



# Combining Multi-Element Analysis with Statistical Modeling for Tracing the Origin of Green Coffee Beans from Amhara Region, Ethiopia

Minbale Endaye<sup>1</sup> · Minaleshewa Atlabachew<sup>1</sup>  · Bewketu Mehari<sup>2</sup> · Melkamu Alemayehu<sup>3</sup> · Daniel Ayalew Mengistu<sup>4</sup> · Bizuayehu Kerisew<sup>5</sup>

Received: 18 June 2019 / Accepted: 6 August 2019 / Published online: 15 August 2019  
© Springer Science+Business Media, LLC, part of Springer Nature 2019

## Abstract

Characterization of coffee *terroirs* is important to determine authenticity and provide confidence for consumers to select the right product. In this regard, Amhara Region, which is located at the northwestern part of Ethiopia, produces various local coffee types with distinct cup qualities. The coffees are, however, not yet registered with certification marks or trademarks for indications of their geographical origins. This study was aimed at developing analytical methodology useful to determine the geographical origin of green coffee beans produced in Amhara Region based on multi-element analysis combined with multivariate statistical techniques. For this, a total of 120 samples of green coffee beans, collected from four major cultivating zones (West Gojjam, East Gojjam, Awi, and Bahir Dar Especial Zones) were analyzed for K, Mg, Ca, Mn, Fe, Cu, Fe, Co, Ni, Zn, Si, Cr, Cd, and Pb using inductively coupled plasma–optical emission spectroscopy. The elemental analysis data were subjected to principal component analysis (PCA) and linear discriminant analysis (LDA). PCA was used to explore the natural groupings of samples and the discriminatory ability of elements. Accordingly, the elements K, Mg, Ca, and Na were found to be the main discriminators among samples. LDA provided a model to classify the coffee samples based on their production zones with an accuracy of 94.2% and prediction ability of 93.4%. Thus, the elemental composition of green coffee beans can be used as a chemical descriptor in the authentication of coffee produced in Amhara Region.

**Keywords** Amhara Region · Green coffee beans · Metals · Geographical origin · Linear discriminate analysis · Principal component analysis

## Introduction

Coffee is the second most important commodity next to petroleum, with annual worth of up to 14 billion USD

worldwide. Coffee is also one of the most widely consumed beverages throughout the world, where its consumption varies widely according to geographical location [1]. Its pleasant taste, aroma, stimulant effect, and

Minbale Endaye, Minaleshewa Atlabachew, Bewketu Mehari, Melkamu Alemayehu, Daniel Ayalew and Bizuayehu Kerisew contributed equally to this work.

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s12011-019-01866-5>) contains supplementary material, which is available to authorized users.

✉ Minaleshewa Atlabachew  
atminale2004@yahoo.com; minaleshawaa@bdu.edu.et

<sup>1</sup> Department of Chemistry, College of Science, Bahir Dar University, P.O. Box 79, Bahir Dar, Ethiopia

<sup>2</sup> Department of Chemistry, College of Natural and Computational Sciences, University of Gondar, Gondar, Ethiopia

<sup>3</sup> Department of Plant Sciences, College of Agriculture and Environmental Sciences, Bahir Dar University, P.O. Box 79, Bahir Dar, Ethiopia

<sup>4</sup> Department of Geography and Environmental Studies, Faculty of Social Sciences, and Geospatial Data and Technology Center, Bahir Dar University, P.O. Box 79, Bahir Dar, Ethiopia

<sup>5</sup> Department of Biology, College of Science, Bahir Dar University, P.O. Box 79, Bahir Dar, Ethiopia

health benefits are the main driving forces for its consumption [2]. Moreover, coffee is used as a raw material in various industries including food and beverage industries. Currently, more than 121 countries are involved in the production, export, and processing of coffee worldwide (<http://www.ecea.org.et/uk/>).

Ethiopia is the birthplace of Arabica coffee (*Coffea arabica* L.). It is the only species grown in the country [3, 4]. Coffee plays a vital role in the Ethiopian economy and is a major source of foreign exchange earnings. Moreover, the livelihood of about 25% of the country's population is directly or indirectly dependent on the activities associated with the production, processing, and transportation of the crop [5]. Although Arabica coffee is the only species grown in Ethiopia, different varieties with distinct and impressive flavor profiles are produced in different regions of the country. These varieties have diverse tastes depending on their geographical origins [6], where climate, altitude, and soil type contribute to the unique characteristics of the varieties. Besides, coffees cultivated in different parts of the country are genetically diverse [7, 8].

In Ethiopia, coffee is mainly produced by smallholder farmers distributed throughout the country. The quantity and quality of the coffee produced, however, varies significantly with its cultivation region. The regional coffee varieties also fetch different prices on the market depending on their production region and quality [7]. Owing to their quality reputations, some of the regional coffees cultivated in the southern and eastern parts of the country are currently marketed with trademarks registered in various countries across the world [7, 8].

Amhara Region, which is found in the northwestern part of Ethiopia, is endowed with suitable agro-ecology and fertile soil for the production of Arabica coffee (<http://newbusinessethiopia.com/index.php/resources/29-agriculture/648/648>). Consequently, coffee is being produced in different parts of the region, with the major production areas located in East Gojjam, West Gojjam, and Awi Zones. Several tons of coffee is being harvested every year from these parts of the region where the majority of the produce is used for local consumption. As perceived by the local people, coffees produced in different areas of the region have diverse flavors and stimulation effect. However, coffees produced in this region have not yet been characterized and included in the country's coffee export classification list while the country is striving to boost its market share by providing high quantity and quality of coffee to the international market, and the coffee varieties cultivated in Amhara Region remain underutilized.

On the other hand, to be competent in the international market, protocols should be developed that help to trace

the sources of green coffee beans at different levels including continental, national, and regional levels. Traditionally, paper-based traceability and cup-quality verification by a panel of trained cup quality testers are used to characterize coffees and their growing areas. However, papers can be easily counterfeited and may create difficulties to differentiate coffees of different origins [8]. Therefore, development of analytical methods that are capable of verifying regional sources of green coffee beans is crucial to protect the market share and reputation of the growing countries including Ethiopia.

In this regard, the fatty acid composition of green coffee beans was used, for instance, to characterize Colombian coffees of different varieties that were grown in different regions [9]. Similarly, Villarreal et al. [10] have shown the effectiveness of fatty acid, elemental and chlorogenic acid compositions to distinguish Colombian coffees according to growing regions. Mehari et al. [7, 8] have also developed chemometric models, based on phenolic and elemental contents, to characterize the geographical origins of commercially important coffee varieties grown in the southern, eastern, and western parts of Ethiopia. According to the results of those authors, chemical constituents of green coffee beans vary significantly with their growing regions and found to be suitable to mark the coffee varieties with respect to their geographical origins.

In addition to chlorogenic acids, alkaloids and major elements, determination of trace elements is important to characterize the geographical origin of agricultural commodities as well as to evaluate their dietary importance and possible harmful effects [11]. In this regard, determination of trace metals like iron, copper, zinc, cobalt, and manganese is quite necessary, as they play important roles in biological systems. According to Liang et al. [12], the essential metals can also have harmful effects when their intakes significantly exceed the required quantities. On the other hand, determination of chromium, lead, and cadmium in food products is crucial as they can be toxic even when present in trace amounts. It is, therefore, necessary to quantify the contribution of coffee to the daily intake of minerals, especially for those individuals who regularly consume coffee beverage.

The chemical constitutes in general and the mineral constitutes in particular of coffee beans produced in Amhara Region have not been investigated and compared with other local as well as overseas coffees. The aim of this study was therefore to evaluate the effectiveness of multi-element and multivariate data analyses for differentiation of green Arabica coffee beans grown in Amhara Region. This will support the ongoing efforts toward certification of the region's coffees with respect to indications of geographical origin, as well as identification, protection, and enhancing their demands in the international market.

## Materials and Methods

### Description of the Study Area

The study was conducted on Arabica coffee grown by small-holder farmers in East Gojjam, West Gojjam, Awi, and Bahir Dar Especial Zones of Amhara Region, Ethiopia (Fig. 1). Coffees produced in these areas are mainly sold at the local markets where the incomes help the farmers to improve their livelihoods.

### Equipment and Chemicals

Inductively coupled plasma optical emission spectrometry (Perkin Elmer Optima 8000), multi-block Kjeldahl digester (DK-20; Velp Scientifica), electrical blender (FW100 High Speed Universal Disintegrator) and nylon sieve (200  $\mu\text{m}$  mesh size), perchloric acid (72%) and nitric acid (69%), and standard solutions of K, Ca, Na, Si, Fe, Mg, Mn, Zn, Ni, Cu, Pb, Cr, Co, and Cd were used in the study.

### Coffee Sample Collection and Preparation

A total of 120 sun-dried coffee cherry samples were directly collected from coffee-growing farmers distributed in the various districts of East Gojjam (48 samples), West Gojjam (39 samples), Bahir Dar Especial (12 samples), and Awi (21 samples) Zones. Each sample weighed 500 g. All the samples

were collected during the 2017/2018 crop season (December 2017 to February 2018) (Table 1). Sampling districts from each zone were selected purposively depending on their experiences in the production of coffee while the selection of individual farmlands within each district was random.

The husks of the coffee cherries were removed and the resulting green beans were powdered using an electrical blender. The coffee powders were then sieved using a 200- $\mu\text{m}$ -mesh-size sieve.

### Standard Solution Preparation and Calibration

Working standard solutions were prepared by serial dilution from stock solutions (1000  $\text{mg L}^{-1}$ ) of each element. The wavelength and calibration conditions corresponding to each element is provided in Table 2.

### Digestion

Digestion of the samples was undertaken by adding 0.5 g of powdered green coffee bean samples into Kjeldahl digestion apparatus (Velp Scientifica DK 20 (20-position with 250 mL tubes,  $\text{O}$  42 mm)) after proper optimization of the digestion procedure. For this, the use of 4 mL  $\text{HNO}_3$  and 4 mL  $\text{HClO}_4$  at variable temperatures (60  $^\circ\text{C}$  for 10 min, 80  $^\circ\text{C}$  for 10 min, 100  $^\circ\text{C}$  for 20 min, 120  $^\circ\text{C}$  for 20 min, and 160  $^\circ\text{C}$  for 30 min) was found to be optimum for coffee sample digestion (Table S1, Supplementary material). All the digested samples

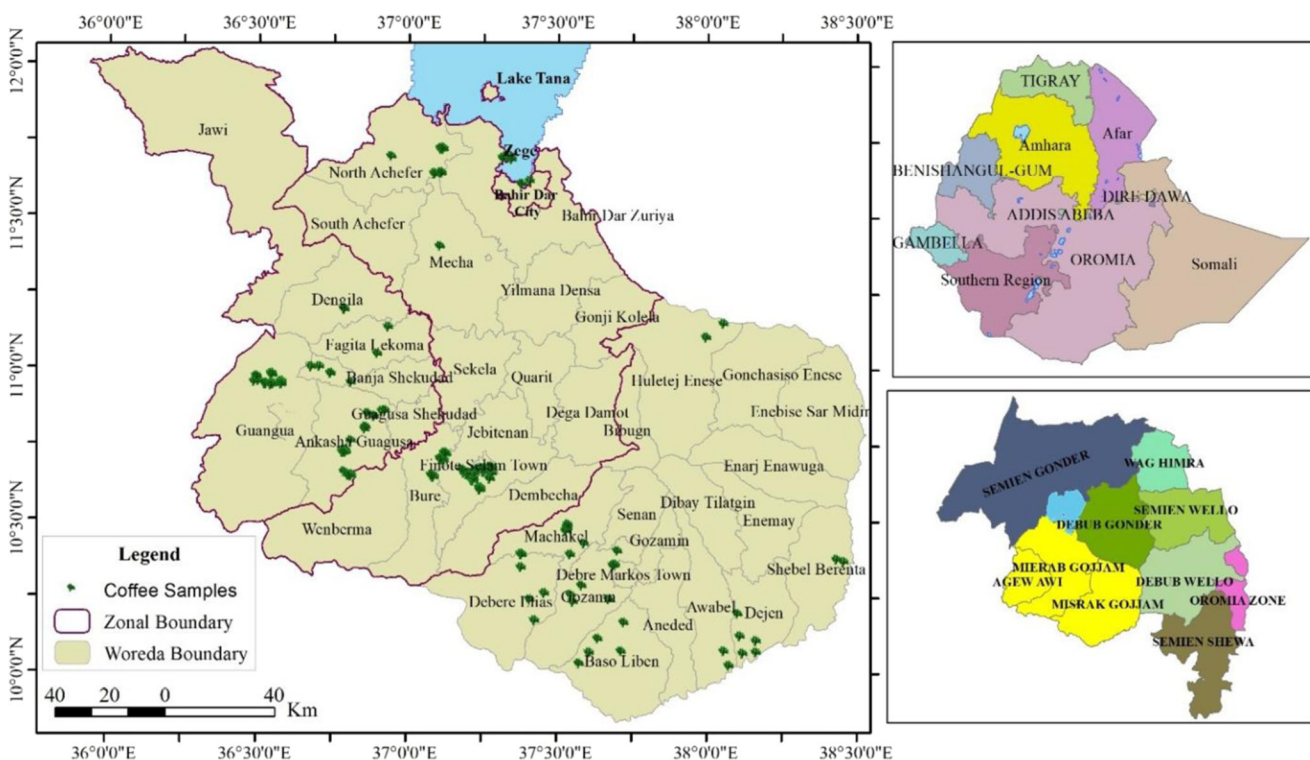


Fig. 1 Location map of sampling sites in Amhara Region

**Table 1** List of sampling districts and number of samples analyzed

Administrative zone	Districts	Number of samples
West Gojjam ( <i>n</i> = 39)	Finotselam	6
	Bure	3
	BureZuria	6
	JabiTehnan	6
	North Achefer	12
Awi ( <i>n</i> = 21)	Mecha	6
	AyhuGagusa	6
	Guangua	3
	Banja	3
	Chagni	3
East Gojjam ( <i>n</i> = 48)	Dangla	6
	Basoliben	15
	Gozamen	9
	Machakel	9
	Deber Elias	9
Bahir Dar Especial Zone ( <i>n</i> = 12)	Dejen	6
	Bahir Dar	6
	Bahir Dar Zuria	6
Total	18	120

were diluted to 50 mL with deionized water, filtered with filter paper, and analyzed by ICP-OES. The ICP-OES operational parameters were power, 1500 kW; plasma flow, 12 L min<sup>-1</sup>; auxiliary argon flow, 1 L min<sup>-1</sup>; and nebulizer flow, 0.8 L min<sup>-1</sup>.

### Accuracy of the Analytical Method

To know the accuracy of the analytical method, 0.5 g of powdered green coffee sample was fortified with a mixed standard

solution at a concentration corresponding to the average measured value for each element and subjected to the digestion and analysis procedure. The fortification was performed in triplicate and the percentage recovery of each element was calculated.

### Statistical Analysis

Chemometric data analysis was performed by using the statistical software packages SIMCA 13 (Umetrics, Sweden) and SPSS 20 (IBM Corp, USA). Each data set consisted of a matrix in which the rows represented the individual green coffee bean samples, and the columns represented the concentration of elements. A preliminary investigation based on principal component analysis (PCA) was applied to visualize sample trends and evaluate the discriminatory characteristics of the determined elements. One-way analysis of variance (ANOVA) was used to test for the presence of significant differences in the mean concentrations of the metals among the regional types of green coffee beans. Differences were considered significant when  $\alpha$  was less than 0.05. In addition, linear discriminant analysis (LDA) was applied to construct classification models useful for the prediction of the geographical production origin of the Arabica coffee produced in Amhara Region.

## Results and Discussion

### Analytical Characteristics of the Method

Detection limit of the method was determined by digesting and analyzing blank samples. Standard deviation of the triplicate blank samples was taken and multiplied by three in order

**Table 2** The wavelengths, calibration equations, and correlation coefficients corresponding to each of the elements determined

Element	Wavelength	Concentrations of standards (mg L <sup>-1</sup> )	Equation	R <sup>2</sup>
Cd	228.802	0.002, 0.02, 0.2, 1.00, 2.00, 4.00	$y = 249,622x + 4276.8$	0.9997
Ca	317.933	0.005, 0.05, 0.5, 1.00, 2.00, 4.00	$y = 253, 066x + 4276.8$	0.9998
Cr	267.716	0.002, 0.02, 0.2, 1.00, 2.00, 4.00	$y = 323, 861x + 5256.5.8$	0.9998
Co	228.616	0.002, 0.02, 0.2, 1.00, 2.00, 4.00	$y = 119, 176x + 1927$	0.9997
Cu	327.393	0.002, 0.02, 0.2, 1.00, 2.00, 4.00	$y = 35, 696x + 2003.4$	0.9998
Fe	238.204	0.002, 0.02, 0.2, 1.00, 2.00, 4.00	$y = 253, 553x + 7440.1$	0.9997
Pb	220.353	0.002, 0.02, 0.2, 1.00, 2.00, 4.00	$y = 16, 154x + 243$	0.9996
Mg	285.213	0.005, 0.05, 0.5, 1.00, 2.00, 4.00	$y = 1, 000,000x + 707.006$	0.9983
Mn	257.610	0.002, 0.02, 0.2, 1.00, 2.00, 4.00	$y = 3,000, 000x + 42, 242$	0.9996
Ni	231.604	0.002, 0.02, 0.2, 1.00, 2.00, 4.00	$y = 106, 031x + 2870.5$	0.9999
Zn	206.200	0.002, 0.02, 0.2, 1.00, 2.00, 4.00	$y = 4041, 1x + 447.74$	0.9995
Si	251.611	0.005, 0.05, 0.5, 1.00, 2.00, 4.00	$y = 4402, 6x + 250.63$	0.9978
Na	589.490	0.005, 0.05, 0.5, 1.00, 2.00, 4.00	$y = 24,611x + 866.11$	0.9994
K	766.490	1.00, 2.00, 4.00, 8.00, 16.00, 32.00	$y = 9274.3x + 5796.8$	0.9986



to obtain the method detection limit for each element. Results are presented in Table 3. Furthermore, accuracy of the method was evaluated by the recovery experiment and the following recovery values were obtained for each metal: K  $110 \pm 9\%$ , Ca  $110 \pm 7\%$ , Cu  $90 \pm 2\%$ , Fe  $105 \pm 7\%$ , Mg  $96 \pm 6\%$ , Mn  $105 \pm 3\%$ , Cr  $95 \pm 1\%$ , Si  $105 \pm 8\%$ , Ni  $95 \pm 2\%$ , Pb  $90 \pm 2\%$ , Cd  $90 \pm 5\%$ , Na  $106 \pm 7\%$ , Co  $95 \pm 3\%$ , and Zn  $110 \pm 8\%$ , which indicates that the method was accurate. The detection limit of the method varied from the lowest  $0.03 \text{ mg kg}^{-1}$  for Cd up to the highest  $100 \text{ mg kg}^{-1}$  for K.

MDL method detection limit

### Concentration of Elements in Coffee Samples

The mean concentration of the studied metals in the samples from each zone is given in Table 4. In all the samples, the major constituents were K, Mg, Ca, Na, Si, and Fe. These elements are considered as macroelements based on their concentrations in various samples [13, 14]. The mean concentrations of these metals in the studied samples were 17,094, 1702, 972, 200, 63.4, and  $52.6 \text{ mg kg}^{-1}$ , respectively, where they are ordered as  $\text{K} \gg \text{Mg} > \text{Ca} > \text{Na} > \text{Si} > \text{Fe}$ . These concentrations are within the concentration range that was also found in Arabica green coffee beans collected from various countries [7, 13, 15]. The results of the present study are generally in agreement with the results of various studies where K was the most abundant element followed by Mg and Ca [7, 9, 13, 14, 16, 17].

Based on the results of the present study, coffee green beans collected from West Gojjam Zone had higher mean concentration of K ( $18,127 \text{ mg kg}^{-1}$ ), Mg ( $1978 \text{ mg kg}^{-1}$ ), Ca ( $1130 \text{ mg kg}^{-1}$ ), Na ( $230 \text{ mg kg}^{-1}$ ), and Si ( $83.0 \text{ mg kg}^{-1}$ ) followed by those coffee samples collected from Bahir Dar Especial Zone ( $17,364 \text{ mg kg}^{-1}$  K,  $1745 \text{ mg kg}^{-1}$  Mg,  $943 \text{ mg kg}^{-1}$  Ca,  $170 \text{ mg kg}^{-1}$  Na, and  $71.2 \text{ mg kg}^{-1}$  Si (Table 4). The elements K and Mg were lowest in samples collected from East Gojjam Zone with mean concentration  $16,403$  and  $1425 \text{ mg kg}^{-1}$ , respectively, while the concentrations of Ca ( $870 \text{ mg kg}^{-1}$ ) and Na ( $170 \text{ mg kg}^{-1}$ ) were lowest in coffee samples collected from West Gojjam Zone.

One-way ANOVA ( $\alpha = 0.05$ ) revealed that K, Mg, Ca, Na, and Si were significantly different in all coffee samples except the samples between East Gojjam and Awi Zones for K, and the samples between East Gojjam and Bahir Dar Especial zone for Ca where the differences were insignificant. The concentrations of the microelements (Cu, Zn, Co, Ni, Mn, and Cr) are in good agreement with results which were reported from Ethiopian and coffee of others countries [7, 13, 14, 16–18]. However, the concentration of Mg recorded in the present study was almost 10 times higher than the data reported by Bertrand et al. [9] in Colombian coffee. This might be due to variations in the soil chemistry and other factors of the study areas.

Coffee is regarded as one possible source of micro and trace metals that are essential for the normal functioning of the human body [11, 13]. The presence of elements like Fe, Zn, Mn, Cu, Co, Ni, and Cr is widely reported in coffee. However, their concentrations were variable with the coffee type [13]. The order of these elements in the present studied coffees was found to be  $\text{Fe} (52.6 \text{ mg kg}^{-1}) > \text{Cu} (23.8 \text{ mg kg}^{-1}) > \text{Mn} (17.2 \text{ mg kg}^{-1}) > \text{Zn} (16 \text{ mg kg}^{-1}) > \text{Co} (5.2 \text{ mg kg}^{-1}) > \text{Ni} (2.79 \text{ mg kg}^{-1}) > \text{Cr} (1.33 \text{ mg kg}^{-1})$ .

Based on the results, coffee samples collected from Bahir Dar Especial Zone contained the highest concentration of microelements like Fe, Mn, and Co while the concentration of Fe in coffee of East Gojjam, Mn in coffee of Awi, and Co in the coffee of West Gojjam were lowest (Table 4). In this regard, the concentration of Fe was significantly varied between sampling zones of coffee. Similarly, the concentration of Co was also significantly different between sampling zones except between samples of West Gojjam Zone and Awi Zone and between East Gojjam and Bahir Dar Especial Zone. A wide variation in the concentration of Zn was also noticed between the sampling zones whereby the lowest mean concentration was found in coffee samples collected from East Gojjam Zone (average  $4 \text{ mg kg}^{-1}$ ) and the highest was recorded in samples collected from Awi Zone (average  $22 \text{ mg kg}^{-1}$ ). Such variations might be aroused due to variations in the physicochemical properties of the soils of the sampling areas and coffee genotypes.

The concentrations of Fe, Cu, and Mn which were found in the present study corroborate with the data reported by Mehari et al. [7] and Habte et al. [13] for other Ethiopian coffee types. However, the concentration of Zn found in the present study was about five times higher than its concentration reported in the other Ethiopian coffee types collected from Eastern, Southern, and Western parts of the country [13]. Compared to the results of Brazilian, Colombian, and Hawaiian coffee varieties, the results of the present study show unpredictable trends (Table 5). While the concentrations of Fe and Mn that were found in the coffee samples are similar with the literature values, the concentrations of Zn, Cu, and Co were widely varied between samples of different countries (Table 5).

The mean concentration of Cr across the coffee samples was from  $0.921$  to  $1.84 \text{ mg kg}^{-1}$  which was not significantly different ( $P > 0.05$ ) among the coffee beans. Nevertheless, the concentration of Cr found in other coffee varieties as indicated by Habte et al. (2016) and Krivan et al. (1993) was about five times less compared to the results of the present study. Lead was detected in all coffee samples of the present study with an overall average concentration of  $0.38 \text{ mg kg}^{-1}$ . On the other hand, cadmium was detected only in samples collected from West Gojjam and Awi Zones with an average concentration of  $0.012 \text{ mg kg}^{-1}$ . The determined concentrations of Pb and Cd were comparable with values reported for some coffee

**Table 3** Method detection limits of each element (mg kg<sup>-1</sup>)

Element	Cd	Ca	Cr	Co	Cu	Fe	Pb
MDL	0.03	0.5	0.06	0.09	0.2	0.2	0.18
Element	Mg	Mn	Ni	Zn	Si	Na	K
MDL	0.4	0.2	0.06	0.09	0.5	0.06	100

types in the literature (Table 5). Contrarily, about 100 times higher concentrations of Pb and Cd were also reported in Brazilian coffee [14].

The results of the present study generally indicated that the amount of metals found in coffee beans varies considerably with geographical production origin which can be associated with soil types, environmental conditions, the fertilizers used, and other conditions which is also supported with research findings of other researchers [17].

*SD* standard deviation, *ND* not detected, *Min* minimum, *Max* maximum

*NR* not reported, *ND* not detected

### Principal Component Analysis

The determined concentrations of metals were used with PCA to explore the presence of trends or patterns in the distribution of minerals among the coffee bean samples from the four

zones. PCA is an unsupervised statistical technique that categorizes samples based on projection of variance and hence avoids prejudice [20, 21]. PCA provides an immediate insight into the data structure and reveals any possible clustering of the observations and identification of any outliers in the data set. Furthermore, PCA provides an opportunity to identify the variables that contribute similar information to the PCA model and provide unique information concerning the samples under investigation [21].

The principal component results are elaborated in terms of scores and loadings. Scores and loadings are the transformed variable values and factors or multipliers that could result in the original scores, respectively [8, 21]. The scores plot (Fig. 2) shows the distribution of the studied coffee samples on the space created using the scores of the first two principal components.

A four-component model was developed with the maximum variation along PC<sub>1</sub> (72.2%). The cumulative chemical variation modeled using four PCs was 96% ( $R^2 X_{cum} = 0.96$ ).  $R^2$  is the proportion of the variation of the data set explained by the model. Though the majority of the variance was captured in the first principal component (PC<sub>1</sub>), accounting for 72.2% of the variation among samples, only 11.5% was explained by the second principal component (PC<sub>2</sub>), which reflects the existence of less variation of samples within the same province or zone. In other words, variation within zone

**Table 4** Concentrations (mg kg<sup>-1</sup> dry weight) determined in Arabica green coffee beans grown in various zones of Amhara Region, Ethiopia

Element	Element													
	K	Mg	Ca	Na	Si	Fe	Cu	Zn	Mn	Co	Ni	Cr	Pb	Cd
West Gojjam ( <i>n</i> = 39)														
Mean	18,127	1978	1130	234	83.0	59.1	27.4	17.6	14.7	3.05	1.50	1.49	0.087	0.012
SD	529	199	261	38.3	11.9	47.7	5.99	8.64	2.13	1.37	0.620	0.940	0.028	0.033
Min	17,155	1690	714	167	54.0	27.5	16.3	6.10	11.3	0.800	0.990	ND	ND	ND
Max	18,770	2299	1479	281	111	174	38.5	30.0	18.2	5.00	3.00	2.70	1.10	0.100
Awi ( <i>n</i> = 21)														
Mean	16,480	1661	870	181	49.4	48.2	29.8	22.2	12.8	3.59	1.56	1.84	0.360	0.080
SD	397	387	87.7	33.7	14.5	13.7	6.98	19.2	2.10	1.46	0.503	0.99	0.370	0.015
Min	15,976	1192	743	107	32.3	27.1	16.2	4.40	10.5	2.10	0.500	ND	ND	ND
Max	16,887	2207	952	216	76.7	70.2	38.0	65.6	15.9	6.00	2.500	2.90	2.20	0.500
East Gojjam ( <i>n</i> = 48)														
Mean	16,403	1425	945	215	49.9	32.5	25.2	4.02	16.7	6.21	5.31	1.06	0.196	ND
SD	493	167	173	54.9	5.99	22.1	10.6	3.99	9.38	5.87	5.16	0.624	0.918	
Min	15,356	1211	505	133	40.5	12.3	0.990	0.900	4.00	1.80	2.60	0.200	ND	
Max	17,396	1724	1190	313	60.0	96.6	36.4	21.6	42.0	19.0	15.4	4.00	4.65	
Bahir Dar Especial Zone ( <i>n</i> = 12)														
Mean	17,364	1745	943	170	71.2	70.6	12.6	20.2	24.4	7.89	9..25	0.921	0.380	ND
SD	552	147	41.8	29.3	12.7	30.4	9.38	15.9	14.5	6.57	3..04	0.383	0.147	
Min	16,712	1478	878	123	56.3	35.9	1.50	6.60	10.9	1.40	5.00	0.500	ND	
Max	18,089	1957	999	199	89.6	102	27.0	45.7	47.0	24.0	15.0	1.50	1.90	

was much less than between zone. The value of the cross-validation parameter,  $Q^2$  (cum), was 66.5%.  $Q^2$  is the proportion of the variation of the data set that can be predicted by the model. Accordingly, the model has good predictive ability as indicated by Atlabachew et al. [20].

Elements that best fit the model, with more than 72%, explained variability, were Ca, Mg, K, Na, Si, and Fe. From the plot (Fig. 2), it is evident that the coffee samples tended to cluster according to their geographical production origins. Coffee samples obtained from West Gojjam Zone showed differences in their mineral composition compared to coffees from the other zones. Accordingly, most of these coffee samples were separated from those collected from East Gojjam, Awi, and Bahir Dar Especial Zones by PC1.

The loading plot for the first two principal components is presented in Fig. 3. The plot displays how the individual elements correlate with each other and contribute to the model. Elements that are far away from the origin have a strong impact on the model, whereas those that are closer to the center have weaker influence [8]. From the plot, it is evident that K, Mg, Ca, and Na play the largest role in discriminating the green coffee beans from various zones of Amhara Region. The first component is highly influenced by K and Mg, whereas Ca and Na contributed to separation by the second component.

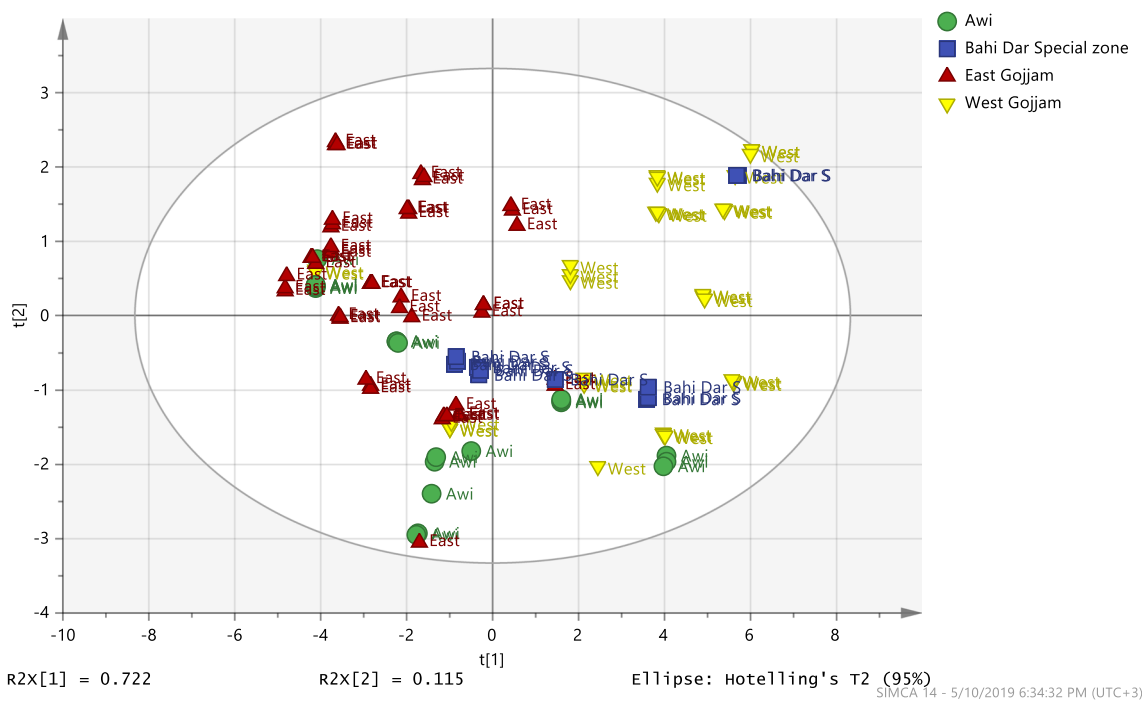
## Linear Discriminant Analysis

Based on the concentrations of the elements, an attempt was made to develop a chemometric method useful for the authentication of the production region of the coffee beans. For that purpose, LDA was applied. Initially, in order to assess the potentials of individual elements toward constructing a linear discriminant analysis model, a test of equality of group means was employed. The test measures each elements potential before the model is created. This was performed by applying one-way ANOVA for each element using the sample regions as the factor. According to the significance values in (Table S2, Supplementary material), each element, except Pb and Cr, was found to be significant ( $\alpha > 0.05$ ) for the discriminant model. Wilks'  $\lambda$  was another measure of a variable's potential. Smaller values indicate that the variable is better at discriminating between regions. The results (Table S2, Supplementary material) suggest that K was best in this regard. Consequently, all of the elements, excluding Pb and Cr, were used simultaneously to construct the LDA model. In addition, all of the four groups of samples were assigned equal prior probabilities.

Accordingly, three discriminant functions were computed. The first explained 71.6% of the variation, the second 18.2%, and the third 10.3%. The significance of the three functions in discriminating among regions was assessed from Wilks'  $\lambda$

**Table 5** Comparison of the elemental composition (mg kg<sup>-1</sup> dry weight) of Arabica green coffee beans reported in the literature with results from the present study

Element	Sample origin							
	This study (Northwestern Ethiopia)	Ethiopia (Eastern, Southern, and Western Ethiopia)	Ethiopia (Eastern, Southern, and Western Ethiopia)	Ethiopia (not specified)	Colombia	Brazil	Worldwide	Worldwide
K	16,403–18,127	16,200–17,300	13,599–16,272	13,010–17,000	13,100–15,900	15,000–21,400	113,816–18,800	17,500–19,600
Mg	1425–1978	1540–1650	1599–1838	1670–1690	144–196	1610–2090	1814–2015	2060–2350
Ca	870–1130	948–1060	712.9–985.9	710–1250	1030–1490	988–1710	938–1200	1010–1230
Na	170–234	NR	26.9–157.5	NR	NR	NR	9.0–22.0	NR
Si	49.4–83.0	16.4–22.1	NR	NR	NR	NR	NR	NR
Fe	32.5–70.6	28.7–45.3	21.1–126.7	26–44	33–59	27–108	39.0–60.0	13–31
Zn	4.02–22.2	NR	3.00–8.27	4–21	NR	NR	4.0–8.0	NR
Co	3.02–7.89	NR	0.019–0.006	2.6–8.4	NR	NR	0.06–0.2	NR
Cu	12.6–29.8	16.0–20.0	13.2–16.9	11–23	13–23	NR	12.0–20.0	12.5–18.1
Mn	12.8–24.4	24.1–38.3	12.6–24.0	13–19	13–31	13–31	23.0–47.0	19–39
Ni	1.50–9.25	NR	0.139–0.509	<0.04–2.5	NR	NR	NR	NR
Cr	0.921–1.84	NR	0.071–0.255				0.4–1.0	
Pb	0.087–0.380	NR	NR	<0.05	NR	NR	NR	NR
Cd	ND–0.012	NR	0.0009–0.004	ND	NR	NR	NR	NR
Ref.	This study	[7]	[13]	[19]	[9]	[16]	[18]	[17]

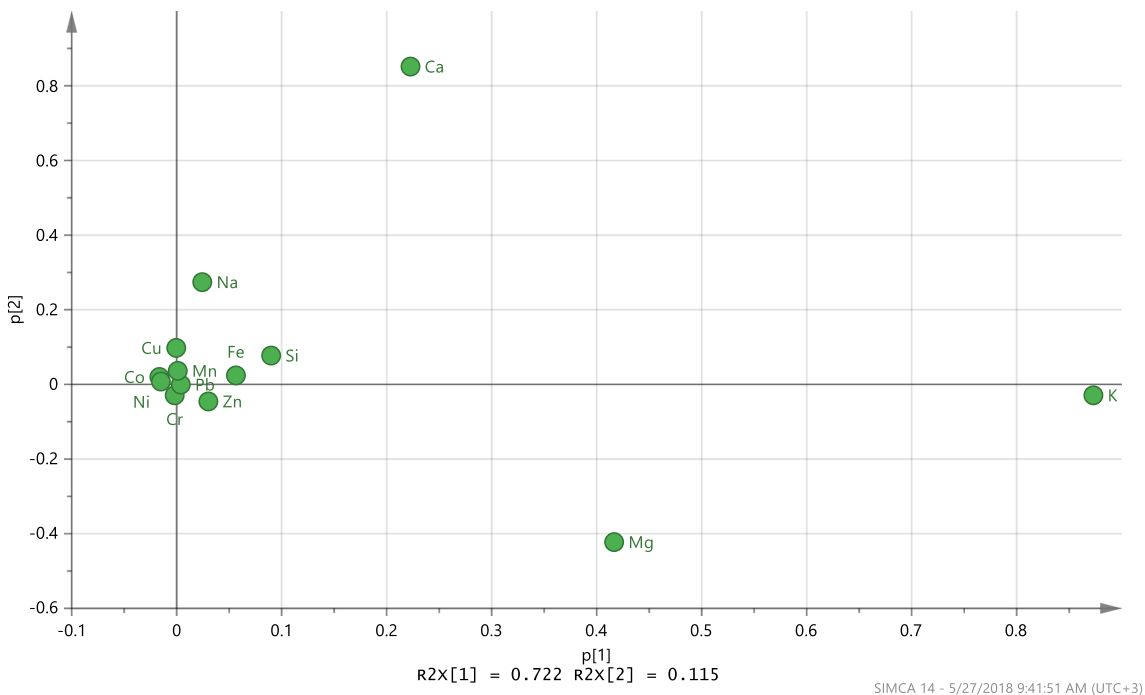


**Fig. 2** Scores plot of the first two principal components showing the distribution of coffee samples from the four zones of Amhara Region

( $\alpha > 0.05$ ), which tests the hypothesis that the means of the functions are equal across groups. The significance values in Table S3, Supplementary material showed that all of the three functions contributed significantly to the model. Additionally, Wilks'  $\lambda$  values indicate the proportion of the total variance in the discriminant functions not explained by differences among

the classified groups of samples. The magnitude of Wilks'  $\lambda$  encompassing the three functions was 0.035, reflecting that the LDA model accounted for almost all of the variation in the data set.

The contribution of each element toward the LDA model was assessed from the computed discriminant function



**Fig. 3** Loadings plot for the first two principal components of the PCA model constructed from the concentration of metals quantified in coffee of Amhara Region



coefficients (Table S4, Supplementary material). The discriminant function coefficients indicate the correlation between each element and the discriminant functions. Elements with higher absolute values of coefficients corresponding to the first discriminant function are the most discriminators among the group of samples. Potassium is the highest contributor to the first function followed by Na and Mg.

Application of LDA with the elemental concentrations resulted in an overall 94.2% correct classification of the coffee beans into their respective regions. The model classified 113 of the 120 samples correctly (Table S5, Supplementary material), while the 7 East Gojam coffee samples were incorrectly classified as Awi (3 samples) and Bahir Dar (4 samples). The distribution of the coffee samples on the plane formed by the first two discriminant function scores is shown in Fig. 4.

The first function has effectively differentiated East Gojam coffees from West Gojam coffees. This function explained 71.6% of the variation in the data set and was strongly associated with K (Table S4, Supplementary material). Hence, the grouping of West Gojam coffees to the positive side of this function can be attributed mainly to their higher content of K relative to East coffees. On the other hand, the second function has effectively differentiated Bahir Dar coffees from Awi coffees. This function was strongly and positively associated with Mn and negatively with Cu (Table S4, Supplementary material). Hence, the grouping of Bahir Dar coffees to the positive side of the second function can be attributed mainly to their higher content of Mn and lower content of Cu relative to Awi coffees.

To check the reliability of the LDA model, its recognition and prediction abilities were computed from samples assigned for training and validation sets. Accordingly, the entire sample set was divided as training (80 samples) and validation sets (40 samples). The prediction ability of the model was found to

be 93.4%, which resulted from the percentage of the samples in the validation set that were correctly classified by using the model constructed in the training step. Similarly, the model had a 94.1% recognition ability, which was computed from the percentage of the samples in the training set that were correctly classified into their respective groups.

## Conclusion

The multi-element composition of coffee samples collected from four major producing zones of Amhara Region, Northwestern Ethiopia were analyzed by ICP-OES and the data were used to discriminate coffees to their respective zones. The elements K, Mg, Ca, and Na were the most abundant constituents in all studied samples. Among the studied elements, K, Mg, and Na were found to be important chemical predictors of production zones of the coffee samples. The LDA model constructed based on the elemental compositions was found to be effective in classifying the coffee samples to their respective production zones. Application of LDA provided a classification model with an overall 94.2% accuracy and 93.4% prediction ability. The results obtained in this study could provide an empirical support for the continued effort for branding the coffee that is produced in Amhara Region like that of other well-known coffee-producing regions of Ethiopia.

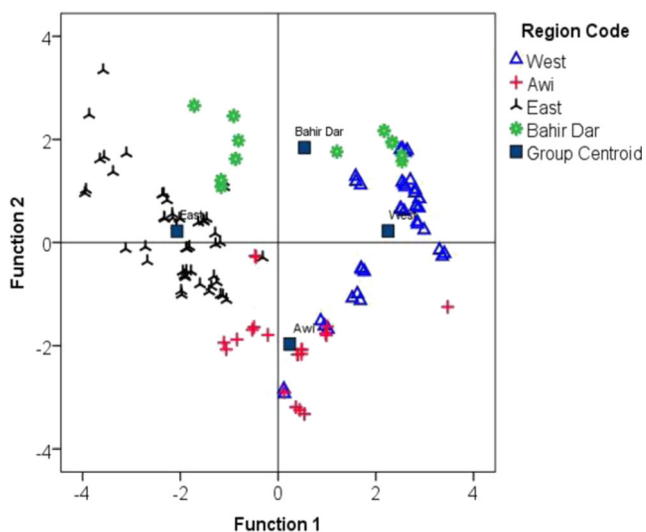
**Acknowledgments** This work was supported by the Ministry of Science and Technology of Ethiopia (3rd round, 2017) for the project title: “Enhancing the Quality and Market Share of the Underutilized Coffee of Amhara Region Through Chemical Analysis, Postharvest Loss Reduction and Mapping for Authentication (Branding)”. The authors are also thankful to Bahir Dar University for the laboratory facilities and the Amhara Region’s Bureau of Agriculture for its logistic support. Furthermore, Mr. Lakachew Emiru from the Amhara Bureau of Agriculture is also acknowledged for his guidance during sample collection.

## Compliance with Ethical Standards

**Conflict of Interest** The authors declare that they have no conflict of interest.

## References

1. Nedzarek A, Torz A, Karakiewicz B, Clark JS, Laszczynska M, Kaleta A, Adler G (2013) Concentrations of heavy metals (Mn, Co, Ni, Cr, Ag, Pb) in coffee. *Acta Biochim Pol* 60(4):623–627
2. Gebeyehu BT, Bikilla SL (2015) Determination of caffeine content and antioxidant activity of coffee. *Am J Appl Chem* 3(2):69–76?
3. Amamo AA (2014) Coffee production and marketing in Ethiopia. *Eur J Bus Manag* 6:109–121
4. Gomez-Ruiz JA, Leake DS, Ames JM (2007) In vitro antioxidant activity of coffee compounds and their metabolites. *J Agric Food Chem* 55:6962–6969



**Fig. 4** Scatter plot showing the distribution of green coffee bean samples on the space created by the scores of the first two discriminant functions

5. Kufa T, Ayano A, Yilma A, Kumela T, Tefera W (2011) The contribution of coffee research for coffee seed development in Ethiopia. *E3 J Agric Res Dev* 1:9–16
6. Farah A (2012) Coffee constituents. In Chu YF (ed) *Coffee: Emerging health effects and disease prevention*. Wiley-Blackwell, Oxford, UK, pp 21–58
7. Mehari B, Redi-Abshiro M, Chandravanshi BS, Combrinck S, McCrindle R (2016) Characterization of the cultivation region of Ethiopian coffee by elemental analysis. *J Anal Lett* 49(15):2474–2489
8. Mehari B, Redi-Abshiro M, Chandravanshi BS, Combrinck S, Atlabachew M, McCrindle R (2016) Profiling of phenolic compounds using UPLC–MS for determining the geographical origin of green coffee beans from Ethiopia. *J Food Compos Anal* 45:16–25
9. Bertrand B, Villarreal D, Laffargue A, Posada H, Lashermes P, Dusser S (2008) Comparison of the effectiveness of fatty acids, chlorogenic acids, and elements for the chemometric discrimination of coffee (*Coffea arabica* L.) varieties and growing origins. *J Agric Food Chem* 56(6):2273–2280
10. Villarreal D, Laffargue A, Posada H, Bertrand B, Lashermes P, Dussert S (2009) Genotypic and environmental effects on coffee (*Coffea arabica* L.) bean fatty acid profile: impact on variety and origin chemometric determination. *J Agric Food Chem* 7(23):11321–11327
11. Ozyigit II, Uras ME, Yalcin IE, Severoglu Z, Demir G, Borkoev B, Salieva K, Yucel S, Erturk U, Solak AO (2019) Heavy metal levels and mineral nutrient status of natural walnut (*Juglans regia* L.) populations in Kyrgyzstan: nutritional values of kernels. *Biol Trace Elem Res* 189(1):277–290
12. Liang J, Wang Q, Huang B (2004) Concentrations of hazardous heavy metals in environmental samples collected in Xiamen, China, as determined by vapor generation non-dispersive atomic fluorescence spectrometry. *Anal Sci* 20:85–88
13. Habte G, Hwang IM, Kim JS, Hong JH, Hong YS, Choi JY, Nho EY, Jamila N, Khan N, Kim KS (2016) Elemental profiling and geographical differentiation of Ethiopian coffee samples through inductively coupled plasma-optical emission spectroscopy (ICP-OES), ICP-mass spectrometry (ICP-MS) and direct mercury analyzer (DMA). *Food Chem* 212:512–520
14. Dos Santos EJ, de Oliveira E (2001) Determination of mineral nutrients and toxic elements in Brazilian soluble coffee by ICP-AES. *J Food Compos Anal* 14:523–531
15. Pohl P, Stelmach E, Welna M, Szymczycha-Madeja A (2013) Determination of the elemental composition of coffee using instrumental methods. *Food Anal Methods* 6:598–613
16. AmorimFilho VR, Polito WL, Gomes Neto JA (2007) Comparative studies of the sample decomposition of green and roasted coffee for determination of nutrients and data exploratory analysis. *J Br Chem Soc* 18(1):47–53
17. Anderson KA, Smith BW (2002) Chemical profiling to differentiate geographic growing origins of coffee. *J Agric Food Chem* 50:2068–2075
18. Krivan V, Barth P, Morales AF (1993) Multi-element analysis of green coffee and its possible use for the determination of origin. *MicrochimActa* 110:217–236
19. Gure A, Chandravanshia BS, Taddese WG (2017) Metals in green coffee beans from major coffee-growing regions of Ethiopia. *J Chem Int* 3(4):359–369
20. Atlabachew M, Combrinck S, Sandasi M, Chen W, Viljoen A (2015) Rapid differentiation of Khat (*Catha edulis* Vahl. Endl.) using single point and imaging vibrational spectroscopy. *VibSpectrosc* 81:96–105
21. Mehari B, Redi-Abshiro M, Chandravanshi BS, Combrinck S, McCrindle R, Atlabachew M (2019) GC-MS profiling of fatty acids in green coffee (*Coffea arabica* L.) beans and chemometric modeling for tracing geographical origins from Ethiopia. *J Sci Food Agric* 99:3811–3823

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.