1) for 5 minutes. Finally, all electrodes were thoroughly rinsed with ultrapure water (milliQ, 18.2MΩ) to be prepared for the next measurement [13]. In cyclic voltammetry, when the potential is applied between the working electrodes and the reference electrode, the response current is collected between the working electrodes and the counter electrode. In the present work, the response current values from each of the three working electrodes were considered for the analysis. All voltages mentioned are versus the Ag/Agcl (3M KCl solution) electrode.

3. MULTIVARIABLE DATA ANALYSIS

A major difference between traditional electrochemistry and the electronic tongue is that traditional electrochemistry only collects redox peak current data while the electronic tongue measures evenly over the whole waveform [23]. In this paper the response currents of the working electrodes immersed in the adulterated and unadulterated milk samples have been collected for data analysis. The number of data points from each working electrode is 640 for each tested sample. The two-dimensional matrix for 52 samples of pure milk and urea-tainted milk and three working electrodes consists of 52×1920 data points, where 52 is the number of tested samples and 1920 is the total number of current response values of the three working electrodes. All the data have been processed and analyzed with Matlab software(Version 7.0, the Math works Inc.). Least square support vector machine (LSSVM) were constructed by the LSSVM toolbox (http://www.esat.kuleuven.be/ sista/lssvmlab/).

3.1 Preprocessing of the raw data

The current responses data obtained from the working electrodes are preprocessed by the following methods. In this paper, the mathematical expressions were adopted in Eqs. (1), Eqs. (2), respectively. Raw data matrix was normalized in Eqs. (1). In Eqs. (2), firstly, the raw data was centered, and then it was normalized. That's because that drift in the electrodes may cause the difference between the measurements of the different samples, which can be reduced by data centered. The method in Eqs. (1) has been found to be an appropriate technique in PCA for urea-tainted milk discrimination. For the LSSVM classifier, the method in Esq. (2) has been applied for preprocessing.

$$A_{ij} = \frac{M_{ij}}{\max(M_{ij})} \tag{1}$$

$$A_{ij} = \frac{M_{ij} - mean(m_j)}{\max\left[M_{ij} - mean(m_j)\right]} \tag{2}$$

Where M is the response current data matrix, Mij is the element of response current data matrix; mj contains current values of all samples in the same potential from the same electrode.

3.2 Principal component analysis

It is well known that the drawback of voltammetric sensors detection is the generation of a large number of experimental data, which hinders their use. In order to remove redundant information and to enhance signal-to-noise ratio of experimental data, the principal component analysis was adopted. Principal component analysis is an unsupervised statistical technique, which can be used in data preprocess and classification. The PCA method was achieved by singular value decomposition, as equation (3).

$$A = U \times S \times V^{-1} \tag{3}$$

$$T = U \times S \tag{4}$$

$$L = V^{-1} \tag{5}$$

Where T is score matrix, which can reflect the difference between samples; L is load matrix, which can reflect the effect of each variable. There are two functions using PCA, one is PCA can reduce the immense data set without losing much information; the other is the score plot that PCA generated can visualize differences between samples measured [24].

3.3 Least square support vector machine

The support vector machine (SVM) has been used successfully in solving many pattern recognition problems as a powerful classifier. The LSSVM is similar to the SVM to implement structural risk minimization according to the margin-maximization principle. For a complicated two-spiral classification problem it is illustrated that a LSSVM with radial basis function (RBF) kernel is readily found [25]. RBF kernel function is given in equation (6).

$$k(x_i, x_j) = \exp(-\frac{\| x_i - x_j \|^2}{2\sigma^2}) \quad (6)$$

The LSSVM classifier can be constructed to separate two classes, as equation (7):

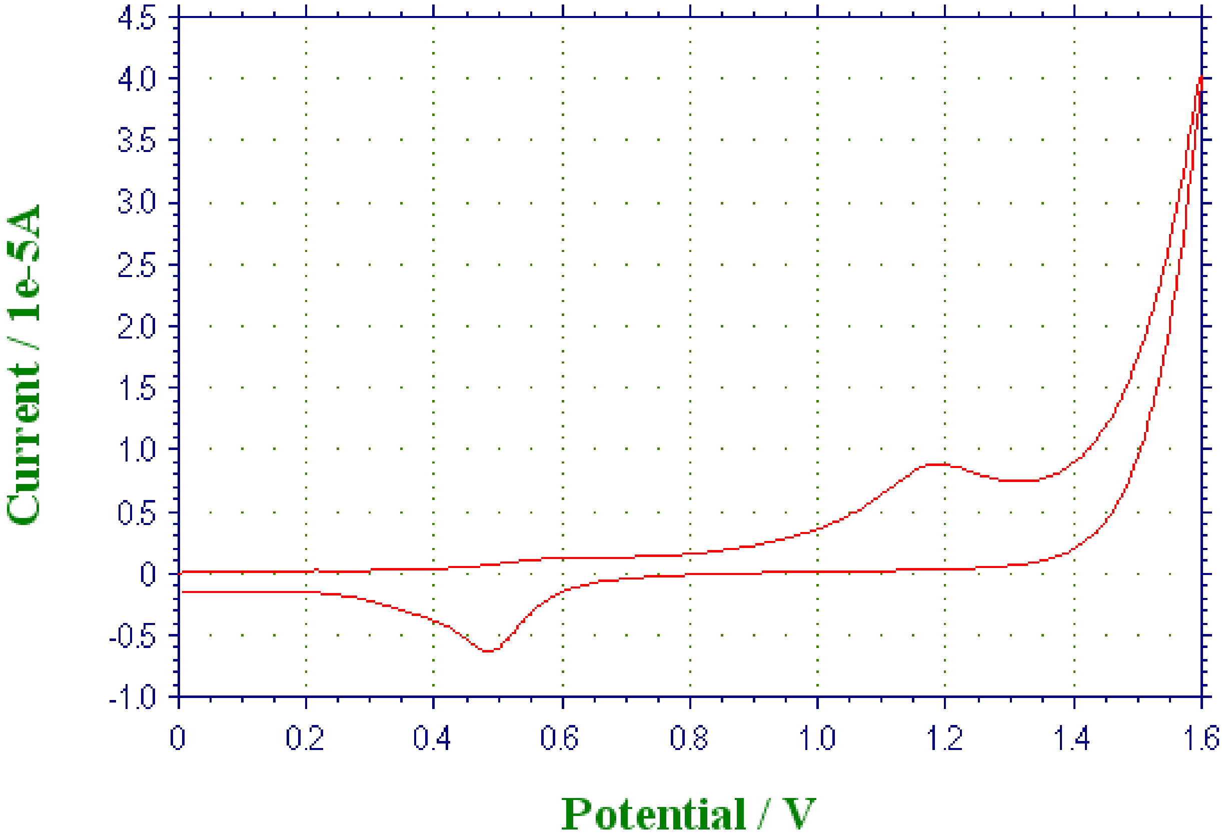
$$y(x) = sign[\sum_{k=1}^{N} a_k y_k K(x, x_k) + b] \quad (7)$$

In there, x_k is kth input and y_k is the kth output pattern, which are belonging to the calibration set; a_k is the Lagrange multipliers; b is real constant, y(x) is output for prediction set x.

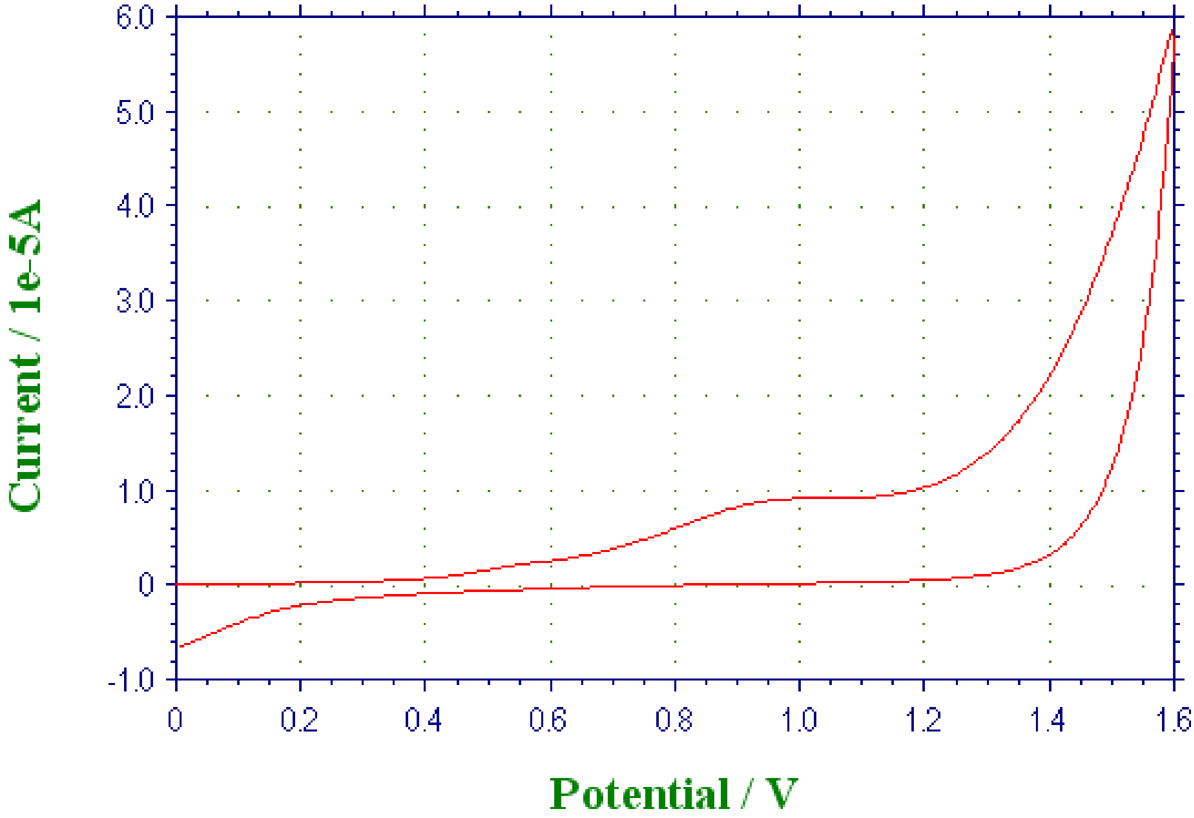
4. RESULTS AND DISCUSSION

4.1 Current responses

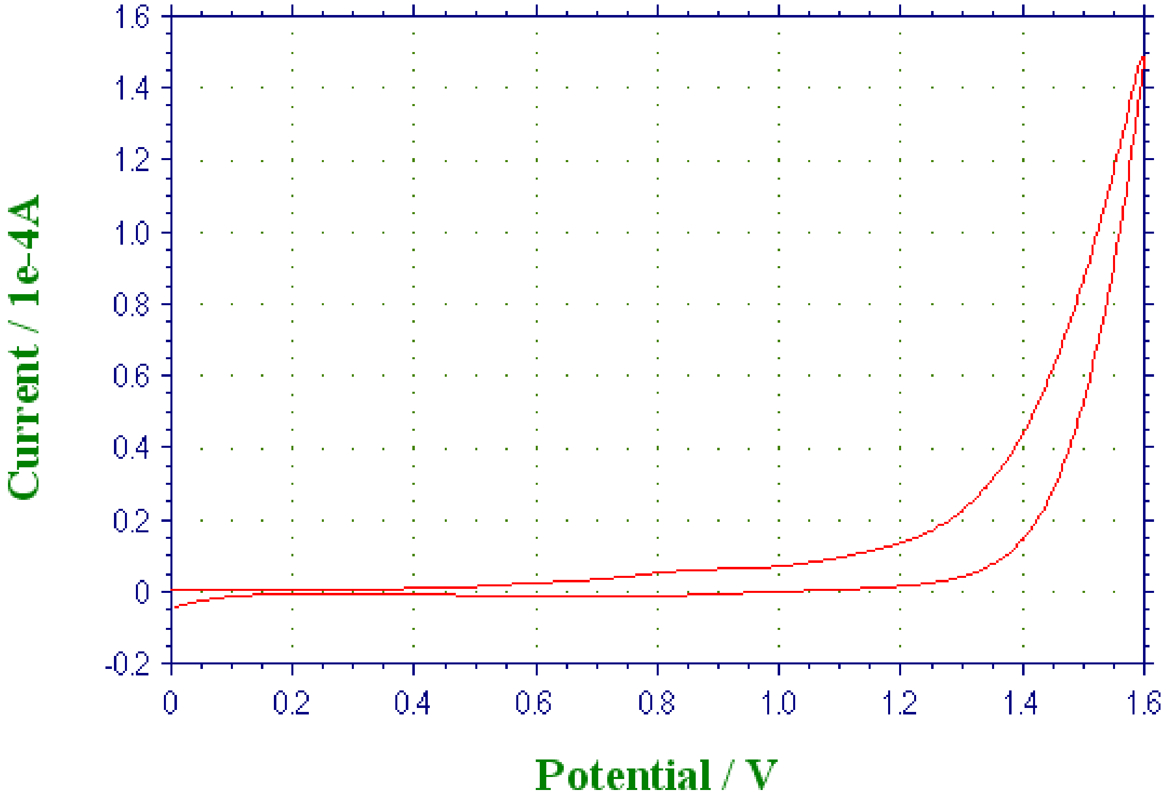
Pure milk samples and urea-tainted milk samples were analyzed using cyclic voltammetry without any samples pretreatment. The three working electrodes gave very similar voltammograms shapes between the unadulterated milk and milk samples adulterated with urea, respectively. In Figure1a-c, the voltammograms of the unadulterated milk sample using gold, platinum, and palladium electrode are shown, respectively. In Figure1a, there was one anodic peak at around +1.18V as well as one cathodic peak at around +0.48V on the voltammogram of the gold working electrode. In Figure1b, only one anodic peak at around +0.9V was shown on the voltammogram of the platinum working electrode. Besides, in Figure1c, there was no obvious peak on the voltammogram for the platinum working electrode.


Current / 1e-5A
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0
-0.5
-1.0
0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
Potential / V

(a)


Current / 1e-5A
6.0
5.0
4.0
3.0
2.0
1.0
0
-1.0
0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
Potential / V

(b)



(c)

Figure 1. Raw transient current response of the 3-voltammetric sensors array: (a) gold, (b) platinum, and (c) palladium in unadulterated milk samples

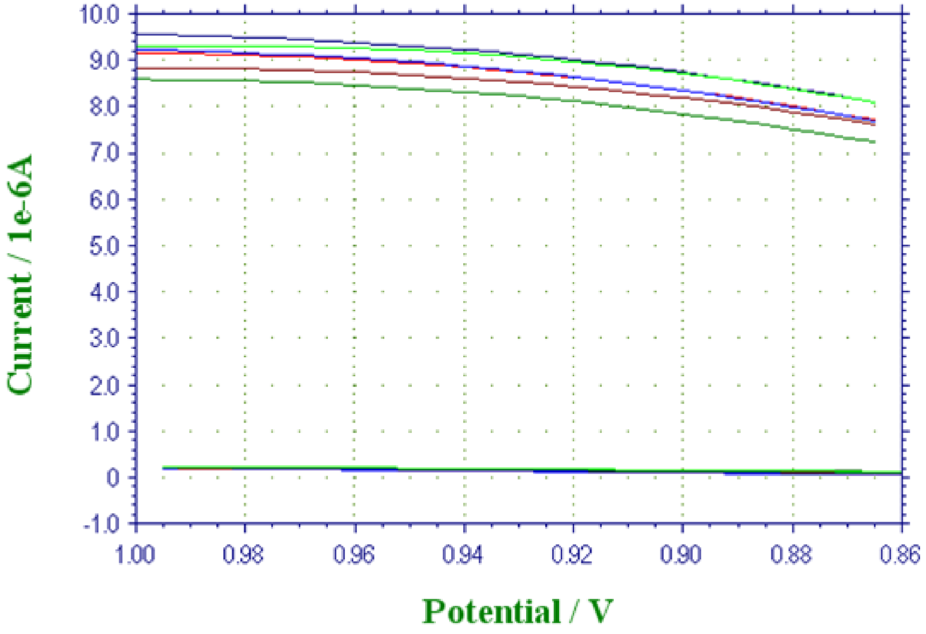


Figure 2. Amplified local current response overlaid curves from the platinum electrode in unadulterated milk samples and milk samples adulterated urea with 300mg/L, 1g/L concentration

Because of the small difference of current response curves between pure milk samples and urea-tainted milk samples, we amplified current response plots. Figure 2 shows the overlaid plots of the amplified local current response curves with voltage range of 0.86~1V from the platinum electrode. The upper two curves are current response curves of pure milk samples, while the middle two curves are current response curves of 300mg/L urea-tainted milk samples, and the two curves at the bottom are current response curves of 1g/L urea-tainted milk samples. From Figure2, in contrast unadulterated milk samples have bigger current value than the ones found for the milk adulterated with urea; 300mg/L urea-tainted milk samples have bigger current than 1g/L urea-tainted milk samples, which is due to passivation of the electrode surface by urea adulterant.

4.2 Classification of the Urea-tainted Milk samples by PCA

Voltammograms of the three working electrodes were recorded between 0 and 1.6V, respectively. The number of current response captured from each working electrode is 640, while response data matrix is 52×640. The two-dimensional matrix of total data from the three working electrodes is 52×1920.

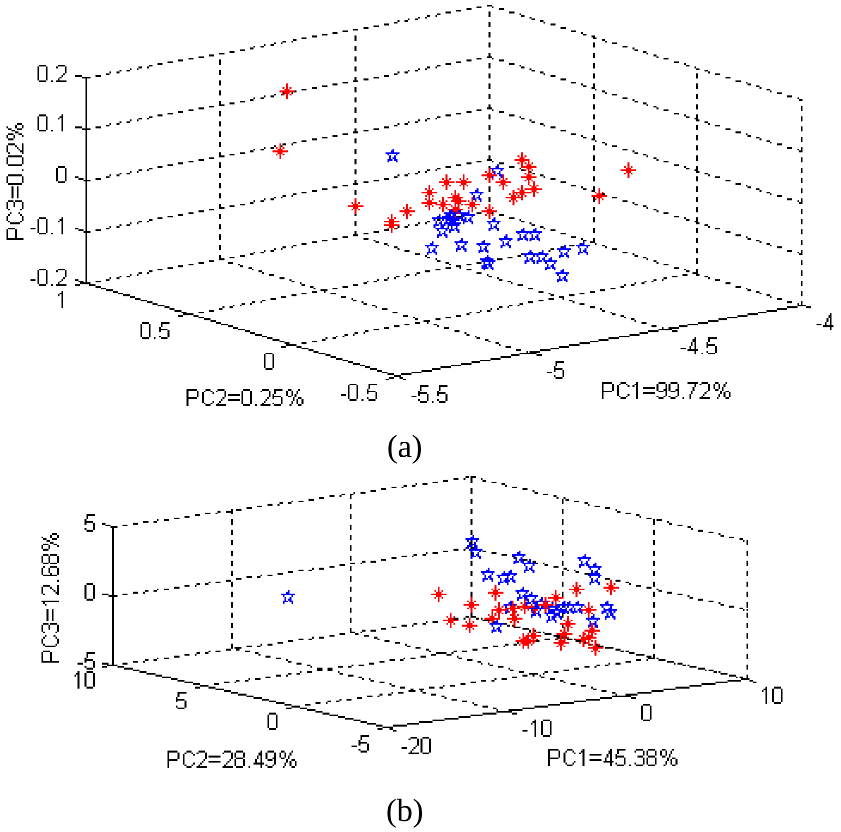


Figure 3. PCA score plot of unadulterated milk samples and milk samples adulterated with urea (blue ☆ represents urea-tainted milk sample, red * represents unadulterated milk sample), (a) PCA score plot of response data of the Pt electrode by normalization of raw data; (b) PCA score plot of response data of the three working electrodes by normalization of data center

The score plot of unadulterated milk samples and milk samples adulterated with urea as shown in Figure3. In the three working electrodes, the current response of the Pt electrode has best stability and reproducibility. In Figure3 (a), the raw data of the Pt electrode was normalized by the maximum value of raw data matrix. The score plot of the Pt electrode shows the best separation ability in all score plots from current response data of each working electrode. PC1, PC2, and PC3 are 99.72%, 0.25%, and 0.02% respectively, i.e., the first three principal components (PCs) accounted for 99.99% of the current response variation. There was a tendency that fifty-two samples from unadulterated milk samples and adulterated milk samples with urea were almost distinguished into two areas, as shown in Figure 3(a), but there are still three urea-tainted milk samples mixed into unadulterated milk samples. In Figure3 (b), the PCA score plot was gained after data centered of total response data from the Pt, Au and Pd electrodes were normalized. The separation ability is lower than that of Figure3 (a). The main reason is that the first three principal components are 45.38%, 28.49%, and 12.68% for PC1, PC2, and PC3, respectively, which only represent 86.55% of raw information. Because content of adulterant is low, the difference between unadulterated milk samples and urea-tainted milk samples is very small, it is not enough to discriminate the difference between unadulterated milk samples and adulterated milk samples with urea by the first three principal components. Therefore, in order to improve classification ability, more PCs are needed to obtain different characteristic information of measured samples. Thus, LSSVM discrimination model was employed by using total response data from three working electrodes to improve separation ability based on the result of PCA analysis.

4.3 PCA-LSSVM Model

In this paper, fifty-two measured samples were divided into two parts: calibration set and prediction set. The calibration set contained 35 samples including 17 pure milk samples and 18 urea-tainted milk samples, which accounted for two-thirds of total samples. And the remaining one-third of the samples had 17 samples including 9 pure milk samples and 8 urea-tainted milk samples, which were used to construct the prediction set. In the calibration set and prediction set, the unadulterated milk samples were set to “-1” and urea-tainted samples were set to “1”.

It is important to select proper number of the principal components, if the number of the principal components is less, the difference between unadulterated milk samples and urea-tainted samples can not be reflected fully; but if the number of the principal components is more, the noise information can be added. According to current response result, the number of these current attributes is 1920 from the three working electrodes for the each sample. When number of these features was reduced to 10 from 1920 by using PCA, the classification effect of measured milk samples is best. LSSVM model was constructed using the scores matrix of the first ten principal components reduced by PCA as new input variables [26]. The LSSVM classifier includes two parameters, which are the relative weight of the classification error γ and the bandwidth of the RBF kernel function σ. They are optimized for building a better classification model. In this study, γ and σ^2 were determined by performing a grid-search technique, and the process of two-step search for optimizing the parameters γ and σ^2 is shown in Figure 4, where $\gamma= 438.78$ and $\sigma^2=106.87$ were selected to build the LSSVM model.

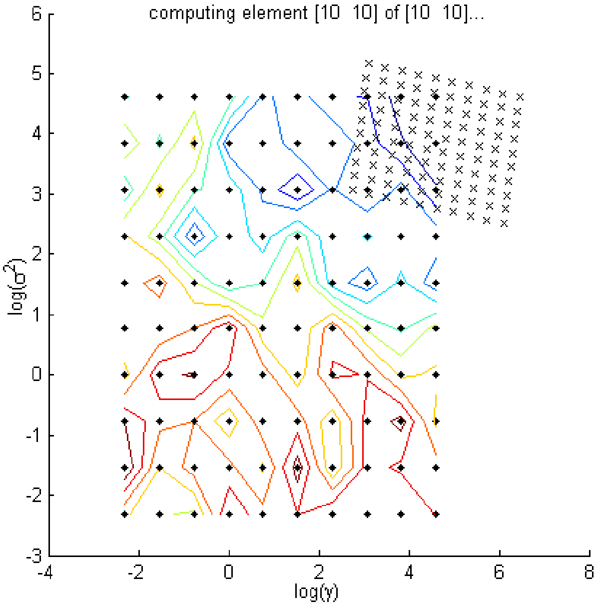


Figure 4. Process of two-step search for optimizing the parameters γ and σ^2

The predicted output targets are set “-1” and “1” for unadulterated milk samples and urea-tainted milk samples in calibration set, the threshold was set to 0. The LSSVM model output value corresponds to the predicted value, the samples with predicted value $y_p<0$ for the samples belong to unadulterated milk samples and $y_p>0$ belong to urea-tainted milk samples.

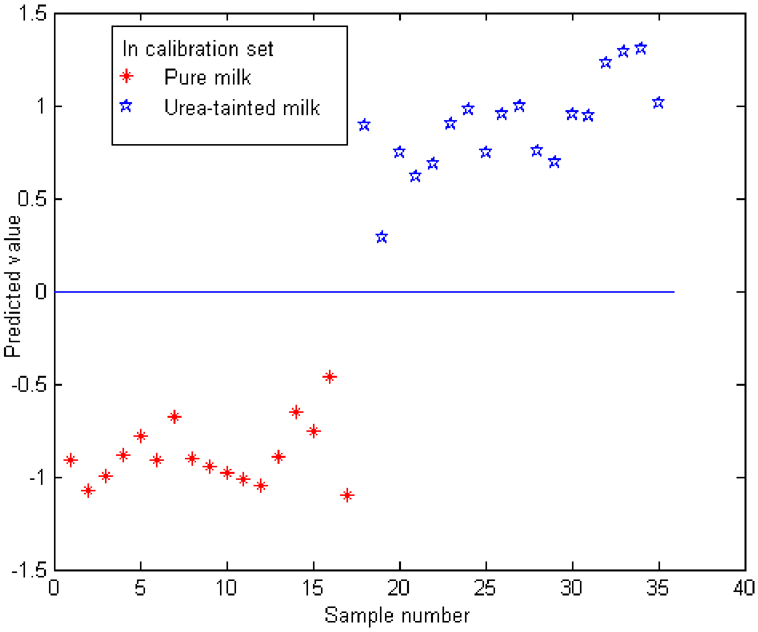


Figure 5. Predicted results in calibration set by PCA-LSSVM model by total response data of three working electrodes

The predicted results in calibration set were obtained using PCA-LSSVM model, as shown in Figure 5. The unadulterated milk and urea-tainted milk samples were properly classified into two chasses. 17 samples in prediction set were predicted by the PCA-LSSVM model, whose predicted results were shown in Figure 6.

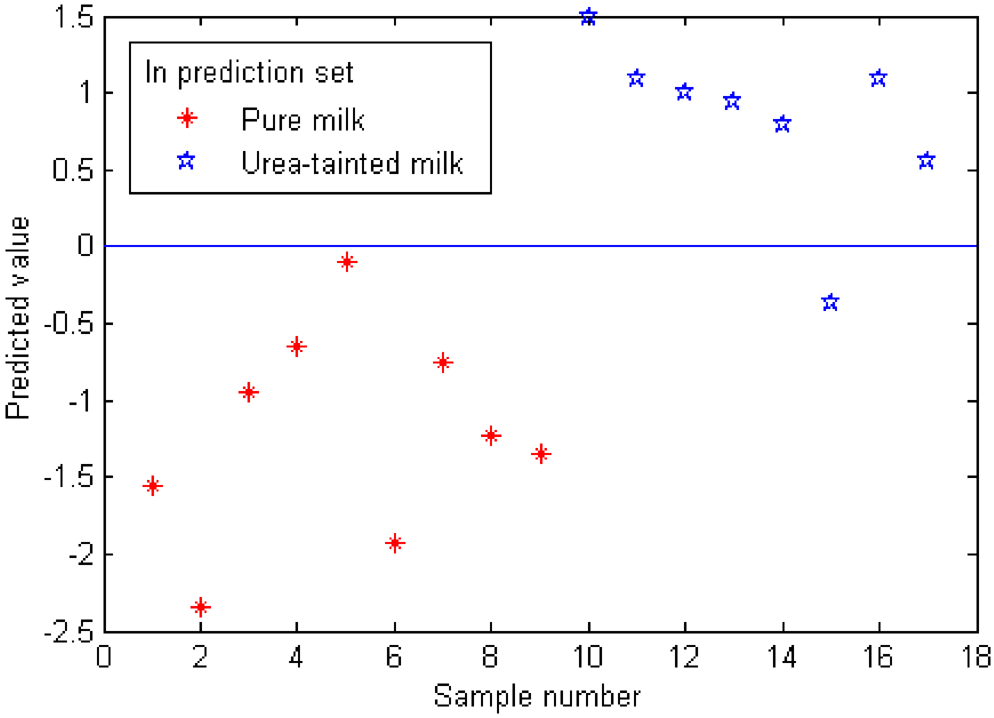


Figure 6. Predicted results in prediction set by PCA-LSSVM model of total response data from three working electrodes

All unadulterated milk samples were properly classified. There was one urea-tainted milk sample that was not correctly identified in predicted samples, so the ratio of correct classification reached 94.12% in prediction set. The wrongly predicted sample is urea-tainted milk sample with concentration of 200mg/L, which is minimum concentration of milk adulterated with urea. Reference [12] proposed that the voltammetric electronic tongue can discriminate urea-tainted milk at concentration above 4.16mmol·L^{-1}(249.8mg·L^{-1}) combined with PCA and HCA (Hierarchical Cluster Analysis). The detection limit is higher 121.4mg·L^{-1} reported by reference [13] combined with f-PCA. But they only discriminated between unadulterated milk samples and adulterated milk samples with different content of urea by unsupervised pattern recognition methods. In this paper, the detection limit for urea is higher than those proposed in those work, but it is possible to detect content of urea in milk samples below intake dose recommended by the health organization and PCA-LSSVM model can be used to predict whether unknown milk samples are adulterated with urea or not.

Receiver operating characteristic (ROC) curve of response data of the three working electrodes is shown in Figure 7, where the horizontal ordinate represents specificity, and the vertical ordinate represents sensitivity. The total ROC area is used to evaluate the classification performance of the discriminant model [27]. When an area under the ROC curve is less than 0.5, the discriminant model is worthless. An area above 0.9 represents a good classifier. For this work, the area under the ROC curve is only 0.847 when response current data of the Pt electrode is adopted; the area under the ROC curve is 0.986 when the total response current data of three working electrodes was used as analysis data, which indicated the LSSVM model constructed by total response data of three working electrodes was more "excellent" at separating adulterated milk samples from unadulterated milk samples.

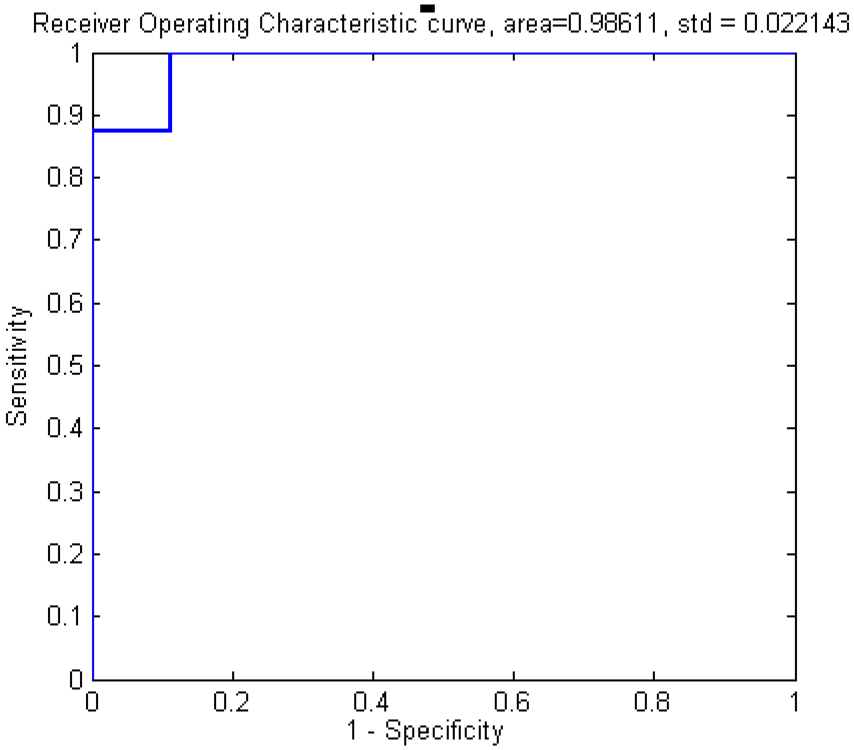


Figure 7. Receiver operating characteristic（ROC）curve of the classifier developed by PCA-LSSVM of the response data from the three working electrodes.

5. CONCLUSIONS

A new approach based on voltammetric electronic tongue in combination with PCA-LSSVM was developed for discrimination between urea-tainted samples and unadulterated milk samples. When the total response current data of three working electrodes (Au, Pt and Pd) were adopted, the correct classification rates were 100% for calibration set, 94.12% for prediction set, and were 98.08% for all samples. The wrongly predicted sample is the urea-tainted milk sample with minimum concentration of 200mg/L. The results show that the proposed method is feasible for discriminating urea-tainted milk.

ACKNOWLEDGMENTS
This research was performed with financial support from Natural Science Foundation of Tianjin (13JCYBJC25300), Innovative Talents Training Plan of Young and Middle- aged Backbone in Tianjin City University (20131213), Innovation Team Training Project of Tianjin University (TD12-2019), Tianjin college students' innovation and entrepreneurship training program (201510061093) and the Second Level Candidates Training Project of Tianjin City 131 Innovative Engineering (20131203).

References

1. S. Das, M. Sivaramakrishna, K. Biswas, B. Goswami, *Sens. Actuators, A* ,167(2011) 273-278.
2. J.C. Moore, J.W. de Vries, M. Lipp, J.C Griffiths, D.R. Abernethy, *Compr. Rev. Food Sci. Food Safe,* 9(2010) 330-357.
3. E.F. Renny, D.K. Daniel, A.I. Krastanov, C.A. Zachariah. *Biotechnol Biotec Eq,* 19(2005) 198-201.
4. G. N. Nikoleli, D.P. Nikolelis, C. Methenitis. *Anal Chim Acta* , 675(2010) 58-63.
5. U.B. Trivedi, D. Lakshminarayana, I.L. Kotharib, N.G. Patel, H.N. Kapsed, K.K. Makhija, P.B. Patel, C.J. Panchal. *Sens. Actuators B,* 140 (2009) 260-266.
6. M. Singh , N. Verma, A. K. Garg, N. Redhu. *Sens. Actuators B,* 134 (2008) 345-351.
7. X. Dai, Y. Zhao, M. Li, X. Fang , X. Li, H. Li, B. Xu. *Research Gate,* 4(2013) 1557-1565.
8. D. I. Ellis, V. L. Brewster, W. B. Dunn, J. W. Allwood, A. P. Golovanovc, R. Goodacre. *Chem. Soc. Rev,* 41(2012) 5706-5727.
9. R. J. Yang, R. Liu, K.X. Xu. *Food Bioscience,* 2 (2013) 61-67.
10. R. J. Yang, W. Y. Zhang , Y. R. Yang , Z. C. Wu , G. M. Dong , Y. H. Du. *Anal Lett,* 47 (2014) 2560-2569.
11. L.A. Liu, Y.P. Yu, J. J. Yang, R. J. Yang, G.M. Dong, T.M. Jin, *Int. J. Electrochem. Sci,* 10(2015) 5970-5980.
12. L. Bueno, W.R. de Araujo, M. O. Salles, M.Y. Kussuda, T. R. L. C. Paixão. *Chemosensors,* 2 (2014) 251-266.
13. A. Hilding-Ohlsson, J. A. Fauerbach, N.J. Sacco, M. C. Bonetto, E. Cortón. *Sensors_Basel,* 12(2012) 12220-12234.
14. Q.S. Chen, J.W. Zhao, S.Vittayapadung. *Food Res Int,* 41(2008) 500-504.
15. A. Ghosh, B. Tudu, P. Tamuly, N. Bhattacharyya, R. Bandyopadhyay. *Chemometr Intell Lab,* 116(2012) 57-66.
16. F. Winquist, J. Olsson, M. Eriksson. *Anal Chim Acta,* 683(2011) 192-197.
17. X. Cetó, J.M. Gutiérrez, A. Mimendia, F. Céspedes, M. D.Valle. Electroanalysis, 6(2013) 1635-1644.
18. X. Cetó, J. M. Gutiérrez , M. Gutiérrez , F. Céspedes , J. Capdevila , S. Mínguez, C. Jiménez-Jorquera, M. del Valle. *Anal Chim Acta,* 732 (2012) 172-179.
19. K. Tiwari, B. Tudu, R. Bandyopadhyay, A. Chatterjee. *J Food Eng,* 117 (2013) 205-210.
20. L. Bueno, T. R. L. C. Paixão. *Int. J. Electrochem.*, 2012 (2012) 1-5.
21. Y.P. Yu, H. Zhao, R.J. Yang, G.M. Dong, L.A. Li, J.J. Yang,; T.M. Jin , W.Y. Zhang, Y. Liu. *Int. J. Electrochem. Sci.,* 10 (2015*)* 4381-4392.
22. F. Winquist . *Microchim Acta,* 163(2008) 3-10.
23. P.Ivarsson, S.Holmin, N.E. Höjer, C. Krantz-Rülcker, F. Winquist. *Sens. Actuators B,* 76(2001) 449-454.
24. X. Cetó, F. Céspedes, M.D. Valle. *Microchim Acta,* 180(2013) 319-330.
25. J.A.K. Suykens, J. Vandewalle. *Neural Process Lett,* 9(1999) 293-300.
26. D. Calisir, E. Dogantekin. *Expert Syst Appl,* 38 (2011) 10705-10708.
27. R. V. Vardhan, S. Pundir, G. Sameera. *Bonfring Intern. J. Data Mining,* 2(2012) 52-56.