www.nature.com/scientificreports

scientific reports

OPEN

Mathematical optimization of multilinear and artificial neural network regressions for mineral composition of different tea types infusions

Yusuf Durmus[1]✉, Ayse Dilek Atasoy[2] & Ahmet Ferit Atasoy[3]

The objective of this study was to investigate the change in mineral composition depending on tea variety, tea concentration, and steeping time. Four different tea varieties, black Ceylon (BC), black Turkish (BT), green Ceylon (GC), and green Turkish (GT), were used to produce teas at concentrations of 1, 2, and 3%, respectively. These teas were produced using 7 different steeping times: 2, 5, 10, 20, 30, 45, and 60 min. It was also aimed to optimize the regression equations utilizing these factors to identify parameters conducive to maximizing Zn, K, Cu, Mg, Ca, Na, and Fe levels; minimizing Al content, and maintaining Mn level at 5.3 mg/L. The optimal conditions for achieving a Mn content of 5.3 mg/L in black Turkish tea entailed steeping at a concentration of 1.94% for 11.4 min. Variations in K and Mg levels across teas were inconsistent with those observed for other minerals, whereas variations in Al, Cu, Fe, Mn, Na, and Zn levels exhibited a close relationship. Overall, mineral levels in tea can be predicted through regression analysis, and by mathematically optimizing the resultant equations, the requisite conditions for tea production can be determined to achieve maximum, minimum, or target mineral values.

Keywords Tea, Mathematical optimization, Regression, ANN, PCA, ISOMAP, Hierarchical clustering

The tea plant, *Camellia sinensis*, is a perennial evergreen shrub and is one of the most important crops[1]. It plays a significant role in food trade in exporting and developing countries. It is cultivated in more than 45 countries worldwide, except for North America. Türkiye and Sri Lanka are the major producers after China, India, and Kenya. However, the major tea exporting countries are Kenya, China, India, Sri Lanka, and Vietnam[2]. A large amount of tea is consumed in Türkiye and even imported from Sri Lanka, because of its preference, traditionality, and popularity. Over the past two millennia, brewed tea has emerged as a globally favored non-alcoholic beverage, appreciated worldwide for its distinct taste[3]. The significant economic and social importance of tea is evident given that around 20 billion cups of brewed tea are consumed worldwide on a daily basis[4].

The process of producing black tea encompasses stages such as plucking, indoor withering, leaf disruption, fermentation, drying, sorting, and packing[5]. During the fermentation step, polyphenol oxidase enzymes play a crucial role in facilitating the blackening of leaves. Subsequently, the drying step is implemented to halt the oxidation process, ensuring the production of a durable and consistent tea product[6,7]. However, the fermentation stage is not performed during green tea manufacturing. The activity of tea enzymes is quickly blocked by heating after plucking. Therefore, the properties of green tea are controlled by the endogenous components of the leaves at the time of plucking[8,9].

Tea infusion contains modest quantities of proteins, vitamins, and carbohydrates, while serving as a potential source of essential dietary components such as Mg, K, Na, and Ca, as well as minor and trace minerals such as Zn, Fe, Cu, and Mn, as well as heavy metals such as Cr, Pb, Al, Ni, and Cd[10]. Certain minerals present in tea infusion, including Ca, K, Mg, Fe, Mn, and Zn, are deemed essential for maintaining proper bodily functions.

[1]Department of Gastronomy and Culinary Arts, Faculty of Tourism, Artvin Çoruh University, Artvin, Turkey. [2]Department of Environmental Engineering, Faculty of Engineering, Harran University, 63300 Sanliurfa, Turkey. [3]Department of Food Engineering, Faculty of Engineering, Harran University, 63300 Sanliurfa, Turkey. ✉email: yusufdurmus@artvin.edu.tr

Conversely, other elements such as Pb, Cr, Hg, Ni, and Cd pose health risks or are considered toxic to human health. The concentration of these elements in tea beverages depends on many factors: the species, soil conditions, rainfall, geographical location of the plantation, genetics, and age of the harvested plant[11]. Black tea plants can accumulate heavy metals and toxic compounds from various sources throughout the cultivation and processing stages. These potential sources include the elemental composition of the soil, irrigation water quality, industrial activities, contamination from processing equipment and storage containers, transportation methods and packaging materials, as well as inadvertent introduction through improper harvesting and processing practices[12,13]. Tea type (characteristic of cultivation region), different manufacturing procedures and degrees of fermentation, as well as the method of preparing tea infusions, i.e. the duration of steeping, temperature, and tea/water ratio, also play a significant role in the diffusion of elements from dry tea leaves to solution[14].

Previous studies have reported varying concentrations of elements in tea infusions. An evaluation of element release during black tea brewing revealed K to be the most abundant element, exceeding 10000 µg/g of tea leaves. This was followed by P at 700–1200 µg/g, Mg at 300–700 µg/g, and Ca with 150–300 µg/g. Na, Si, and Zn were all present in lower amounts, with levels between 6 and 50 µg/g. Trace elements, including B, Ba, Fe, and Pb, were detected at concentrations below 1 µg/g. Sr, Co, and Gd were found in quantities less than 1 µg/g[15]. Al exhibited with levels documented between 0.06 and 16.82 mg/L. Cu follows with a range of 0.02–40.0 mg/L, and F varies from 0.2 and 4.54 mg/L. Mn concentrations were reported to be between 0.1 to 250 mg/L. As (< 1.53 µg/L), Cd (< 0.79 µg/L), and Cr (< 43.2 µg/L) were typically detected in trace amounts, while Ni was rarely found, with a maximum reported level of 0.16 mg/L[12].

It is feasible to regulate tea content by mathematically optimizing production conditions such as tea concentration, steeping time, and temperature. Based on optimization algorithms and statistical techniques, mathematical optimization scans the space of potential variable subsets and selects the subset that is optimal or nearly optimal in reference to an objective function[16].

The objective of this study was to investigate the change in mineral composition depending on tea variety, tea concentration, and steeping time. It was also aimed to optimize the regression equations using these variables to find the parameters that allow us to reach the optimal mineral contents. The study targeted to optimize the levels of essential elements, including Zn, K, Cu, Mg, Ca, Na, and Fe, while minimizing the presence of Al. Additionally, Mn concentration was targeted for stabilization at 5.3 mg/L.

Materials and methods

Tea sampling, infusion, and analyzing

Black Ceylon (BC), black Turkish (BT), green Ceylon (GC), and green Turkish (GT) tea samples were collected from local stores in Sanliurfa, Türkiye. Prior to tea steeping, all the samples were dried at 60 °C for 12 h until reaching a constant weight was reached.

For tea infusions, three different tea concentrations (1, 2, and 3%) were used. 1.00, 2.00, and 3.00 g of each tea leaf was kept in contact with 100 mL of boiling distilled water (traditional beverage method applied by tea drinkers) and allowed to steep for 2, 5, 10, 20, 30, 45, and 60 min. The procedure described in Atasoy et al.[6] was followed for the preparation of tea infusions and other processes. The tea infusion was stirred with a glass rod, covered with watch glasses to facilitate the extraction of tea leaf components, and steeped for specified durations. Following steeping, the infusion was filtered through Whatman 42 filter paper (125 mm diameter) into a test tube and subsequently cooled. Tea steeping was repeated 2 times on different days for each sample.

Tea infusion samples were analyzed immediately for Al, Ca, Cd, Cr, Cu, Hg, Fe, K, Mg, Mn, Na, Pb, and Zn via inductively coupled plasma-optical emission spectrometry (ICP-OES) (Perkin Elmer Optima 5300 DV model). Pyrolytically coated tubes were employed for sample analysis within an inert argon (Ar) atmosphere. The instrument operating parameters were optimized daily to enhance sensitivity and stability. These parameters included: instrument power (1.35×10^3 W), plasma argon gas flow rate (15 L min^{-1}), auxiliary argon gas flow rate (1.5 L min^{-1}), nebulizer gas flow rate (0.75 L min^{-1}), solution uptake rate (1.5 mL min^{-1}), and dwell time (100 min). All the tests were conducted in duplicate. Cd, Cr, Hg and Pb were not evaluated due to their extremely low concentrations. The analytical methods were further validated through the determination of limit of detection (LOD) and limit of quantification (LOQ) for each target analyte. The LOD and LOQ values for Cd, Cr, Hg and Pb were determined as 0.027–0.081, 0.032–0.096, 0.026–0.078, and 0.030–0.090, respectively. The architecture diagram of the research framework is illustrated in Fig. 1. Collecting of tea leaf samples and methods were carried out in accordance with the IUCN policy statement on research involving species at risk of extinction.

Multilinear regression (MLR)

This study employed two approaches, MLR and ANN, to estimate mineral content within the samples. The performance of these models was then evaluated through a comparative analysis of their respective metrics. Notably, both models were optimized to achieve the previously established mineral content targets. The main purpose of MLR is to build a mathematical model between the independent variables and a single dependent variable and to provide predictions by revealing the relationships between variables[17]. The MLR model with n number of independent variables is shown in Eq. 1.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n + \varepsilon \tag{1}$$

where Y is the dependent variable, $X_1, X_2, \ldots, X_n$ are the independent variables, $\beta_0, \beta_1, \beta_2, \ldots, \beta_n$ are the unknown coefficients and ε is the error term[18]. In this study, the concentration, time, and type of tea were designated as the independent variables, and the mineral contents of the tea samples were designated as the dependent variables.

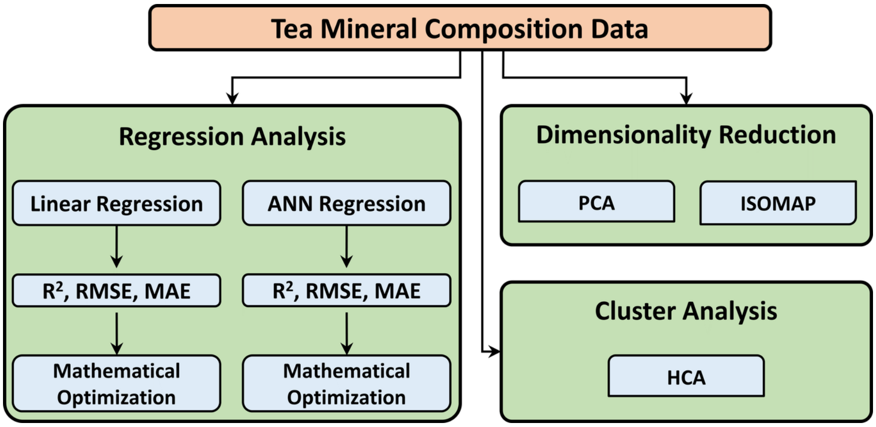

Figure 1. Architecture diagram of the research framework. *ANN* artificial neural network, *R^2* coefficient of determination, *RMSE* root mean square error, *MAE* mean absolute error, *PCA* principal component analysis, *ISOMAP* isometric mapping, *HCA* hierarchical cluster analysis.

Artificial neural network (ANN) regression

A multilayer perceptron (MLP) type ANN was used for the estimation of mineral contents. The architecture of the ANN model is shown in (Fig. 2). The number of inputs was set to 5, the number of hidden layers as 4, the number of neurons in hidden layers as 18, 8, 12, and 18 and the number of outputs as 1. The "Adam" solver was used to optimize the weights with a learning rate of 0.0112. Leaky rectified linear unit (LeakyRelu) was used as the activation function and the maximum iteration for the solver was set to 100[19].

Performance metrics

Three different performance metrics were used to evaluate the MLR and ANN models; coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE). These metrics are shown in Eqs. 2, 3, and 4.

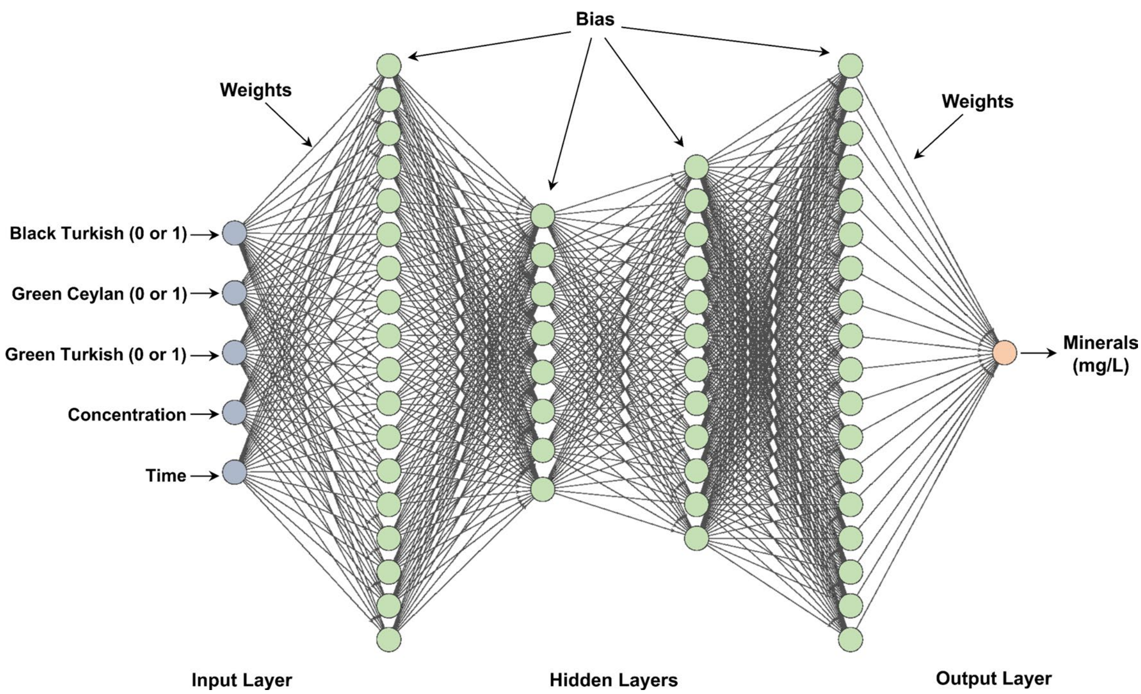

Figure 2. Representative MLP-type ANN architecture for regression of mineral contents. *If Black Turkish = 0, Green Ceylon = 0, and Green Turkish = 0, tea type is Black Ceylon.

$$R^2 = 1 - \frac{\sum_i^n \left(y_i - \hat{y}_i\right)^2}{\sum_i^n \left(y_i - \overline{y}\right)^2} \tag{2}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n \left(y_i - \hat{y}_i\right)^2}{n}} \tag{3}$$

$$MAE = \frac{1}{n}\sum_{i=1}^n \left|y_i - \hat{y}_i\right| \tag{4}$$

where y_i is the ith observed value, $\overline{y}$ is the average of observed values, $\widehat{y}_i$ is the predicted value for the ith observed value, and n is the number of observed values.

Mathematical optimization

Mathematical optimization is the process of finding the maximum or minimum point of an objective function by optimally adjusting values across a set of inputs[20]. In this study, two objective functions were produced with regression analysis: MLR and ANN models. Concentration, time, and tea type are the independent variables for which the regression functions' optimum values were aimed to be determined, and the minerals contents are the dependent variables whose values are aimed to be maximized or minimized. It was aimed to maximize Zn, K, Cu, Mg, Ca, Na, and Fe levels; minimize Al level; and stabilize Mn level at 5.3 mg/L. Atasoy, et al.[6] reported that the adequate intake of Mn is 5.3 mg/day. In this optimization problem, tea type is a categorical variable while tea concentration and time are continuous variables. Since tea type is a categorical variable, each tea was binary coded (e.g., if tea is BC, BC = 1 and all others 0). The tree-structured Parzen Estimator algorithm described by Bergstra, et al.[21] was used as the optimization algorithm. The numbers of trials for optimizing the MLR and ANN objective functions were 500 and 100, respectively.

Principal component analysis (PCA) and isometric mapping (ISOMAP)

PCA was employed to analyze the mineral composition of the tea samples, aiming to ascertain the linear relationships among the different types of tea and among the minerals themselves. The dataset underwent standardization and decomposition into principal components (PCs), with a biplot generated utilizing the first two PCs[22]. The nonlinear dimensionality reduction method ISOMAP was used to reveal the nonlinear relationships between tea types. The ISOMAP model was built with 5 nearest neighbors, and the "Euclidean" distance metric shown in Eq. 5 Ref.[23].

$$|X - Y| = \sqrt{\sum_{i=1}^n \left(x_i - y_i\right)^2} \tag{5}$$

where x_i and y_i are the ith observed values and n is the number of observed values.

Hierarchical cluster analysis (HCA)

HCA was conducted to cluster the minerals, and the results are depicted in a dendrogram. The "Ward" method, also referred to as the "Incremental algorithm" as delineated in Eq. 6, was employed to compute the distance between two clusters. This method aims to minimize the variance between the clusters being merged. Additionally, the "Euclidean" metric specified in Eq. 5 was used to calculate the distance between two points.

$$D(A,B) = \sum_{x \in A\cup B} D(x, \mu_{A\cup B})^2 \tag{6}$$

where A and B are the first and second clusters, respectively, and x is the object belonging to each cluster.

Statistical analysis

The experimental design was constructed on three factors. These factors were the type of tea with four levels (BC, BT, GC, or GT), the concentration of tea with three levels (1%, 2%, or 3%), and steeping time with 7 levels (2, 5, 10, 20, 30, 45, or 60 min). Mineral composition analysis was carried out in duplicate ($n = 4 \times 3 \times 7 \times 2$). Statistical analysis was performed using Minitab 20 statistical software (Minitab, Inc., 2021, State College, PA, USA). Analysis of variance (ANOVA) was conducted using the general linear model at a significance level of $\alpha = 0.05$. Following this, any significant differences ($p < 0.05$) observed among the means were subjected to further examination using Tukey's multiple range tests.

Results and discussion

Mineral composition

The mineral compositions of the four tea types infused at concentrations of 1, 2, and 3% are shown in (Table 1). As expected, the amounts of all the minerals increased with increasing tea concentration. The mineral contents

Sample	Concentration (%)	Al	Ca	Cu	Fe	K	Mg	Mn	Na	Zn
BC	1	2.52h	4.23h	0.036fg	0.062h	186.5c	10.26c	1.64e	1.87ef	0.14e
BC	2	4.32gh	4.82gh	0.041efg	0.072gh	263.2b	18.33b	3.28de	3.10cde	0.23cd
BC	3	6.29fg	5.86e	0.070bcd	0.089fgh	429.9a	26.21a	3.71de	4.21bc	0.35a
BT	1	4.65gh	5.11fg	0.031g	0.122def	151.7cde	7.40cde	3.62de	1.42f	0.19cde
BT	2	9.71de	6.26de	0.062cde	0.172abc	258.9b	14.90b	6.80bc	2.43def	0.25bc
BT	3	15.14ab	7.55b	0.079bc	0.199ab	381.2a	20.49a	9.70a	3.51cd	0.35a
GC	1	3.47h	6.13de	0.058cdef	0.128cdef	99.8e	5.74e	3.12de	2.48def	0.17de
GC	2	7.33ef	7.07bc	0.092b	0.164bcd	174.6cd	11.41cd	5.47cd	4.14bc	0.24bc
GC	3	12.87bc	8.30a	0.144a	0.212a	277.9b	16.86b	8.33ab	5.62a	0.30ab
GT	1	4.91fgh	4.92fgh	0.052defg	0.109efg	122.9de	7.24de	3.30de	2.47def	0.19cde
GT	2	11.6cd	5.56ef	0.049defg	0.144cde	250.5b	14.31b	6.60b	3.69cd	0.24bc
GT	3	17.54a	6.77cd	0.054def	0.209ab	305.2b	20.31b	10.47b	5.26ab	0.36a

Table 1. Mineral composition (mg/L) of tea samples prepared in different concentrations. *BC* black Ceylon, *BT* black Turkish, *GC* green Ceylon, *GT* green Turkish. Values in the same column with the same letter are not significantly ($p < .05$) different.

in green teas were found to be significantly ($p < 0.05$) higher than those in black tea, with the exception of K, Mg, and Zn. While Al, Fe, and Mn were significantly ($p < 0.05$) higher in Turkish tea, only Cu was significantly ($p < 0.05$) higher in Ceylon tea. Olivier, et al.[24] also noted that while Zn levels were comparable between black and green tea, black tea had higher Mg and K contents. The tea types and concentrations were steeped for 2, 5, 10, 20, 30, 45, and 60 min, and it was found that the mineral contents increased in step with the steeping period. The amount of Ca in black tea was higher than the amount of Ca found by Derun[25] in black tea. However, the amount of Ca measured by the same researcher was higher when green tea was taken into account. The Cu concentrations in black and green tea were determined to be 0.05 and 0.07 mg/L, respectively, while Derun[25] determined them to be 2.6 and 2.4 mg/L, respectively. The average Fe content was 0.12 and 0.16 mg/L for black and green tea, respectively. These findings were observed to be exceptionally high, rather low, or comparable to those of some previous investigations[26–30]. The diversity of the mineral contents is influenced by variations in tea infusion techniques. Mineral levels might vary according to the plant type, steeping time, temperature, granule size, and concentration[14].

Principal component analysis (PCA) was used to explain the variation among the minerals and changes in mineral composition based on tea variety. PCA is a method aimed at reducing the number of variables in a dataset by computing principal components (PCs) along directions that exhibit the greatest deviation from the original data[31]. In this study, two distinct PCA analyses were conducted: one focused on tea types and the other on mineral compositions. The biplot of PCA results for the tea types is shown in (Fig. 3A). The first two PCs explained 86.0% of the total variance. The BC and GC teas were separated into two clusters and aligned along the x-axis. This is due to the fact that BC and GC teas exhibit substantially different mineral compositions compared to other teas, as well as to each other. Additionally, the distributions of the BT and GT teas within one another suggests that their mineral compositions are similar. In the case of Turkish tea, black or green tea did not cause major differences, whereas in the case of Ceylon teas, the processing of the tea considerably changed the mineral content.

Isometric mapping (ISOMAP) is another dimensionality reduction method that can be used to explain nonlinear interactions. ISOMAP is a type of manifold learning which looks for a lower-dimensional embedding that keeps the geodesic distances between all of the points[23]. Since PCA is a linear dimensionality reduction method and has a total unexplained variance of 14%, a nonlinear method was considered appropriate. Similar outcomes were reached with PCA, even though the tea varieties displayed a distinct distribution on the plot with ISOMAP (Fig. 3B). The BC and GC teas formed two separate clusters on the plot, and the BT and GT teas had comparable mineral contents. PCA was used to reveal patterns of relationships between minerals (Fig. 3C). Zn, Al, Mn, Na, Fe, and Cu were highly correlated. K, Mg, and Ca were located at distant points in the plot, and accordingly, the variances of these minerals were not similar to other minerals. This is because K exhibited quite different variances for tea variety and steeping time. Table 1 shows that K and Mg have exactly the same post hoc letters, but these two minerals are located in different places in (Fig. 3C). This was due to the interactions between tea variety/concentration and tea variety/time. The variation in Mg concentration was almost the same in the BT and GT samples, while K increased more in the BT samples. Similarly, the variation in Mg with time was almost the same for the BT and GT teas, while K was more abundant in the BT teas. Since almost all the variance in the data could be explained by PCA (99.97%), ISOMAP analysis was considered unnecessary.

The relationships between minerals were also investigated using hierarchical cluster analysis (HCA). Increased cutting points suggest a greater degree of separation between clusters on the dendrogram, whereas lower cutting points signify closer associations in HCA[32]. A dendrogram can be used to visually count the number of clusters that are present. The relationship between the minerals was graphically represented by a dendrogram (Fig. 3D). Zn, Cu, and Fe exhibited close proximity on the dendrogram, indicating a high degree of similarity among these three minerals. Unlike PCA, Ca and Mn were found to have close similarities.

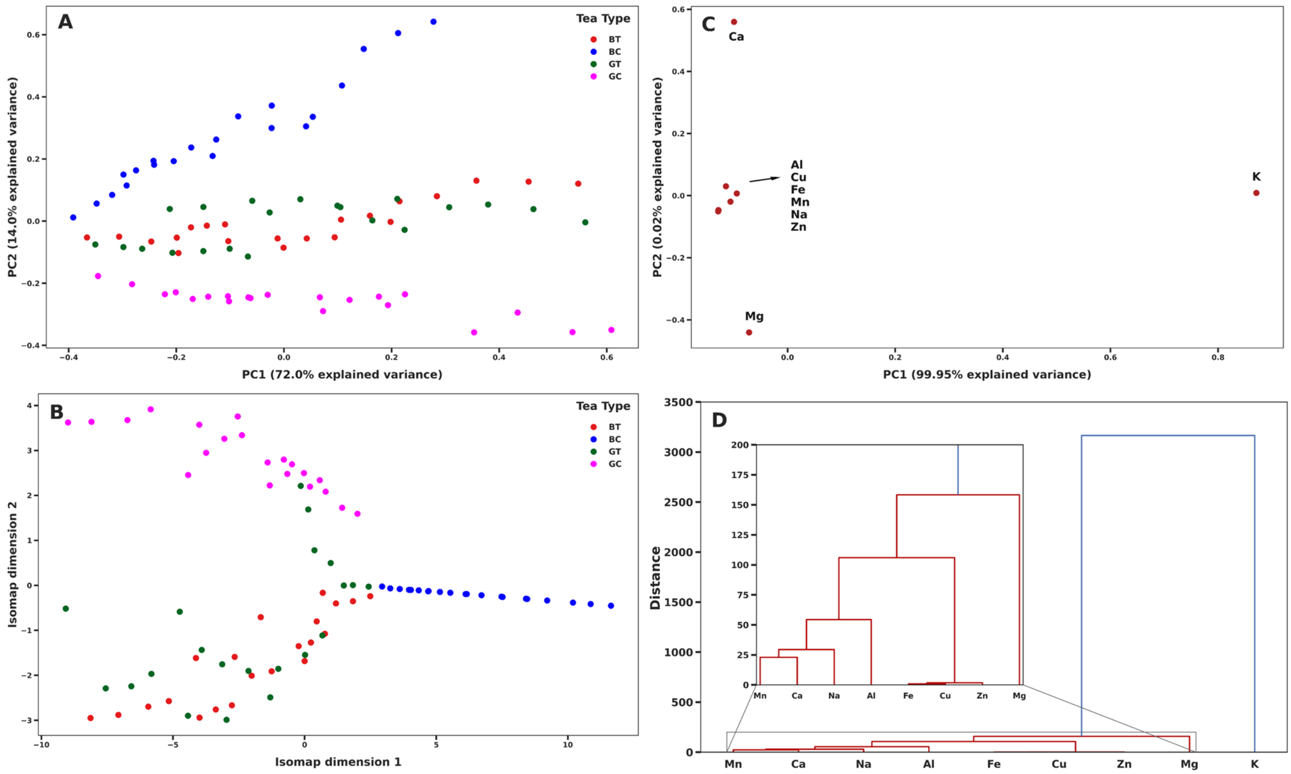

Figure 3. Biplots of PCA (**A**) and ISOMAP (**B**) illustrating the overall variation among the tea types, Biplot of PCA illustrating the overall variation among the minerals (**C**) and Dendrogram of cluster analysis based on minerals (**D**). BC black Ceylon, BT black Turkish, GC green Ceylon, GT green Turkish.

Regression analysis

Regression analysis is the mathematical procedure that computes parameters describing a relationship between two or more variables. Regression analysis provides a function that allows the value of the dependent variable to be determined using the independent variables[33]. In this study, two different regression analyses were performed and the regression equations are given in (Table 2). These equations are first-order linear equations, and by altering production factors such as tea variety, tea concentration, and steeping time, one can determine the quantity of each mineral. The R^2, RMSE, and MAE values for each equation are given in (Table 3). Within the multiple linear regression (MLR) equations, the most notable R^2 values were attained for K (0.9276) and Ca (0.9463), signifying strong correlations. Conversely, the weakest correlations were observed for Al (0.8426) and Cu (0.6515), indicating lower explanatory power. For the ANN, the R^2 values of the equations were quite high. The Mg (0.9890) and Na (0.9878) equations had the highest R^2 values, while the Cu (0.9076) and Al (0.9431) equations had low R^2 values. Considering that R^2 ranges between 0 and 1 and that a R^2 closer to 1 is more reliable, we can conclude that the ANN is quite successful.

Response	Model equation*
Al	= 4.53741071* C + 0.06750285 * T + 5.4587619 * BT + 3.51252381 * GC + 6.97483333 * GT−6.357677227827578
Ca	= 1.01193393 * C + 0.02485085 * T + 1.3354381 * BT + 2.19806667 * GC + 0.7754 * GT + 2.3354683222910264
Cu	= 0.02141574 * C + 0.00041644 * T + 0.00802373 * BT + 0.04871869 * GC + 0.00223009 * GT−0.0037603238663258542
Fe	= 0.03591782 * C + 0.00152479 * T + 0.09012643 * BT + 0.09357882 * GC + 0.07960263 * GT—0.03490458399165197
K	= 104.15169643 * C + 1.9788357 * T−29.295 * BT−109.12619048 * GC−66.99452381 * GT + 36.282120537969575
Mg	= 6.65731786 * C + 0.22008861 * T−3.99711905 * BT−6.92673571 * GC−4.31009524 * GT−0.4587652964823494
Mn	= 2.56860268 * C + 0.08107707 * T + 3.82796429 * BT + 2.76560714 * GC + 3.91236905 * GT−4.253944389413339
Na	= 1.29484821 * C + 0.04418365 * T−0.60716667 * BT + 1.02128571 * GC + 0.74728571 * GT−0.6152089179260685
Zn	= 0.0864491071 * C + 0.00235847647 * T + 0.0248095238 * BT−0.0000714285714 * GC + 0.0250071429 * GT−0.0060316018682909645

Table 2. MLR regression model equations as a function of independent variables. *C* concentration (%), *T* time (min), *BT* black Turkish (0 or 1), *GC* green Ceylon (0 or 1), *GT* green Turkish (0 or 1). *If BT = 0, GC = 0, and GT = 0, tea type is black Ceylon.

Regression	Mineral	R^2	RMSE	MAE	Mineral content (mg/L)	Concentration (%)	Time (min)	Tea type
MLR	Al	0.8426	2.0445	1.6375	1.934	1.33	33.3	BC
	Ca	0.9463	0.2990	0.2316	9.060	3.00	60	GC
	Cu	0.6515	0.0203	0.0165	0.134	3.00	60	GC
	Fe	0.8893	0.0202	0.0163	0.258	3.00	60	GC
	K	0.9276	28.6229	23.0744	467.46	3.00	60	BC
	Mg	0.8987	2.4953	1.9927	32.719	3.00	60	BC
	Mn	0.8446	1.3260	1.0821	5.30	1.20	31.7	GT
	Na	0.8588	0.6168	0.4756	6.942	3.00	60	GC
	Zn	0.8880	0.0305	0.0241	0.432	3.00	60	GT
ANN	Al	0.9854	0.6232	0.4153	2.136	1.00	2	GC
	Ca	0.9692	0.2263	0.1696	9.241	3.00	60	GC
	Cu	0.9076	0.0104	0.0078	0.183	3.00	60	GC
	Fe	0.9824	0.0081	0.0062	0.278	3.00	60	GC
	K	0.9431	25.3699	19.6957	489.38	3.00	60	BC
	Mg	0.9890	0.8222	0.6019	36.535	3.00	60	BC
	Mn	0.9874	0.3776	0.2914	5.30	1.94	11.4	BT
	Na	0.9878	0.1817	0.1438	8.177	3.00	60	GC
	Zn	0.9823	0.0121	0.0097	0.445	3.00	60	GT

Table 3. Performance outcomes and mathematical optimization results of regression models for the prediction of mineral contents. *MLR* multilinear regression, *ANN* artificial neural network, R^2 coefficient of determination, *RMSE* root mean square error, *MAE* mean absolute error, *BC* black Ceylon, *BT* black Turkish, *GC* green Ceylon, *GT* green Turkish.

The RMSE and MAE are highly beneficial because they provide the error rate of the equation in the same units as the target variable, whereas the R^2 allows comparisons of different target variables. If deviations in a dataset are likely to be large, i.e. there are outliers, the RMSE is more sensitive in indicating these errors. The MAE, on the other hand, gives more interpretable results when the variance in the deviations is small. The variance in the errors may be identified by combining the MAE with the RMSE. The RMSE will always be greater than the MAE or equal to it; the greater the difference between the two is, the greater the variance in the individual errors in the data[34]. The MAE of Al for the MLR equation was found to be 1.6375, which means that this equation calculates the amount of Al with an error of 1.6375 mg/L (Table 3). The error levels in the regression with ANN are approximately 2.5 times smaller than those in MLR for both the RMSE and MAE for all minerals. Al exhibited the greatest decrease in MAE with the use of ANN instead of MLR, at approximately 4 times, while K exhibited the least decrease in MAE, at approximately 1.2 times.

Mathematical optimization

Many studies have been conducted on the importance of minerals in tea for human nutrition and for the ability to meet human needs[6,24,25,30,35,36]. It is clear that teas contain only trace levels of minerals; thus, they do not significantly help people satisfy their needs for these nutrients. For this reason, in the present study, the tea production conditions required for minerals other than Al and Mn to reach maximum values were determined. Mathematical optimization can be used to maximize or minimize a target variable, fix it to a certain value, or ensure that it is within a certain range. Due to the toxic effects of Al on humans, the aim was to minimize the amount of Al[37,38]. Mn is present at significant levels in most tea varieties and can meet human needs. Since the adequate intake of Mn was reported to be 5.3 mg/L for adults, this value was fixed by optimization, and the required production conditions were determined[6].

Figure 4 shows the optimization history of Fe and the change in Fe content as a function of concentration and time. The Fe content increased with increasing tea concentration and steeping time. This also applies to all other minerals. The plots for the other minerals could not be given due to the lack of space in the paper. The highest Fe value (0.258 mg/L) for MLR optimization was attained in trial 241, whereas for ANN, the highest Fe value (0.278 mg/L) was obtained in trial 86 (Fig. 4A,B). Table 3 shows the optimization results for each mineral according to the regression models. Although the production parameters were found to be the same in both regression models, it is more possible to reach a Fe content closer to 0.278 mg/L when tea is produced in these parameters because the ANN model is more reliable as shown in (Table 3). The optimum tea concentration of 3%, and a steeping time of 60 min was required for the maximum values of Zn, K, Cu, Mg, Ca, Na, and Fe. This makes sense because the mineral content is directly proportional to the concentration and extraction time. The ability to pick the right type of tea is the main advantage of optimization in this situation. For instance, Ca, Cu, Fe, and Na require GC tea, K and Mg require BC tea, and Zn requires GT tea for the ANN model. In the MLR model, 1.33% concentration and 33.3 min are required for Al to reach a minimum (1.934 mg/L), while 1% concentration and 2 min are required in the ANN (2.136 mg/L). The values of 1% concentration and 2 min for minimum Al are more appropriate because the performance metrics of the ANN model are superior, and it is more convenient to use less tea and time to restrict extraction. The adequate intake of Mn is 5.3 mg/L, and

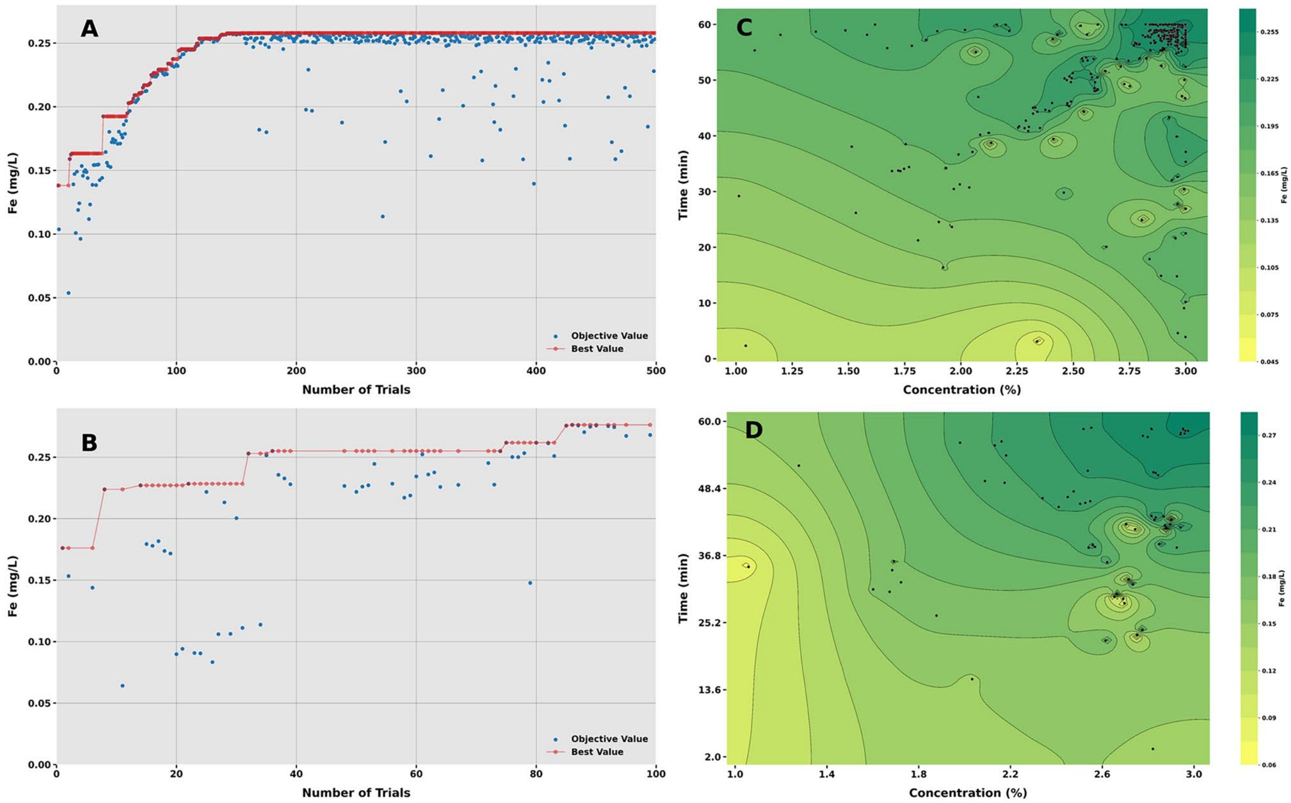

Figure 4. Optimization history for maximizing the Fe content (**A**,**B**) and variation of Fe content with time and concentration during optimization (**C**,**D**). A and C = MLR regression, B and D = ANN regression.

to produce a tea containing this amount of Mn, it is necessary to use 1.94% of the total amount of BT tea and apply a steeping time of 11.4 min. Considering that 1 cup holds 250 mL of liquid, 4 cups of tea will provide an adequate intake of Mn.

Conclusion

In this study two different regression models were successfully built for each mineral using the MLR and ANN methods. The R^2, RMSE, and MAE values of the models built with ANN were much superior to those of MLR. According to the MLR regression equations, the R^2 values ranged between 0.6515 and 0.9463, with an average of 0.8608, while according to the ANN, the R^2 values ranged between 0.9076 and 0.9890, with an average of 0.9705. Mathematical optimization was performed for the regression models. Mathematical optimization was used to determine the tea type, tea concentration, and steeping time parameters needed for the highest Zn, K, Cu, Mg, Ca, Na, and Fe contents. Similarly, the tea production parameters required for a minimum Al content and Mn content of 5.3 mg/L were determined by optimization. The results suggest that a tea concentration of 3% and a steeping time of 60 min are optimal for maximizing the mineral content. It was determined that 1.94% concentration, 11.4 min, and BT tea were required to reach the adequate intake of 5.3 mg/L of Mn. PCA and ISOMAP analyses revealed distinct mineral compositions for the BC and GC teas in comparison to both each other and the Turkish teas. Processing Turkish teas from green to black did not drastically alter the mineral content. In this study, it was ensured that the mineral contents in tea can be calculated with regression equations. The mathematical optimization established the production parameters necessary for the minerals to have a maximum, minimum, or desired value. Similar studies can be carried out for various tea types and tea compounds, utilizing of mathematical optimization and machine learning approaches.

Data availability

Data is provided within the supplementary information files.

Received: 23 February 2024; Accepted: 1 August 2024
Published online: 07 August 2024

References

1. Atasoy, A., Yesilnacar, M. & Atasoy, A. Evaluation of fluoride concentration and daily intake by human from tea infusions. *Harran Tarım ve Gıda Bilimleri Dergisi* **20**, 1–6. https://doi.org/10.29050/harranziraat.194280 (2016).
2. FAO. *International tea market: market situation, prospects and emerging issues*, <https://www.fao.org/3/cc0238en/cc0238en.pdf> (2022).

3. Atasoy, A., Yesilnacar, M. I. & Atasoy, A. Essential element contents of Turkish black tea. In *Non-Alcoholic Beverages* (eds Grumezescu, A. M. & Holban, A. M.) (Elsevier, 2019).
4. Mandiwana, K. L., Panichev, N. & Panicheva, S. Determination of chromium (VI) in black, green and herbal teas. *Food Chem.* **129**, 1839–1843. https://doi.org/10.1016/j.foodchem.2011.05.124 (2011).
5. Polechońska, L., Dambiec, M., Klink, A. & Rudecki, A. Concentrations and solubility of selected trace metals in leaf and bagged black teas commercialized in Poland. *J. Food Drug Anal.* **23**, 486–492. https://doi.org/10.1016/j.jfda.2014.08.003 (2015).
6. Atasoy, A. D., Yesilnacar, M. I., Yildirim, A. & Atasoy, A. F. Nutritional minerals and heavy metals in tea infusions and daily intake of human body. *Turkish J. Agric. Food Sci. Technol.* **7**, 234–239. https://doi.org/10.24925/turjaf.v7i2.234-239.2217 (2019).
7. Rumaisa, N., Hanim, M. N. & Hii, C. Superheated steam drying of black tea and quality improvement. *Int. J. Food Eng.* https://doi.org/10.1515/ijfe-2018-0185 (2018).
8. Mehra, A. & Baker, C. Leaching and bioavailability of aluminium, copper and manganese from tea (*Camellia sinensis*). *Food Chem.* **100**, 1456–1463. https://doi.org/10.1016/j.foodchem.2005.11.038 (2007).
9. Wang, M., Yang, R. & Zhao, W. Effects of heat and pulsed electric fields on bioactive components and color of green tea infusions. *Int. J. Food Eng.* https://doi.org/10.2202/1556-3758.1332 (2008).
10. Szymczycha-Madeja, A., Welna, M. & Pohl, P. Elemental analysis of teas and their infusions by spectrometric methods. *TrAC, Trends Anal. Chem.* **35**, 165–181. https://doi.org/10.1016/j.trac.2011.12.005 (2012).
11. Klink, A., Dambiec, M., Polechońska, L. & Rudecki, A. Evaluation of macroelements and fluorine in leaf and bagged black teas. *J. Food Meas. Charact.* **12**, 488–496. https://doi.org/10.1007/s11694-017-9662-7 (2018).
12. Karak, T. & Bhagat, R. Trace elements in tea leaves, made tea and tea infusion: A review. *Food Res. Int.* **43**, 2234–2252 (2010).
13. Hu, C., Zhang, X., Zhan, N. & Liu, Y. Current status and health risk assessment of heavy metals contamination in tea across China. *Toxics* **11**, 662 (2023).
14. Saletnik, B. *et al.* Effect of infusion time and addition of lemon juice on the mobility of selected macroelements and aluminium during aqueous extraction of quality brands of leaf tea. *J. Elementol.* https://doi.org/10.5601/jelem.2017.22.3.1449 (2018).
15. Natesan, S. & Ranganathan, V. Content of various elements in different parts of the tea plant and in infusions of black tea from southern India. *J. Sci. Food Agric.* **51**, 125–139 (1990).
16. Yun, Y. H., Li, H. D., Deng, B. C. & Cao, D. S. An overview of variable selection methods in multivariate analysis of near-infrared spectra. *TrAC Trends Anal. Chem.* **113**, 102–115. https://doi.org/10.1016/j.trac.2019.01.018 (2019).
17. Kayakus, M. & Terzioglu, M. Yapay sinir ağlari ve çoklu doğrusal regresyon kullanarak emeklilik fonu net varlik değerlerinin tahmin edilmesi. *Bilişim Teknolojileri Dergisi* **14**, 95–103. https://doi.org/10.17671/gazibtd.742995 (2021).
18. Aiken, L. S., West, S. G. & Pitts, S. C. Multiple linear regression. In *Handbook of Psychology* (eds Aiken, L. S. *et al.*) (Wiley, 2003).
19. Sabanci, K., Kayabasi, A. & Toktas, A. Computer vision-based method for classification of wheat grains using artificial neural network. *J. Sci. Food Agric.* **97**, 2588–2593. https://doi.org/10.1002/jsfa.8080 (2017).
20. Zomorrodi, A. R., Suthers, P. F., Ranganathan, S. & Maranas, C. D. Mathematical optimization applications in metabolic networks. *Metab. Eng.* **14**, 672–686. https://doi.org/10.1016/j.ymben.2012.09.005 (2012).
21. Bergstra, J., Bardenet, R., Bengio, Y. & Kégl, B. Algorithms for hyper-parameter optimization. *Adv. Neural Inform. Process. Syst.* **24** (2011).
22. Jolliffe, I. T. & Cadima, J. Principal component analysis: A review and recent developments. *Philos. Trans. A Math. Phys. Eng. Sci.* **374**, 20150202. https://doi.org/10.1098/rsta.2015.0202 (2016).
23. Tenenbaum, J. B., de Silva, V. & Langford, J. C. A global geometric framework for nonlinear dimensionality reduction. *Science* **290**, 2319–2323. https://doi.org/10.1126/science.290.5500.2319 (2000).
24. Olivier, J., Symington, E. A., Jonker, C. Z., Rampedi, I. T. & Van Eeden, T. S. Comparison of the mineral composition of leaves and infusions of traditional and herbal teas. *S. Afr. J. Sci.* **108**, 1–7. https://doi.org/10.4102/sajs.v108i1/2.623 (2012).
25. Derun, E. M. Determination of essential mineral concentrations in some Turkish teas and the effect of lemon addition. *Food Sci. Biotechnol.* **23**, 671–675. https://doi.org/10.1007/s10068-014-0091-7 (2014).
26. Gezgin, S., Ozcan, M. M. & Atalay, E. Determination of minerals extracted from several commercial teas (*Camellia sinensis*) to hot water (infusion). *J. Med. Food* **9**, 123–127. https://doi.org/10.1089/jmf.2006.9.123 (2006).
27. Garbowska, B., Wieczorek, J., Garbowska, B., Wieczorek, Z. & Polak-Śliwińska, M. The content of minerals, bioactive compounds and anti-nutritional factors in tea infusions. *J. Elementol.* https://doi.org/10.5601/jelem.2017.22.2.1306 (2017).
28. Gillies, M. E. & Birkbeck, J. A. Tea and coffee as sources of some minerals in the New Zealand diet. *Am. J. Clin. Nutr.* **38**, 936–942. https://doi.org/10.1093/ajcn/38.6.936 (1983).
29. Musa Özcan, M., Ünver, A., Uçar, T. & Arslan, D. Mineral content of some herbs and herbal teas by infusion and decoction. *Food Chem.* **106**, 1120–1127. https://doi.org/10.1016/j.foodchem.2007.07.042 (2008).
30. Gallaher, R. N., Gallaher, K., Marshall, A. J. & Marshall, A. C. Mineral analysis of ten types of commercially available tea. *J. Food Compos. Anal.* **19**, S53–S57. https://doi.org/10.1016/j.jfca.2006.02.006 (2006).
31. Achten, E. *et al.* Classification of grain maize (*Zea mays* L.) from different geographical origins with FTIR spectroscopy—A suitable analytical tool for feed authentication?. *Food Anal. Methods* **12**, 2172–2184. https://doi.org/10.1007/s12161-019-01558-9 (2019).
32. Malalgoda, M., Manthey, F. A. & Simsek, S. Reducing the celiac disease antigenicity of wheat. *Cereal Chem. J.* https://doi.org/10.1094/cchem-05-17-0100-fi (2017).
33. Carpenter, R. P., Lyon, D. H. & Hasdell, T. A. *Guidelines for Sensory Analysis in Food Product Development and Quality Control* (Springer, 2000).
34. Chai, T. & Draxler, R. R. Root mean square error (RMSE) or mean absolute error (MAE)?—Arguments against avoiding RMSE in the literature. *Geosci. Model Dev.* **7**, 1247–1250. https://doi.org/10.5194/gmd-7-1247-2014 (2014).
35. Martín-Domingo, M. C. *et al.* Determination of metalloid, metallic and mineral elements in herbal teas. Risk assessment for the consumers. *J. Food Compos. Anal.* **60**, 81–89. https://doi.org/10.1016/j.jfca.2017.03.009 (2017).
36. Powell, J. J., Burden, T. J. & Thompson, R. P. In vitro mineral availability from digested tea: A rich dietary source of managanese. *Analyst* **123**, 1721–1724. https://doi.org/10.1039/a802131g (1998).
37. Salahinejad, M. & Aflaki, F. Toxic and essential mineral elements content of black tea leaves and their tea infusions consumed in Iran. *Biol. Trace Elem. Res.* **134**, 109–117. https://doi.org/10.1007/s12011-009-8449-z (2010).
38. Basgel, S. & Erdemoglu, S. B. Determination of mineral and trace elements in some medicinal herbs and their infusions consumed in Turkey. *Sci. Total Environ.* **359**, 82–89. https://doi.org/10.1016/j.scitotenv.2005.04.016 (2006).

Author contributions

Y.D.: data curation, formal analysis, methodology, writing-original draft, software. A.D.A.: investigation, funding acquisition, project administration, review and editing. A.F.A.: supervision, conceptualization, review and editing. All authors have read and agreed to the published version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-69149-1.

Correspondence and requests for materials should be addressed to Y.D.