Authenticating Geographical Origin of Tea from the North-East Region of India Using ICP-MS and Agilent Mass Profiler Professional Chemometrics Software

Vinay Jain[1], Partha Sen[1], Prasenjit Kar[1], Shuofei Dong[2], and Ed McCurdy[3]*

[1] Agilent Technologies (international) Pvt., Ltd., India
[2] Agilent Technologies Co., Ltd. (China), China
[3] Agilent Technologies LDA (UK) Ltd., UK

* Corresponding Author:
ed.mccurdy@agilent.com

Agilent
Trusted Answers

Introduction

India is the world's second largest producer of tea (after China), with most Indian tea being produced in the north-eastern states of Assam, West Bengal, and Tripura. Tea producers blend different varieties to create consistent products with a well-balanced flavor. Blending helps to take the pressure off high-demand tea varieties such as Darjeeling that is only grown in West Bengal. However, if a consumer buys a packet of tea labelled as Darjeeling, it is difficult for them to be certain of the quality – and origin – of the tea, based solely on the labeling.

Foodstuffs with a high market value – including high-value tea varieties – are increasingly a target for food fraudsters, who routinely substitute, adulterate, or mislabel food items for financial gain. Tea varieties with a value that depends on their region of origin are more vulnerable to adulteration, especially where demand for the product is higher than the amount produced. To protect consumers and producers from food fraud, the analytical methods employed need to be able to distinguish the geographical origin of tea and verify the product's authenticity.

Recent studies have shown that the multielement technique of ICP-MS can provide an "elemental fingerprint" for each sample that is indicative of the sample's geographic origin.[1] In this study, elemental data obtained using an Agilent 7850 ICP-MS was processed using Agilent Mass Profiler Professional (MPP) chemometrics software. Principal Component Analysis (PCA) enabled the geographical origin of tea to be differentiated based on the concentrations of 18 of the 68 elements measured. Two MPP approaches (SVM and LDA), that are often applied to food-authentication studies, were used to build class prediction models based on the 18 elements. Both prediction models were able to correctly identify the geographical origin of 24 'unknown' tea samples, including high-value Darjeeling.

The ICP-MS elemental fingerprint method has the potential to characterize the geographical origin of a wide range of foodstuffs.[2] The results can be used to authenticate the origin of genuine products or identify products that have been adulterated or mislabeled.

Experimental

Instrumentation: An Agilent 7850 ICP-MS was used for the analysis of digested tea samples. The standard instrument configuration was used, with MicroMist glass concentric nebulizer, Peltier-cooled quartz double-pass spray chamber, quartz torch with 2.5 mm injector, nickel interface cones, and x-Lens. For speed and simplicity of method setup, helium (He) KED cell mode was used for all elements. Operating conditions are shown in Table 1.

Table 1. Agilent 7850 ICP-MS operating conditions

Parameter	Value
RF Power (W)	1550
Sampling Depth (mm)	9.0
Nebulizer Gas Flow (L/min)	0.92
Spray Chamber Temperature (°C)	2.0
Lens Tune	Autotune
He Flow Rate (mL/min)	4.3
KED (V)	3

Agilent ICP-MS instruments have certain advanced features, which simplify the generation of the high-quality, multi-element data sets needed for food authenticity studies:

1. **Robust, low CeO/Ce plasma:** Provides market-leading matrix tolerance, reduces the formation of common polyatomic ions, minimizes sample dilution, and saves time by reducing the need for screening.

2. **Ultra High Matrix Introduction (UHMI) system for direct analysis of 25% total dissolved solids (TDS):** Practically eliminates matrix suppression, so results are more consistent for variable, high matrix samples, and labs don't need to waste time preparing matrix-matched calibrations. Provides better matrix decomposition, so less signal drift, fewer QC failures, and less frequent maintenance, another time saving.

3. **Octopole Reaction System (ORS[4]) cell with Shield Torch System for ion energy control:** This unique combination provides the most effective control of common polyatomic ion interferences using helium collision mode. This saves time by simplifying method setup and provides accurate and precise data, without the need for reaction gases in typical applications.

4. **Automated dual-mode detector with 10 orders linear dynamic range:** Saves time by reducing over-range results and simplifies calibration by enabling the use of consistent major element calibration ranges. Users don't need to setup custom signal attenuation to bring each major element within the detector range.

5. **Chemometrics software compatibility:** Agilent ICP-MS data is compatible with Agilent MPP, a chemometrics software package that is designed to process large sets of mass spectrometry (MS) data. MPP includes a choice of data classification methods, enabling analysts to build robust models to predict attributes such as sample origin of unknown samples.

During method development, a total of 68 elements were measured, of which 18 elements showed enough variation to be used in the class-prediction modeling.

Calibration standards: Calibration standards for the 18 "indicator" elements were prepared in 5% HNO_3 by volume (v/v). Calibration standards were prepared from Agilent standard solutions including multi-element calibration standard-1 (p/n 8500-6944), multi-element calibration standard-2A (p/n 8500-6940), environmental spike mix (p/n 5183-4687), and multi-element calibration standard-4 (p/n 8500-6942).

Twelve of the target elements (B, V, Cr, Co, Ni, Zn, Se, Rb, Sr, Mo, Cs, and Ba) were calibrated from 0.1 to 200 ppb; La, Ce, and Nd were calibrated from 0.1 to 20.0 ppb; Mn and Al from 0.1 to 2.0 ppm; Mg from 1.0 to 20 ppm. The rhodium (Rh) internal standard (ISTD) solution was prepared from an Agilent single element Rh standard (p/n 5190-8509). The ISTD solution (1.00 mg/L) was mixed with the sample online using the standard T-connector.

Sample preparation Samples of 150 teas from eight tea-producing regions of India (Dooars, Cachar, Terai, Darjeeling, Tocklai, Tripura, Upper Assam, and North Bank) were provided by the Tea Research Association, India. Each sample was dried and ground, and approx. 0.50 ± 0.01 g was accurately weighed into a polytetrafluoroethylene (PTFE) microwave digestion vessel.

The samples were predigested at room temperature in 5 mL HNO_3 and 0.1 mL HCl before microwave digestion (Anton Paar, Austria) shown in Table 2. Once cooled to ambient temperature, the solutions were transferred to 50 mL centrifuge tubes and made up to 40 mL using de-ionized water. All samples, method blanks, and spiked samples were prepared using the same procedure (Figure 1).

Table 2. Microwave digestion program

Steps	Ramp Time (minutes)	Temperature, °C	Hold Time (minutes)
I	15	105	3
II	10	160	3
III	10	190	10
IV	-	Ambient Temperature	-

Spiked Sample preparation To test the accuracy of the quantitative method, a spike recovery study was carried out by spiking a randomly selected sample from the North Bank region with known concentrations of the target elements.

Of the elements analyzed, tea samples generally contain high levels of Zn, Al, Mg, and Mn, moderate concentrations of B, V, Cr, Co, Ni, Rb, Sr, Mo, and Se, and lower levels of Se, Cs, La, Ce, and Nd. Therefore, before digestion, the tea sample from the North Bank region was spiked with the three groups of elements at 1 ppm, 100 ppb, and 10 ppb, respectively, relative to the final sample digest solution. The selected tea sample was spiked and prepared in triplicate.
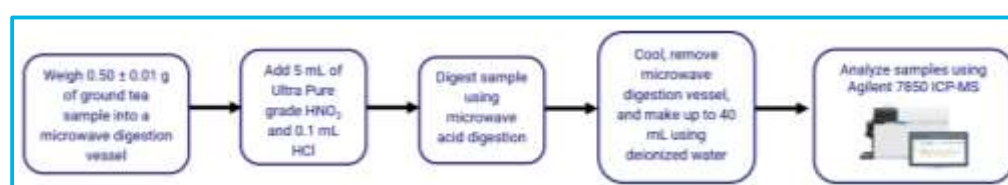

Figure 1. Analytical workflow used for microwave digestion sample preparation and ICP-MS analysis of tea samples.

Results and Discussion

Detection limits Instrument detection limits (IDLs) for the 18 elements were calculated as 3 x standard deviation (SD) of the concentration of the calibration curve blanks (n=10). Dilution corrected "method" DLs (MDLs) were calculated by applying the dilution factor of 80 (0.50 g of tea digested and diluted to a final volume of 40 mL). The IDLs and MDLs are shown in Table 3. The low DLs demonstrate the suitability of the 7850 ICP-MS for the determination of a wade range of elements in tea.

Table 3. Calibration coefficient (r), IDLs, and MDLs of the 18 elements used to distinguish the geographical origin of tea.

Element	R Value	IDL, µg/L	MDL, µg/L
11 B	0.9999	0.069	5.48
24 Mg	0.9999	0.003	0.24
27 Al	0.9997	0.007	0.55
51 V	0.9999	0.0003	0.02
52 Cr	0.9999	0.0002	0.01
55 Mn	0.9999	0.017	1.35
59 Co	0.9999	0.003	0.24
60 Ni	0.9997	0.0004	0.03
66 Zn	0.9999	0.001	0.10
82 Se	0.9999	0.004	0.32
85 Rb	1.0000	0.0003	0.02
88 Sr	1.0000	0.0002	0.01
95 Mo	1.0000	0.0003	0.02
133 Cs	0.9999	0.0001	0.01
137 Ba	1.0000	0.0005	0.04
139 La	0.9997	0.00005	0.004
140 Ce	0.9997	0.002	0.13
146 Nd	1.0000	0.001	0.08

Internal standard stability Rh was used as the ISTD as it was not expected to be present in the tea samples, unlike other potential ISTD elements such as Li and Sc. The analytical sequence comprising 150 tea sample digests, method blanks, quality control (QC) checks, and spike recovery solutions was analyzed repeatedly over 11 hours (280 runs in total).

As shown in Figure 2, all the ISTD recovery measurements were within the ± 20% QC limits (indicated by the red dotted lines). The recovery test shows that the very robust plasma of the 7850 decomposed the variable sample matrices effectively, enabling excellent stability to be maintained over the long run. Also, no drift occurred during the sequence, confirming little or no matrix deposition on the interface.
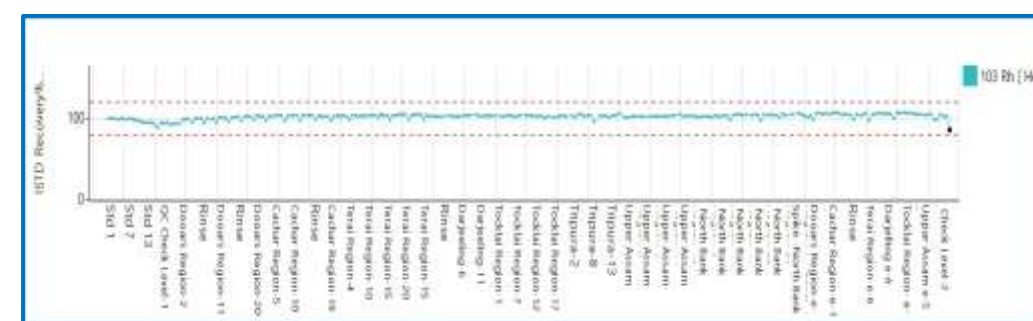

Figure 2. Recovery of ISTD measured over 11 hours (280 analyses) using the Agilent 7850 ICP-MS. Due to limited space, not all the sample names are shown

Quality control check: As a QC check, a standard containing 1.0 µg/L of La, Ce, Nd, and 10.0 µg/L of the other elements was measured before and after the batch analysis. The drift (% difference) between the standards run at the start and end was less than 5.0% for all elements, again demonstrating the stability and robustness of the Agilent 7850 ICP-MS (Table 4)

Table 4. QC check solution measured before and after the 11-hour run using the Agilent 7850 ICP-MS

Element	QC Conc., µg/L	Measured Concentration, µg/L Run 1	Run 2	Drift (% change)
11 B	10.0	9.94	9.70	2.5
24 Mg	10.0	10.37	9.87	5.0
27 Al	10.0	10.70	10.46	2.4
51 V	10.0	9.93	9.61	3.2
52 Cr	10.0	9.87	9.55	3.1
55 Mn	10.0	9.82	9.44	3.8
59 Co	10.0	9.88	9.66	2.2
60 Ni	10.0	9.59	9.33	2.6
66 Zn	10.0	10.07	9.91	1.6
82 Se	10.0	10.53	10.04	4.9
85 Rb	10.0	9.69	9.36	3.2
88 Sr	10.0	9.67	9.35	3.2
95 Mo	10.0	9.60	9.67	0.7
133 Cs	10.0	9.58	9.41	1.8
137 Ba	10.0	9.97	9.93	0.4
139 La	1.0	0.95	0.95	0.1
140 Ce	1.0	0.95	0.95	0.0
146 Nd	1.0	1.00	1.00	0.0

Quantitative results for tea samples Digested tea samples from eight different regions of India were analyzed. As shown in Table 5, the % variation of the quantitative results for samples from within each region was <5%.

Table 5. Mean concentrations and variation of 18 elements in tea samples from eight regions, corrected for dilution, mg/kg.

Element	Dooars Region Mean Conc. (n=20)	Variation (%)	Cachar Region Mean Conc. (n=20)	Variation (%)	Terai Region Mean Conc. (n=20)	Variation (%)	Darjeeling Region Mean Conc. (n=15)	Variation (%)	Tocklai Region Mean Conc. (n=15)	Variation (%)	Tripura Mean Conc. (n=20)	Variation (%)	Upper Assam Mean Conc. (n=20)	Variation (%)	North Bank Mean Conc. (n=20)	Variation (%)
11 B	-12	2.0	15	1.8	9.7	2.1	12	2.1	6.4	2.3	14	3.2	9.9	2.4	11	2.7
24 Mg	2245	0.6	1989	0.9	2183	1.3	1951	0.8	2337	1.5	1418	2.9	2162	1.5	1988	2.2
27 Al	853	0.7	734	0.8	922	1.3	559	1.2	579	1.5	349	3.0	630	1.5	759	2.0
51 V	0.9	4.9	0.22	2.8	0.25	3.6	0.25	2.7	0.23	2.4	<MDL	-	0.10	3.6	8.3	3.0
52 Cr	7.9	4.6	1.9	1.0	2.1	3.6	2.4	7.9	1.6	0.32	4.9	0.44	3.7	2.8	2.0
55 Mn	639	0.8	637	1.0	504	1.2	384	1.0	817	1.3	540	3.0	308	1.3	1088	2.0
59 Co	0.18	4.4	0.34	1.6	0.39	2.8	0.14	4.2	0.30	1.3	2.9	0.12	2.6	1.2	2.8
60 Ni	5.9	4.3	7.7	1.1	7.6	2.7	8.2	0.9	8.0	1.7	3.4	2.8	1.5	8.9	3.8
66 Zn	32	4.4	20	0.9	33	4.9	41	0.9	37	1.5	27	2.9	19	1.6	28	2.3
82 Se	0.85	3.9	0.27	4.9	0.19	4.8	0.19	4	<MDL	-	0.20	4.1	0.31	4.6	0.19	3.1
85 Rb	23	0.5	44	1.0	67	1.6	71	1.4	29	1.9	16	1.3	15	93	2.2
88 Sr	3.1	4.3	1.2	1.0	1.2	1.0	16	13	3	2.9	1.6	1.1	2.3	4.9
95 Mo	0.90	2.9	0.25	3.1	0.23	3.6	<MDL	-	<MDL	-	<MDL	-	1.0	4.8
133 Cs	0.36	2.9	0.16	1.6	0.17	1.9	0.60	<MDL	-	0.12	2.9	0.13	2.1	0.13	2.0
137 Ba	19	4.9	3.1	1.0	3.8	2.0	7.2	2.4	3	2.9	37	1.4	41	2.2
139 La	0.13	1.7	0.40	1.1	0.05	0.16	4.2	0.14	5.1	0.29	0.15	1.9	0.21	4.3
140 Ce	0.21	1.1	0.33	1.0	0.28	1.8	0.28	1.5	0.24	5.1	0.24	1.4	0.38	2.2
146 Nd	0.10	2.4	0.20	2.2	0.12	4.8	0.12	3.8	<MDL	-	0.21	1.7	0.24	3.0

Spike recoveries and evaluation of matrix effects To test the accuracy of the quantitative method, a spike recovery study was carried out on one of the tea samples from the North Bank region. Three replicate samples were prepared, spiked and analyzed and the spike recoveries were calculated, as shown in Table 6. The average of the spike recoveries (n=3) was within 100 ± 10% with %RSD of the measurements less than 5%. The excellent recoveries confirm minimal matrix effects, due to the robust (low CeO/Ce) plasma of the 7850.

Table 6. Accuracy and precision of analysis of a North Bank spiked tea sample prepared in triplicate

Element	Conc. Unspiked Sample, µg/g	Spike Conc., µg/g	Conc. Spiked Sample, µg/g	Recovery, % (n=3)	RSD, %
11 B	11.4	8	18.7	92	3.7
24 Mg	1987	80	2071	105	4.8
27 Al	763	80	837	93	4.0
51 V	8.3	8	16.5	103	0.7
52 Cr	12.0	8	20.6	107	1.1
55 Mn	1087	80	1160	92	2.1
59 Co	1.4	8	9.5	101	2.0
60 Ni	8.7	8	16.8	102	1.0
66 Zn	33	80	118	107	1.4
82 Se	0.18	0.8	0.99	101	1.4
85 Rb	94	8	101	97	1.7
88 Sr	11.5	8	18.9	93	3.3
95 Mo	0.98	8	8.7	96	0.5
133 Cs	0.14	0.8	0.87	92	0.8
137 Ba	42	8	49	95	2.2
139 La	0.23	0.8	0.99	95	3.6
140 Ce	0.39	0.8	1.21	103	4.0
146 Nd	0.25	0.8	1.1	101	3.5

Data analysis using MPP software The ICP-MS data for 18 elements measured in 150 tea samples were combined and imported into Agilent MPP chemometric software for statistical analysis. The MPP software includes various tests such as t-test, analysis of variance (ANOVA), model building algorithms, BoxWhisker, correlation, and clustering analysis.

Principal Component Analysis A PCA (with a P cut-off value <0.5) was performed to see if there were significant differences between tea sample groups from the different geographical regions. The PCA technique evaluates the relative contribution of elements to the separation of the groups. The three dimensional (3D)-PCA scores in Figure 3 show significant variation between tea samples from different geographical origin based on elemental profiling of 18 elements. A total of 72% of the variance ratio was captured in first three dimensions PC1 (X-axis, 38.68%), PC2 (Y-axis, 19.87%), and PC3 (Z-axis, 13.66%). According to the PCA loading values, the separation between the geographical origin of tea was driven by Sr, Ba, and B for PC1; Cs, La, and Rb for PC2; and Mo, Ce, and Nd for PC3.
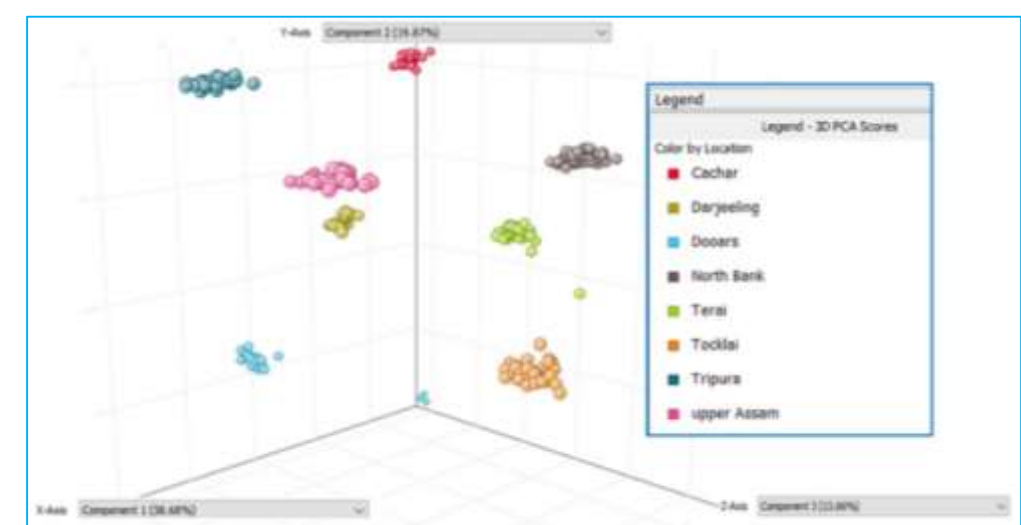

Figure 3. 3D-PCA based on 18 elements in 150 tea samples showing discrimination in geographical origin. The plot axes represent the three major PCA principal components.

Class prediction analysis: Class prediction analysis is a useful technique that enables unknown samples to be assigned to a previously determined group in an unbiased fashion. The MPP software includes several class prediction algorithms including soft independent modeling of class analogy (SIMCA), random forest, linear discriminant analysis (LDA), partial least squares discriminant analysis (PLSDA), decision tree, support vector machine (SVM), naive Bayes, and neural network. A classification model can be built using any of these class prediction algorithms.

Building a prediction model: LDA and SVM, which are commonly used class prediction algorithms for foods, were used to build class prediction models for identifying the geographical origin of tea samples. All the interpretations were carried out using non-averaged location values. To build the models, tea samples were selected at random from each of the eight regions (12 samples from Darjeeling and Tripura, and 18 from all other regions). The remaining 24 samples (three from each region) were used as unknowns to test the accuracy of both prediction models.

Testing the prediction model: The predicted location results obtained using the LDA and SVM models are shown in Table 7. The locations of all 24 'unknown' tea samples were correctly identified using both prediction models, with a high degree of confidence (values close to 1) in all cases.

Table 7. Predicted locations for 24 'unknown' tea samples.

Sample Name	Actual Location	SVM Prediction Model Predicted Location	Confidence Measure	LDA Prediction Model Predicted Location	Confidence Measure
Unknown C1	Cachar	Cachar	0.907	Cachar	0.986
Unknown C2	Cachar	Cachar	0.910	Cachar	0.967
Unknown C3	Cachar	Cachar	0.907	Cachar	0.977
Unknown D1	Dooars	Dooars	0.912	Dooars	0.976
Unknown D2	Dooars	Dooars	0.916	Dooars	0.966
Unknown D3	Dooars	Dooars	0.905	Dooars	0.964
Unknown Da1	Darjeeling	Darjeeling	0.886	Darjeeling	0.981
Unknown Da2	Darjeeling	Darjeeling	0.891	Darjeeling	0.979
Unknown Da3	Darjeeling	Darjeeling	0.906	Darjeeling	0.980
Unknown NB1	North Bank	North Bank	0.919	North Bank	0.978
Unknown NB2	North Bank	North Bank	0.921	North Bank	0.989
Unknown NB3	North Bank	North Bank	0.919	North Bank	0.989
Unknown T1	Tripura	Tripura	0.898	Tripura	0.985
Unknown T2	Tripura	Tripura	0.886	Tripura	0.947
Unknown T3	Tripura	Tripura	0.877	Tripura	0.940
Unknown Te1	Terai	Terai	0.902	Terai	0.986
Unknown Te2	Terai	Terai	0.905	Terai	0.983
Unknown Te3	Terai	Terai	0.911	Terai	0.983
Unknown To1	Tocklai	Tocklai	0.884	Tocklai	0.987
Unknown To2	Tocklai	Tocklai	0.887	Tocklai	0.981
Unknown To3	Tocklai	Tocklai	0.895	Tocklai	0.993
Unknown UA1	Upper Assam	Upper Assam	0.901	Upper Assam	0.985
Unknown UA2	Upper Assam	Upper Assam	0.913	Upper Assam	0.988
Unknown UA3	Upper Assam	Upper Assam	0.910	Upper Assam	0.985

Conclusions

- Sixty-eight major and trace elements were measured in 150 tea samples from eight regions of India using the Agilent 7850 ICP-MS.

- The wide dynamic range and helium collision cell mode of the 7850 enabled all elements to be measured in a single analysis using the same operating conditions.

- The multielement data batch was processed using Agilent MPP chemometrics software.

- Principal Component Analysis (PCA) was applied to the data, indicating that the concentrations of 18 elements in tea varied significantly depending on the region.

- Two class prediction models (SVM and LDA) that are often used in food-authentication studies were used to build class prediction models based on the 18 elements.

- Both models correctly identified the geographical origin of 24 'unknown' samples, including high-value Darjeeling.

The ICP-MS elemental fingerprint method can be used to confirm the geographical origin of a wide range of foodstuffs and to identify when a product has been adulterated or mislabeled for the purpose of fraud.

References

1. Bappaditya Kanrar et.al., Food Chemistry Advance 1, 2022

2. Jenny Nelson et al., Authentic Spices: Method for Identifying the Country of Origin, Food Qual. Safety, 2019.