Full Length Research Paper

Classification of tea consumed in Senegal using XRF techniques and chemometric based on their country of origin

Modou Mbaye, Alassane Traoré*, Ababacar Sadikh Ndao and Ahmadou Wagué

Department of Physics, Faculty of Sciences and Techniques, University Cheikh Anta Diop, UCAD, Applied Institute of Nuclear and Technology, BP 5005. Dakar, Senegal.

Accepted 30 October, 2013

Discrimination of tea varieties from different countries is demanding task and to address this requirement, we demonstrate the feasibility of using x-ray fluorescence (XRF) spectrometry combined with chemometric, to classify geographical origin of different based on their elemental contents. Both direct non-pre-treatment XRF spectra and elemental concentration were used to achieve discrimination of tea. The classification was carried out on the basis of chemical information contained in tea samples. In total, 13 elements (Mg, Al, P, S, Rb, Sr, K, Ca, Mn, Fe, Cu, Zn, As, Pb) were used as chemometric descriptors for classification purpose of tea types based to their geographical origin. Different pattern recognition techniques such as singular value decomposition (SVD), hierarchical clustering analysis (HCA) and artificial neural network (ANN) were applied to differentiate tea varieties. The most dominant element descriptors features were calcium Ca, potassium K, manganese Mn and iron Fe.

Key words: Portable X-ray fluorescence, tea, singular value decomposition; hierarchical clustering analysis, artificial neural network.

INTRODUCTION

Tea is one of the most popular beverages in the world, especially in Africa and Asia where people started drinking tea thousands of years ago (Stragg and Millin, 1975). Tea contains flavonoids, minerals, and trace elements that are essential to human health. There are broad scientific publications about medicinal value and beneficial health effects of tea (Dufresne and Farnworth, 2001; Weisburger, 1997). Some of the possible beneficial effects of drinking tea are: anti-oxidative activity (Luczaj and Skrzydlewska, 2005), immune system boosting (Mark, 2007), protective effect against a range of cancers including lung, prostate, and breast cancer (Siddiqui et al., 2005; Record and Dreosti, 1998) and reduction of

blood cholesterol levels (Fujita and Yamagami, 2008) effective drug against Plasmodium (Ferreira and Janick, 1995).

Zinc is an essential component of over 200 enzymes having catalytic and structural roles. Zn deficiency is characterized by recurrent infections, lack of immunity performance and poor growth. Low intake of zinc may cause coronary artery disease. Clinical materials prove that Zn can have good effect on eliminating ulcer and promoting healing wounds. Following zinc and iron, copper is the third most abundant trace element in the body. Copper is involved in the oxidation of Fe²⁺ Fe³⁺ during haemoglobin formation. It is an important catalys

for iron absorption. Copper deficiency may be risk factor for cardiovascular disease. When over copper deficiency occurs, symptoms include neutrogena, cardiac disorders, osteoporosis and anaemia. Iron is needed for a healthy immune system and for energy production. Essential element as halogen Cl, among others, is employed as fingerprints in the diagnosis of neoplastic tissues (Fung et al., 2003).

Due to the importance of minerals in tea, several studies including atomic absorption spectroscopy (AAS) were performed (Marcos et al., 1998). However inductively coupled plasma mass spectrometry (ICP-MS) is the most commonly used techniques of differentiation of tea other ICP (Yingxu et al., 2009) and others related techniques such as liquid chromatography, electronic nose (Dutta et al., 2003), electronic tongue (Ivarsson et al., 2001), near infrared reflectance (Chen et al., 2007) and infrared spectroscopy (Budinova et al., 1998) can be found as well. These kinds of techniques are very time consuming due to various sample preparation steps. However elemental fingerprint techniques, based on elemental composition and multivariate analysis, has primarily been used to determine the geographic origin of teas (Fernandez-Caceres et al., 2001). Based on tea sensory properties such as flavor, texture, smell and perceived color tea samples from different countries are difficult to classify due to their leaves similarities and their environmental impact. The most common procedures used for the classification of teas brands and determination of the geographic origin of different teas include several pattern recognition techniques, such as principal component analysis (PCA), linear discriminant analysis (LDA), cluster analysis (CA), soft independent modeling of class analogy (SIMCA) and artificial neural networks (ANNs) (Herrador Gonzalez, 2001). However, in our knowledge, there are no previous studies on differentiation of tea types from different countries using no time consuming portable energy X-ray fluorescence analyzers.

In this present work thirteen tea samples from Cameroon, China and Luxembourg were analyzed by fluorescence energy dispersive x-ray (EDXRF) spectroscopy for the purpose of using elemental composition of tea to discriminate among geographical origin. Elemental content (Al, As, Ca, Cu, Fe, K, Mg, Mn, P, Pb, Rb, S, Sr and Zn) were determined and major elements are chosen as variable predicators. The chemometric procedure performed were singular value decomposition (SVD) for data exploration, hierarchical analysis HCA for identification discrimination and artificial neural network. These methods achieved 100% correct classification for the thirteen tea types from Asia China, Europa Luxembourg and Africa Cameroon. For the detection of present elements, X-ray fluorescence has been performed to evaluate environmental contamination of tea leave and the verification of the bio-accumulation potential. The classification chemometric model developed in this work

will be able to achieve an authentication of tea according to their geographical origin but also an authentication of brands falsification of tea for further study.

MATERIALS AND METHODS

Samples were dried-out in an oven at 80°C until constant mass was obtained. To be ready for energy dispersive X-ray fluorescence (EDXRF) analysis, the tea was ground to 120 µm using an agate mortar. In addition for homogenization and to avoid grain size effect in the quantitative analysis, a vibrating planetary mill was used to obtain to reduce grain size less than < 70 µm. One the advantageous features of XRF techniques is the possibility of performing analysis directly on solid samples, which usually implies simpler sample preparation with considerable reduction in usage of reagents. Determination of the composition of solid samples, without any sample preparation, is possible for samples that are homogeneous in all three dimensions and with flat surface. This is the case for direct analysis of metals and alloys that has been one of the main applications of XRF. However, most solid materials require sample pre treatment to make them homogeneous and to ensure the quality and the reproducibility of measurements (Tertian and Claisse, 1982). All The samples were prepared as a mixture of tea and Ethylene bis-Stearamide (known as Licowax or Hoechtwax HWC) used as binder agent, in a proportion of 1 g of tea sample to 10% of HWC of the mass of the sample with 0.13 g/cm³ of superficial density. The mixture was thoroughly homogenized once more in a planetary mill and pressed (under 5 tones) into a disc pellet of 32 mm diameter. The pellets were measured using an EDXRF portable analyzer spectrometer (Niton XL3t900s with Geometrically Optimized Large Drift Detector) with Ag x-ray excitation source and several filters as secondary targets for excitation. This configuration allows the attainment of improved sensitivity and signal to noise ratio by sequentially selecting appropriate combinations of filters as secondary targets and different groups of elements were carried out depending on element interest. These pellets were (three replicate measurements for each tea sample were performed) placed within the stand lid system. The use of portable analyzer for excitation and the characteristic x-rays emitted by the constituents of the samples were carried out for a period of 200 s (Table 1). The portable XRF uses both quantification techniques: normalization and the full fundamental method

However, the presence of some potential toxic elements in tea sample as lead (Pb), arsenic (As), chromium (Cr), cadmium (Cd) or aluminium (Al) constitutes a reason for concern. In the case of Cadmium we pay attention to the overlapping spectra between cadmium (Cd-L) and potassium (K) lines. For this purpose, the sandwich of molybdenum (Mo), Al and titanium (Ti) as filter was selected for Cd excitation and these additional filters with Ag was performed to improve the signal-to-noise ratio by reducing the measured continuum. On the other hand, Mo filter is the best choice, instead of silver direct excitation source alone, to attain larger signal-to-noise ratios.

The precision of the method was verified by using three different replicate samples due to the fact that measurements have been made in reproducible conditions. Table 2 shows the results of precision attained after the analysis of the INCT-Tea Leaves-1 standard. In this study enhancement effects and/or absorption of characteristic X-ray emitted by the sample constituents make the quantification of the detected elements difficult. On the other hand, the use on thin sample was carried out by assessing the proportion between area density and attenuation factor and assuming monochromatic excitation source (Karimi et al., 2008; Markowick, 1993; Salvador et al., 2002; Matsumoto et al., 1976). Tea name labelled, geographical origin and corresponding abbreviation are

Table 1. Spectrometer specifications and operating conditions.

Resolution	162.15 eV (910 cps Mn Ka)
Window Thickness	12.7 µm Be
Rating	50kV, 40 μA maximum power of the tube 2 W
Diameter of the beam	7 mm
Filters	List of element interest
Ag excitation source	Sb, Sn, Cd, Pd, Ag, Mo, Nb, Zr, Sr, Rh, As, Ti, Ni, Co, Fe, Mn, Cr, V, Se, Au, Pb, W, Zn, Ta

Table 2. Results of average and certified values INCT-Tea Leaves-1 (mg/kg).

Element	Certified value ± uncertainty	\overline{X}	$\sigma_{\scriptscriptstyle X}$	$(\frac{\sigma_X}{\overline{X}})*100$
Al	2290±3300	2400	250	10.4
Cd	0.03±0.004	ND		
Pb	1.78±0.24	15.6	0.6	3,8
As	0.106±0.021	ND		
Fe	432	593	78.5	13,2
K	17000±1200	8917.3	1692.1	19,0
Ca	5820±520	3145	576.1	18,3
Р	1810	2100	185	8.8
Mn	1570±110	1761.6	247.2	14,0
S	2470±250	2575.3	519.1	20,2
Sr	20.8±1.7	23.8	1.9	8,0
Rb	81.5±6.5	89.2	7.5	8,4
Cu	20.4±1.5	33.3	3.8	11,4
Zn	34.7±2.7	51.9	5.5	10,6
Mg	2240±170	N.D [*]		

^(*) Not Detected.

given in Table 3.

Multivariate statistical analysis

Singular value decomposition (SVD)

SVD is a chemometric method, where an orthogonal linear transformation is applied on the XRF spectra data. The original spectra data can be represented by a few principal components (PCs) (Mei et al., 2012; Varmuza and Filmoser, 2009). The dataset was reduced by first applying a linear combination of the standardized original variables possessing the greatest possible variance, thereby creating the first principal component (PC1). The second component (PC2) was created based on the linear combination of the standardized original variables having the greatest possible variance and was uncorrelated with all previous defined components. This procedure facilitates the extraction of important chemical information and leads to substantial reduction of redundant data.

Hierarchical clustering analysis (HCA)

Unsupervised HCA is a widely used data analysis tool in applied

spectroscopy in combination with SVD. In this present work HCA was used to classify tea while no prior information about any group membership is available, and not even the number of clusters is known. BLANK tea was considered as an unknown tea sample from china, thus the predict chemometric model developed in this work achieved correct classification to unknown tea sample.

Artificial neural networks (ANN)

ANN derived from biological nerve systems, have been used in this paper for both classification and clustering. The core technique of ANNs is neural net (or network in short) that consists of layers of neurons (nodes or units) and links (or interconnections) between neurons on neighbouring layers for nonlinear computational (Karim et al., 2008). These pattern recognition techniques SVD, HCA and ANN were perform with software in version of R and Matlab.

RESULTS AND DISCUSSION

Reference data from elements

The average values for each element were calculated

Sample	Label	Commercial area	Growing area	
TEA A	CMR	Cameroon	Cameroon	
TEA B	LUX	Luxembourg	Luxembourg	
TEA 1	FLECHA	Senegal	China	
TEA 2	CHEVAL	Senegal	China	
TEA 3	FORCE	Senegal	China	
TEA 4	FORTE	Senegal	China	
TEA 5	SAVA	Mauritania	China	
TEA 6	EL HELLA	Mauritania	China	

Table 3. Commercial green tea leaves samples selected in this study and their growing area.

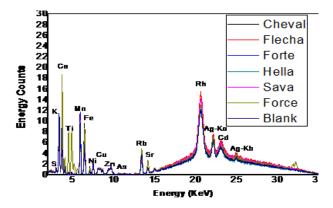


Figure 1. Energy dispersive X-ray fluorescence and scatter spectra of 05 tea samples from China at 35KV.

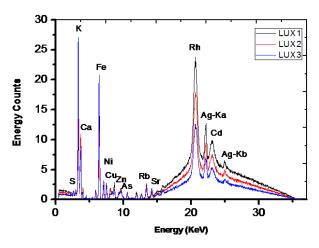


Figure 2. Energy dispersive X-ray fluorescence and scatter spectra of 03 tea samples from Luxembourg (LUX) at 35 KV.

only for those elements for which results had very high concentrations of interest. The values of Se, W, Ni and Co are below the detection limit. The cases presenting missing or wrong values are excluded. The concentrations of rare earth elements (Dy, Eu, Gd, Sm, Tb and Yb) were not detected in any tea leaves. The list of element determined includes macro, micro, toxic and trace elements in tea are presented in Table 4. Along these macro elements, K has the higher concentration in

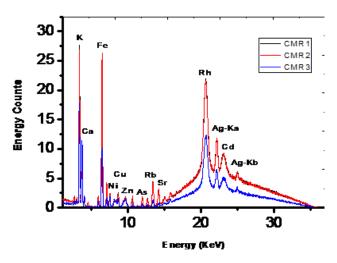


Figure 3. Energy dispersive X-ray fluorescence and scatter spectra of 03 tea samples from Cameroon (CMR) at 35KV.

TEA В. Due to the experimental setup, measurements were carried out under air instead of vacuum or helium flushes. For light elements with the concentration slightly higher to the detection limit it becomes scarce to get good sensitivity and we only detected Mg in this tea. In some tea samples Ti was detected, but is not an essential element for the plant growth. The concentration of Zn in TEA 6, TEA 5 and TEA 3 is the same. This elemental tea variability can be explain in fact that elemental content of teas depends on several factors, including soil characteristic and environmental conditions e.g. mineral composition, cation exchange, particle size, temperature, rainfall and air quality.

Singular value decomposition (SVD)

Prior to the application of SVD, the data were autoscaled, auto-scaling procedure consists, for each value, to the subtraction of the variable mean. The obtaining result is a new variable with zero mean. The original data matrix (Figures 1, 2 and 3) is transformed into a new

Table 4. Element concentration in leaves of thirteen teas presented as mean ±SD (n=3).

Elements (mg.kg-1 dry)	TeaA	TeaB	Tea1	Tea2	Tea3	Tea4	Tea5	T ea6
Al	12830±5780	23400±6660	<8040	10830±3470	18470±5390	10920±5010	8950±4350	<8040
As	<20	30±10	<20	<20	<20	<20	<20	<20
Ca	93540±1040	64850±890	35480±490	90710±890	121480±1110	87340±950	45220±580	50550±670
Cu	250±30	100±20	160±20	180±30	170±30	180±30	190±30	180±30
Fe	6280±190	12200±260	1730±110	4520±160	4340±170	2950±160	1600±110	1870±130
K(x10 ³)	219.7±1.64	298.74±2.11	134.04±1.16	119.27± 1.03	116.97±1	141.54±1.24	147.45±1.25	163.41±1.43
Mg	<88140	89690±43680	<65250	<70010	<73970	<80010	<69050	<79960
Mn	1410±100	1780±110	7300±220	9660±260	8700±260	11940±320	9060±250	10170±290
P(x10 ³)	18.97±0.53	15.18±0.45	14.36±0.5	11.88±0.46	13.52±0.46	15.61±0.53	15.49±0.51	14.78±0.54
Pb	50±10	100±10	<20	<20	<20	<20	<20	<20
Rb	50±10	100±10	100±10	130±10	190±10	150±10	130±10	160±10
S(x10 ³)	24.82±0.43	21.91±0.4	25.94±0.45	19.95±0.4	26.11±0.42	25.18±0.46	25.9±0.45	27.06±0.5
Sr	20±10	60 ±10	30±10	40±10	90±10	40±10	40±10	40±10
Zn	270±20	280±20	130±20	140±20	150±20	140±20	150±20	150±20

coordinate system. The results of singular value decomposition are summarized in Figure 4. The first three extracted components accounted for 99.29% of the total variability of the data in Figure 4. The largest variability in the data set (89.31%) is accounted for by transition and alkali elements *Ca, K,* Mn and Fe which showed the highest weights in the loadings plot in Figure 5. Whereas other group of elements tea contents Mg, Al, P, S, Cu, Zn, As and Pb are associated with less dominant features discriminate for tea classification. Chemically the first two PCs can probably associate of influence of soil composition (Ca, K and Fe).

Hierarchical clustering analysis (HCA)

Hierarchical cluster analysis was applied in the present study to achieve classification. A data matrix was built which consist of the concentration

of thirteen elements in each tea sample. This data matrix of concentration was subject to hierarchical cluster analysis. Thus, elementals content data were used as input to perform the characterization. The similarities between samples were calculated on the basis of the squared Euclidean distance, whereas a hierarchical agglomerative system was used to establish the cluster after redundant information was removed with SVD. The results obtained are presented in Figure 6 as can be seen in the dendrogram representation, three main clusters with 100% accurate grouped sample based on their geographical area are obtained: Cameroon (CMR1, CMR2 and CMR3), China (ELHELLA, BLANK, FORTE, FLECHA and SAVA) and Luxembourg (LUX1, LUX2 and LUX3).

For direct XRF spectra analysis by chemometric without pre-treatment to achieve unsupervised classification based on hierarchical cluster analysis, a data matrix was built which consist of

511 spectral data and 47 tea samples. This data matrix of 47 rows and 511 columns (511 channels) was subject to singular value decomposition (SVD) for dimension reduction and cross- validation to obtain the optimal component for data matrix reconstruction. Figure 7 illustrates the dendrogram plot of direct XRF spectra and as shown in the Figure 7 the tea samples clearly fall into three groups Cameroon, Luxemburg and china, except Tea 3 which is misclassified comparing the result we have obtained using elemental content as input instead of spectra data.

Artificial neural network (ANN)

There are many different kinds of ANN but two types are most commonly used in chemometric to perform regression tasks: multilayer Feed-Forward Neural Network (MLF-NN) and some

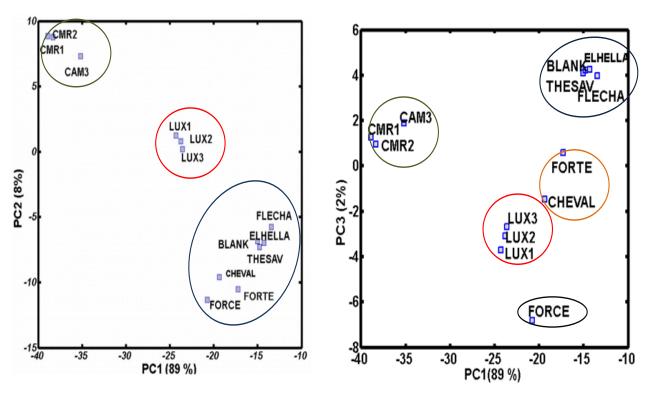


Figure 4. The score plot associated with PC1, PC2 and PC3 from principal components analysis of elemental composition of tea sample with total variability. Figures show a clear distinction between tea samples from different geographical area.

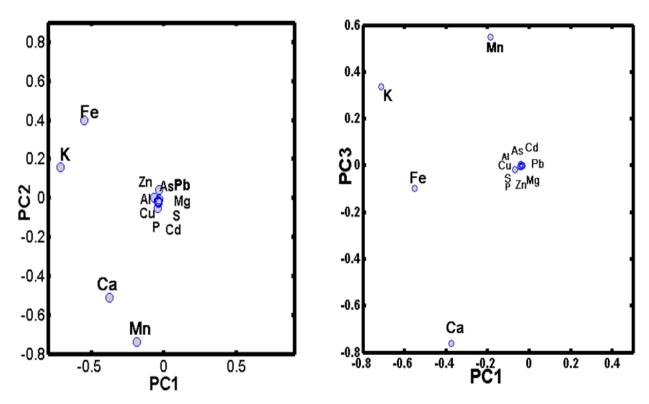


Figure 5. The loadings plot associated with PC1, PC2 and PC3 from principal components analysis of the elemental composition of tea sample.

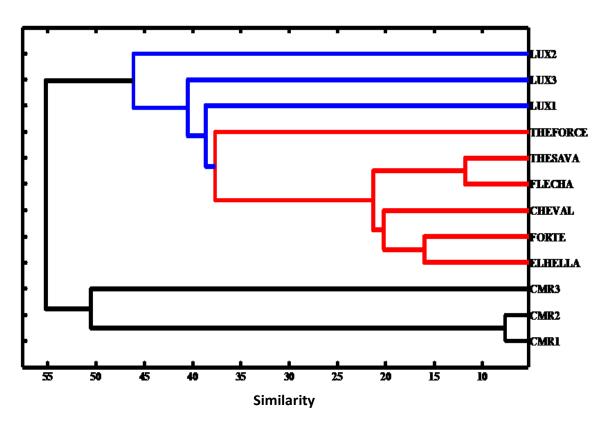


Figure 6. Dendrographic classification of tea sample from elemental contents according to their geographical origin.

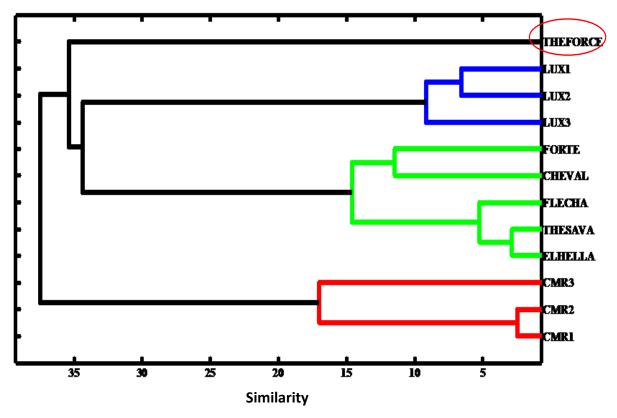


Figure 7. Dendrographic classification of tea sample from direct XRF spectra according to their geographical origin.

supervised modification of kohonen Self-Organizing Maps (SOMs). In this work a two-layer feed-forward network with sigmoid input hidden and output neurons were applied to predict the class of tea according to their geographical origin. The network was trained with scaled conjugate gradient back propagation. The data set was divided into three subset, training (70%), validation (15%) and testing (15%). Different neurons combinations were tried. The best result was obtained when neural network architecture with two layers, 8-4, was applied. The number of epochs was 40 with a learning rate and momentum of 0.25 and 0.5 respectively. The predictive MLF-NN model gives 100% correct classification of tea varieties with 0% error and 3.5×10⁻⁷ mean squared error.

Conclusion

The technique of energy dispersive X-ray fluorescence combined with multivariate statistic is a rapid method of elemental analysis of tea sample and classification based on their geographical origin, according to the calcium, magnesium, iron and potassium concentration. The differences in the concentration of the various elements within the different tea is attributed to the preferential absorbability of a particular plant for the corresponding element and the mineral composition of the soil in which the plant grows as well as its surrounding climatological conditions. The use of elemental concentration as inputs has shown that each cultivar presents distinctive element content. The classification of thirteen tea samples were 100% accurate in total by single value decomposition (SVD), hierarchical cluster analysis (HCA) and Artificial Neuron Networks analysis (ANNs). The present study through the establishment baseline of the relationship between elemental concentration and geographical origin will allow us in the second phase of our project to be able to assess the best quality of tea by taking into account both parameters (elemental concentration geographical origin).

Further work could involve quality assessment and safety aspects of teas and others local foods and beverages. This present work gives possibility for customs service applications to control the origin of tea.

ACKNOWLEDGMENTS

The authors acknowledge Dr. Mouhamadou Diallo Malaria Section, Molecular Biology Unit, Bacteriology and Virology Laboratory at Aristide Le Dantec Hospital in providing technical support and gratefully acknowledge Dr. Lutgen Pierre Luxembourg IFBV NGO in providing tea sample. The authors would like also to gratefully acknowledge the International Atomic Energy Agency IAEA for their kind help. We would like to thank Pr. Piet Van Espen (University of Antwerpen Belgium) and Mikkel Brydegaard (Lund University).

REFERENCES

- Budinova G, Vlacil D, Mestek O, Volka K (1998). Application of infrared spectroscopy to the assessment of authenticity of tea. Talanta 47:255-260.
- Chen Q, Zhao J, Fang CH, Wang D (2007). Feasability study on identification green, black and oolong teas using near-infrared spectroscopy based on support vector machine (SVM) Spectrochim. Acta A, 66:568-574.
- Dufresne CJ, Farnworth ER (2001). A Review of latest research findings on the health promortion properties tea. J. Nutr. Biochem.12:404-421.
- Dutta R, Hines EL, Gardner JW, Kashwan KR, Bhuyan M (2003). Tea quality prediction using a tin oxide based electronic nose: an artificial intelligence approach Sens. Actuators, B. 94:228-237.
- Fernandez-Caceres PL, Martin MJ, Pablos F, Gonzalez AG (2001). Study of catechin and xanthine tea profiles as geographical tracers. J. Agric. Food Chem. 49:4775-4779.
- Ferreira JFS, Janick J (1995). Production and detection of artemisinin from Artemisia annua. Acta Hortic. 390:41-49.
- Fujita H, Yamagami T (2008). Antihypercholesterolemic effect of Chinese black tea extract in human subjects with borderline hypercholesterolemia. Nutr. Res. 28:450-456.
- Fung KF, Zhang ZQ, Wong JWC, Wong MH (2003). Aluminium and fluoride concentrations of three tea varieties growing at lantau Island, Hong Kong Environ. Geochem. Health 25:219-232.
- Herrador MA, Gonzalez AG (2001). Pattern recognition procedures for differentiation of green, black and oolong teas according to their metal content from inductively coupled plasma atomic emission spectrometry. Talanta 53:1249-1257.
- Ivarsson P, Holmin S, Hojer NE, Krantz-Rulcker C, Winquist F (2001). Discrimination of tea by means of a voltammetric electronic tongue and different applied waveforms Sens. Actuators B 76:449-454.
- Karimi G, Hasanzadeh MK, Nili A, Khashayarmanesh Z, Samiei Z, Nazari F, Teimuri M (2008). Concentrations and health risk of heavy metals in tea samples marketed in Iran. Pharmacol. Online 3:164-174
- Luczaj W, Skrzydlewska E (2005). Antioxidative Properties of Black Tea. Prev. Med. 40:910-918.
- Marcos A, Fisher A, Rea G, Hill SJ (1998). Preliminary study using trace element contrations and chemometrics approach to determine the geographical origin of tea. J. Anal. Atom. Spectrom. 13:521-525
- Mark H (2007). The beneficial effects of tea on immune function and inflammation: A Review of evidence from in vitro, animal and human research. Nutr. Res. 27:373-379.
- Markowick AA, Grieken V (1993). Handkbook of X-ray Spectrometry, Marcel Dekker, Inc., New York.
- Matsumoto H, Hirasawa E, Morimura S, Takahashi E (1976). Localization of aluminum in tea leaves. Plant Cell Physiol. 17:627-631.
- Mei L, Lundin P, Brydegaard M, Gong S, Tang D, Somesfalean G, He S, Svanberg S (2012). Tea classification and quality assessment using laser-induced fluorescence and chemometric evaluation. Opt. Soc. Am. 7:51.
- Record IR, Dreosti IE (1998). Protection by black tea and green tea againts UVB and UVA+B induced skin cancer in hairless mice. Mutat. Res. 422:191-199.
- Salvador MJ, Lopes GN, Filho VFN, Zucchi OLAD (2002). Quality control of commercial tea by X-ray fluorescence. X-Ray Spectrometry 31:141-144.
- Siddiqui IA, Raisuddin S, Shukla Y (2005). Protective effects of black tea extract on testosterone induced oxidative damage in prostate. Cancer Lett. 227:125-132.
- Stragg GV, Millin DJ (1975). The nutritional and therapeutic value of tea: A Review. J. Sci. Food Agric. 26:1439-1442.
- Tertian R, Claisse F (1982). Principles of quantitative X-ray fluorescence Analysis, Heyden & Son Ltd, London, UK.
- Varmuza K, Filmoser P (2009). Introduction to Multivariate Statistical Analysis in Chemometric, CRC Press, Boca Raton, FL, USA.
- Yingxu C, Mingge Y, Jie X, Xincai C, Jiyan S (2009). Differentiation of eight tea (camellia sinensis) cultivars in china by elemental fingerprint of their leaves. J. Sci. Food Agric. 89:2350-2355.
- Weisburger JH (1997). Tea and Health: A Historical Perspective. Cancer Lett. 114:315-317.