Analytica Chimica Acta 1400 (2026) 345337

Contents lists available at ScienceDirect

Analytica Chimica Acta

journal homepage: www.elsevier.com/locate/aca

Optimized HSSE-GC-MS and machine learning for forensic gender classification of human scent from clothing

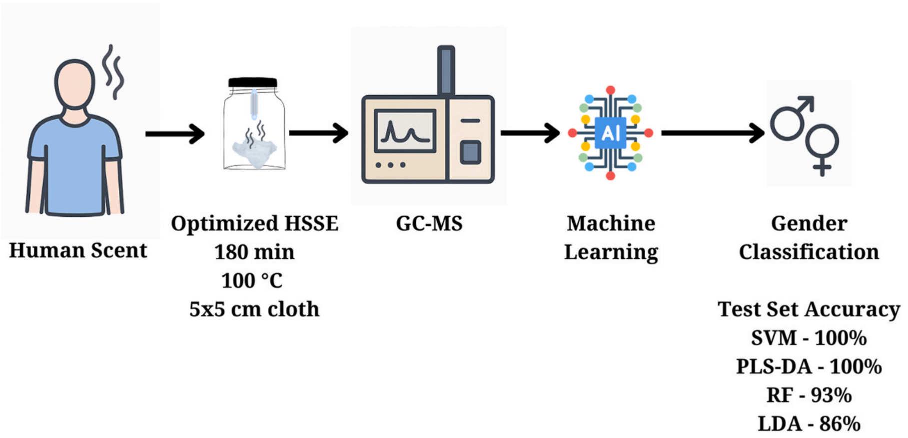
John Marty C. Mateo [a,b], José Luis P. Calle [a], Enrique Durán-Guerrero [a,*], Miguel Palma [a], Marta Ferreiro-González [a]

[a] Department of Analytical Chemistry, Faculty of Sciences, University of Cádiz, Agri-food Campus of International Excellence (ceiA3), IVAGRO, 11510, Puerto Real, Cádiz, Spain
[b] Institute of Legal Medicine, University Hospital Frankfurt am Main, Goethe-University, Kennedyallee 104, D-60596, Frankfurt am Main, Germany

HIGHLIGHTS

- Stir bar HSSE-GC-MS method was optimized for gender classification of human scent.
- TIS data and ML provides fast and automatic gender classification of human scent.
- Supervised learning (SVM, PLS-DA) achieved 100% accuracy for gender classification.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
Human scent
HSSE-GC-MS
Machine learning
VOCs
SVOCs
Forensics

ABSTRACT

Background: Clothes are considered a valuable piece of evidence in forensic investigations because of the underlying information that can be extracted from them. Human scent is one kind of trace evidence that can be extracted from clothes, which can link a specific person or location to the crime. This study optimized a headspace sorptive extraction (HSSE) method coupled with gas chromatography-mass spectrometry in combination with machine learning for objective discrimination of gender from human scent traces on clothing for forensic applications.

Results: The HSSE was optimized using the Box-Behnken design and Response Surface Methodology to maximize the Euclidean distance between the total ion sum spectrum (TIS) of two independent samples. The optimized extraction parameters were determined at 100 °C extraction temperature, 180 min extraction time, and 5 × 5 cm cloth dimension at m/z range 101 – 250 which produced an R^2 of 0.9874. Method validation demonstrated an acceptable absolute error of 6.60%. Intra- and inter-day precisions were calculated for each m/z from 101 to 250

---

* Corresponding author.
E-mail address: enrique.duranguerrero@uca.es (E. Durán-Guerrero).

https://doi.org/10.1016/j.aca.2026.345337
Received 26 November 2025; Received in revised form 27 February 2026; Accepted 2 March 2026

to remove outliers from the dataset for the multivariate analysis. Using the optimized extraction method, 62 samples were analyzed and the pre-processed TIS from the samples were subjected to various supervised machine learning algorithms. Support vector machine (SVM) and partial least squares-discriminant analysis (PLS-DA) models achieved perfect 100% accuracy for gender classification in both training and test sets. Random forest (RF) and linear discriminant analysis (LDA) also showed high test set accuracies of 93% and 86%, respectively.
Significance: This integrated approach, combining optimized stir bar HSSE-GC-MS with different machine learning algorithms using the TIS untargeted approach, provides a fast and automatic gender classification from human scent traces on clothes that can be used as a rapid screening technique in forensic investigations to narrow down suspect or victim profiles.

1. Introduction

Forensic investigations mostly rely on physical evidence to connect individuals or locations to crimes, exonerate innocent parties, and reconstruct events [1,2]. Among the various types of evidence found at crime scenes, human clothing is particularly valuable due to its capacity to retain a diverse array of information, including physical, biological, and chemical traces [3,4]. Human scent is a critical form of trace evidence that can be extracted from clothing, which can be analyzed and offer potential to link specific persons to a crime [5]. It is a complex chemical signature composed of volatile (VOCs) and semi-volatile organic compounds (SVOCs) emitted primarily from human skin, and influenced by interactions with dead skin cells, air currents, and symbiotic bacteria [6]. The human scent profile is considered highly individualistic, transferable, and collectable, with the potential for computerized database development, which are characteristics of an ideal trace evidence [5].

Historically, canines trained for human scent detection have been employed in forensic investigations for their speed, versatility, and discriminatory ability in identifying individual scents [7]. However, the reliability of canine-based methods is inherently dependent on dog training and the consistency of sample handling and preservation [7]. The admissibility of scent evidence by canines in court requires adherence to validated scientific procedures and reliable dog handler methods [8]. Advancements in analytical chemistry have now enabled the identification and quantification of human scent chemical profiles. Gas chromatography-mass spectrometry (GC-MS) is the most common analytical technique used for the analysis, identification, and profiling of VOCs from human scent [8–10]. Other analytical techniques employed include ion mobility spectrometry (IMS) [11], proton transfer reaction-mass spectrometry (PTR-MS) [12], and electronic noses [13,14], which detect gases through pattern recognition systems. Two-dimensional GC (GCxGC-MS) has also gained popularity in recent years for its increased peak capacity, enhanced sensitivity, and potential for group analysis in profiling semi-volatile compounds from human scent [15–18].

Despite these advancements, one of the major challenges in collecting and chemical profiling human scent is the absence of optimized or standardized protocols for the collection and analysis of volatile and semi-volatile organic compounds, especially for forensic purposes [8]. Solid phase microextraction (SPME) is a widely preferred method for human scent extraction due to its simplicity, efficiency, and solvent-free nature, allowing for single-step extraction and concentration using various sorbent materials like polydimethylsiloxane (PDMS), polyacrylate (PA), carboxen/PDMS (CAR/PDMS), and divinylbenzene/CAR/PDMS (DVB/CAR/PDMS) [19–27]. However, existing SPME methods specifically for forensic purposes often involve long preparation times, like 24-h aging of samples to simulate conditions for canine identification [28], followed by 15-21 h of SPME extraction at room temperature before GC-MS analysis [19–22,24,25,29,30]. The total analysis time to obtain just the chromatographic and spectroscopic data can go up to 48 h. Stir bar sorptive extraction (SBSE), which uses polydimethylsiloxane (PDMS) stir bars, can be used as an alternative to SPME. SBSE offers higher sorptive capacity and enhanced sensitivity compared to SPME due to its larger surface area for extracting VOCs and SVOCs from fabric matrices in the headspace [31]. Although SBSE has been used only for direct contact sampling from skin in previous studies [16,32], its application in headspace extraction (Head Space Sorptive Extraction, HSSE) from fabrics for human scent analysis is also possible.

Previous works on objective gender discrimination of human scent involve a targeted approach that relies on the identification of specific marker compounds or chromatographic peaks in tandem with different machine learning techniques. Penn et al. [32] successfully separated the gender from human scent samples using direct SBSE-GC-MS with principal component analysis (PCA). Pojmanova et al. [33] used palm VOCs and headspace-GC-MS with linear support vector machine (SVM) and quadratic discriminant analysis (QDA), while Sabilla et al. [34] used sweat odor and electronic nose with linear SVM, to discriminate gender from human scent. The study of Colon-Crespo et al. [25] and Frazier et al. [22] utilized hand scent and SPME-GC-MS analysis in combination with linear discriminant analysis (LDA) to generate a predictive model for gender. Moreover, Dixon et al. [35] had a similar study using hand scent and SPME-GC-MS analysis to classify gender using partial least squares-discriminant analysis (PLS-DA). While most of these studies have high accuracies in gender discrimination of human scent, this particular approach is time-consuming, especially in the identification and quantification of target compounds or peaks, requires extensive expertise in data analysis, and may overlook subtle differences in concentrations and intensities or be complicated by masking compounds [36].

The present study introduces an “MS-sensor” approach that utilizes the Total Ion Spectrum (TIS) as a holistic chemical fingerprint, harnessing the superior sensitivity, precision, and reproducibility of GC-MS instrumentation. The untargeted TIS-based strategy offers rapid, objective, and data-driven profiling, already validated in applications such as ignitable liquid residue classification [37]. It is important to clarify that, unlike traditional e-nose systems, the key distinction in our approach arises from the ionization step in the mass spectrometer. Specifically, the compounds do not reach the mass spectrometer simultaneously, which provides an additional dimension of analytical information and results in high-quality spectroscopic data. The incorporation of chromatographic separation, even when not fully exploited for individual compound identification, could contribute positively to the overall performance of the method. This partial separation reduces signal overlap and enhances the quality of the spectral patterns used for classification, thereby improving the robustness and reliability of the final model.

In this context, the present study seeks to optimize and validate a human scent extraction method from clothing using stir bar HSSE-GC-MS, and to integrate TIS-derived chemical fingerprints with machine learning algorithms—support vector machines (SVM), random forest (RF), linear discriminant analysis (LDA), and partial least squares discriminant analysis (PLS-DA)—to objectively discriminate gender from human scent. This non-target global profiling significantly reduces analysis time, minimizes reliance on complex spectral libraries and standards, and decreases the need for specialized expertise, while maintaining the efficiency of narrowing suspect profiles within 3–4 h before confirmatory methods are employed.

2. Materials and methods

2.1. Sample collection

The lower back was selected as the sampling site because of its direct contact with clothing and fewer contaminations from exogenous odors like deodorants or perfumes. Before scent collection, the volunteers were advised to clean their lower back with unscented olive oil soap followed by thorough washing with water and drying with paper to avoid any exogenous compounds in the scent profile. The volunteers were also asked not to use deodorants, lotions, perfumes, and other fragrant or cosmetic products 48 h before sample collection. The human scent samples were collected from a total of 62 volunteers. A 5 × 5 cm white cotton and polyester cloth was patched at the lower back part of the individual using medical tape. This material was selected based on the bibliography [6]. Volunteers were asked to do a light physical exercise (walking) for 20 min to induce sweating and activate the sweat glands to produce more VOCs [38]. The clothes with the scent samples were then collected after the exercise and stored in a 10 mL screwcap headspace glass vial, (Agilent CrossLab, Santa Clara, CA, USA) which was then stored in a freezer (-20 °C) prior to extraction and analysis [21].

The volunteers, completely anonymous, agreed to fill out consent forms and questionnaires, answering questions regarding their biological sex, age, race, and specific diet restrictions.

2.2. Experimental design

Box-Behnken design (BBD) and response surface methodology (RSM) were used to optimize the extraction parameters of human scent from clothes. The optimization aimed to maximize the Euclidean distance (ED) between the TIS of two samples (one male, one female) to select the working conditions that allow for maximizing the discriminating power between the two genders. The independent variables optimized were extraction temperature, extraction time, and cloth dimension. The Euclidean distance (ED) calculated from the normalized TIS of the two samples was used as the response variable. Three levels were used for each independent variable which yielded thirteen experiments including one center point (Table 1). Different mass ranges including the initial full scan from m/z 50 – 350 were evaluated for maximizing Euclidean distance between two individual scents. The predictive model generated from the TIS mass range which yielded the highest coefficient of determination was selected.

A second-order polynomial function was built based on the RSM according to the following equation:

$$Y = B_0 + \sum_{i=1}^{n} B_i X_i + \sum_{i=1}^{n} B_{ii} X_i^2 + \sum_{i=1}^{n} \sum_{j=i+1}^{n} B_{ij} X_i X_j$$

Where Y is the response variable; B_0, B_i, B_{ii}, and B_{ij} are the regression coefficients for the intercept, linear, quadratic, and interaction terms, respectively; X_i and X_j are the independent variables.

2.3. Sample extraction

The optimized independent variables were used to extract the VOCs and SVOCs from the clothes. In summary, a GERSTEL twister® (Gerstel, Mülheim a/d Ruhr, Germany) stir bar sorptive extraction device (PDMS, film thickness 0.5 mm, length 20 mm) was placed in the headspace of a 10-mL screwcap glass vial (Agilent CrossLab) containing the 5 × 5 cm clothes with the scent sample. HSSE was carried out for 180 min at 100 °C in a Memmert UNE 400 convection oven (Schwabach, Germany).

Table 1
Levels of independent variables used for the experimental design.

Independent Variable	−1	0	+1
Extraction Temperature (°C)	40	70	100
Extraction Time (min)	30	105	180
Dimension (cm^2)	9 (3 x 3)	16 (4 x 4)	25 (5 x 5)

2.4. Gas-chromatography mass spectrometry (GC-MS) analysis

Agilent Technology 7890 Gas Chromatography in tandem with Agilent 5975C Mass Spectrometry Detector (Agilent) was used to determine the volatile compounds extracted from the SBSE stir bar. The column used was a 60 m × 0.25 mm x 0.25 μm film thickness DB – WAX (Agilent). The VOCs and SVOCs were desorbed from the stir bar in the injection port of the GC with an inlet temperature of 250 °C. The method started with an initial oven temperature of 40 °C for 5 min, followed by a ramp of 10 °C/min until the oven temperature reached 300 °C and ended with a 2 min hold. The helium carrier gas was controlled at a flow rate of 1.0 mL/min. Electron-ionization was employed in the quadrupole mass analyzer in full scan mode over a mass range of 50 – 350 m/z [20]. The TIS was calculated by summing the abundances across the retention time at each m/z value and normalizing the TIS to a maximum of one [36], after which it was used for experimental design experiments and chemometric analysis. TIS provides a global profile instead of resolving specific compounds.

2.5. Data analysis and software

Experimental design was carried out using Minitab® 21.4.1 (Minitab LLC, State College, Pennsylvania, USA). Analysis of variance (ANOVA) was used to determine the statistical significance of variables and interactions with respect to the model ($p < 0.05$). The coefficient of determination, R^2, was used to determine the fitness of the polynomial function. Supervised machine learning algorithms (SVM, RF, LDA, PLS-DA) were done using *caret* and *mixomics* packages of R Studio 2024.04.1 (R version 4.4.1, Boston, Massachusetts, USA). All datasets were divided into training (75%) and test sets (25%) and the generated models from hyperparameter optimization were tested using 5-fold cross-validation. This 5-fold cross-validation step ensures robustness of the model and avoids data overfitting. The training set was utilized to develop the models, while the independent test set was used to assess the predictive model performance. Both sets were randomly and evenly selected to ensure the representation of samples from the different classes.

The performance of the predictive models generated from all the supervised machine learning methods was assessed using the training and test set accuracy, sensitivity, and specificity calculated based on the following equations (Eq. (1), Eq. (2) and Eq. (3)) where TP, FP, TN, and FN represent True Positive, False Positive, True Negative, and False Negative, respectively. The 5-fold cross-validated accuracy, defined as the average accuracy obtained by training and testing the model across five different splits of the dataset, was also reported for each model.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \qquad \text{(Eq. 1)}$$

$$Sensitivity = \frac{TP}{TP + FN} \qquad \text{(Eq. 2)}$$

$$Specificity = \frac{TN}{TN + FP} \qquad \text{(Eq. 3)}$$

3. Results and discussion

3.1. Method optimization for the extraction of human scent

BBD and RSM were performed to optimize the extraction parameters (extraction time, extraction temperature, and cloth dimension) for discriminating genders using human scent (Table 1). Different mass ranges were used to generate predictive models starting from the initial full scan of m/z 50 – 350, however the predictive model generated from the TIS of mass range of m/z 101 – 250 yielded the highest coefficient of determination ($R^2 = 0.9874$) and adjusted R^2 (0.9495), indicating an excellent fit of the experimental data to the model and suitability of the experimental design for maximizing distances between two individual scents (Table 2).

Analysis of variance (ANOVA) was used to assess the effect of the extraction variables on the response and the results are shown in Table 3. Results showed that extraction time and temperature have a significant effect on the response, while cloth dimension was non-significant. The interaction between temperature-dimension and the square function of time also significantly affected the response ($p < 0.05$). The generated model was significant with p-value = 0.011 and was able to explain 98.74% of the response variance.

A standardized Pareto chart, shown in Fig. 1, illustrates the effect of the variables and their interactions in the generated model. Extraction temperature, the interaction between extraction temperature and dimension, and the square function of time had significant positive effects (green) on the response, while extraction time had a significant negative effect (yellow) on the response. The significant positive effect of the variables and their interactions maximizes the differences between two different samples (male and female) and therefore favors the discrimination between two individuals.

The model derived from the BBD and RSM is characterized by the polynomial function shown below in Eq. (4).

$$\begin{aligned}\boldsymbol{ED} = & -0.123 + (0.004{*}\mathit{Temperature}) - (0.001227{*}\mathit{Time}) \\ & + (0.0838{*}\mathit{Dimension}) - \left(0.00003{*}\mathit{Temperature}^2\right) \\ & + \left(0.000008{*}\mathit{Time}^2\right) - \left(0.0164{*}\mathit{Dimension}^2\right) \\ & + (0.00001{*}\mathit{Temperature}{*}\mathit{Time}) \\ & + (0.001013{*}\mathit{Temperature}{*}\mathit{Dimension}) \\ & - (0.000059{*}\mathit{Time}{*}\mathit{Dimension})\end{aligned} \tag{Eq. 4}$$

Direct proportionality of extraction temperature with the ED was observed in Eq. (4). High extraction temperatures are known to increase the volatility of VOCs and increase their concentration in the headspace while also improving the mass transfer from the sample matrix to the headspace [39]. High extraction temperature is important in solid matrices where VOCs are harder to release [40]. Increasing the temperature to 100 °C will not only extract VOCs but also SVOCs which were recently discovered to play an important role in the formation of human scent, with recent studies trying to include SVOCs in the sampling and analysis of overall human scent [8].

It is also observed that time has a significant negative effect on the response, however, this effect is countered by the significant positive effect of the square function of time. Therefore, a non-linear relationship was found for time. This can be because VOCs are readily released at low extraction temperatures and therefore only need a little time to equilibrate. However, at higher extraction temperatures and times, trapped VOCs in the matrix were more easily released and the SVOCs were also extracted in larger amount. This may also imply that SVOCs play a significant role in increasing the differences between the two samples which can be beneficial in classification studies. The optimized extraction time was significantly shorter compared to previous studies of Frazier et al. [22], Curran et al. [20], and Prada et al. [19] which implemented 15 h, 21 h, and 21 h extraction times using SPME at room temperature, respectively.

Table 2
Performance of the models derived from different mass ranges.

Mass Range (m/z)	Performance Parameter	
	R^2	R^2 adjusted
50 – 350	0.7161	0.00
50 – 250	0.7750	0.00
50 – 200	0.7228	0.00
101 – 300	0.9413	0.7654
151 – 300	0.9129	0.6516
201 - 300	0.9804	0.6337
101 – 250	0.9874	0.9495
75 – 250	0.9394	0.7576

Table 3
Results of ANOVA of the model generated using Box-Behnken experimental design.

Source	F-Value	p-Value
Regression Model	26.09	0.011
Temperature	45.46	0.007
Time	117.38	0.002
Dimension	8.50	0.062
Temperature*Temperature	6.13	0.090
Time*Time	15.59	0.029
Dimension*Dimension	2.28	0.228
Temperature*Time	7.49	0.071
Temperature*Dimension	13.64	0.034
Time*Dimension	0.29	0.626
$R^2 = 0.9874$, adj. $R^2 = 0.9495$		

Using Eq. (4), the predicted response values were calculated. Table 4 summarizes the BBD matrix with the experimental and predicted responses. The difference between the actual and predicted models was small which demonstrates the good performance of the generated model.

Surface response plots (Fig. 2) were also generated from Eq. (4) to show the effect of the main variables on the response. Based on the surface plots, the highest response is achieved with high extraction time, high extraction temperature, and high cloth dimension. However, the surface plots suggest that increasing the extraction temperature will not be beneficial because of the minimal positive effect on the response. The optimization space was constrained based on both mathematical modeling and physical limitations: according to Eq. (4), the negative quadratic coefficients for temperature and sample dimension indicate that exceeding current limits would result in a decline in ED due to the concave nature of the response surface (Fig. 2). Physically, preliminary experiments revealed that temperatures above 100 °C and an increase in extraction time (e.g., 5 h) caused sample degradation (burning), which introduces significant noise and compromises signal integrity. Furthermore, statistical analysis (Table 3) confirmed that increasing sample dimension does not significantly improve ED and leads to poorer sample separation.

A validation experiment was carried out applying the optimal conditions (100 °C extraction temperature, 180 min extraction time, and 5 × 5 cm clothes dimension) yielded an absolute error of 6.60% from the predicted response, which is well within acceptable limits for prediction models (<10%).

3.2. Repeatability and intermediate precision of the method

The intra-day precision, expressed as relative standard deviation (RSD) of the summed TIS intensities of the mass spectral range, m/z 101 – 250, for one individual, was 9.75% while the inter-day precision for three days ranged from 5.61 to 12.23%. The intermediate precision for inter-day analysis (RSD_{pooled}) was calculated to be 8.76 % (Table 5). These RSD values are considered acceptable for human scent analysis using the commonly accepted threshold of 10% [41], especially given the multi-step extraction and analysis method which propagates variation and error.

Subsequently, m/z outliers were identified and excluded from the dataset because they exhibited high variability and unstable TIS intensities in both inter-day and intra-day measurements. This filtering strategy was based on the forensic odor framework, in which the endogenous and genetically influenced primary human odor are known to be generally more stable over time than secondary and tertiary odor components, which are strongly affected by diet, environmental

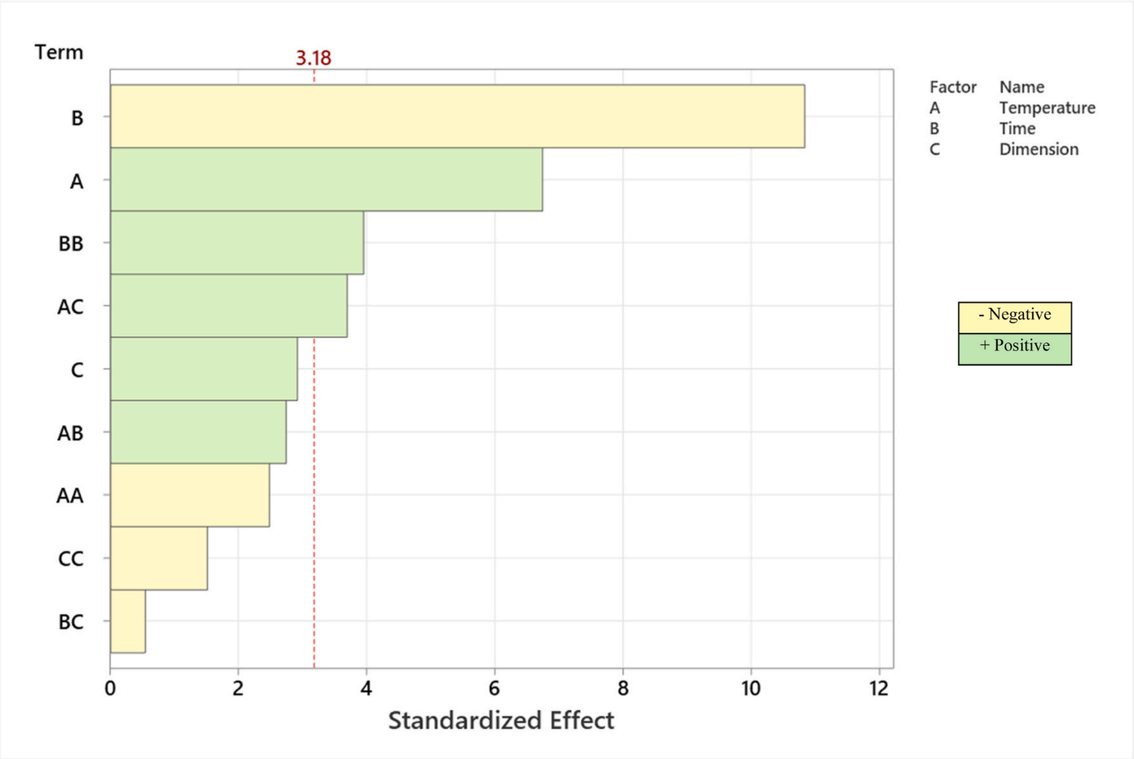


Fig. 1. Standardized Pareto chart of the generated prediction model.

Table 4
BBD matrix with actual and predicted response.

Experiment	Temperature	Time	Dimension	Response (ED)	
	°C	min	cm^2	Actual	Predicted
1	70	105	16	0.1174	0.1225
2	70	30	9	0.0702	0.0606
3	70	30	25	0.0937	0.1041
4	70	180	9	0.2031	0.2070
5	70	180	25	0.2088	0.2328
6	40	105	9	0.0360	0.0531
7	40	105	25	0.0285	0.0270
8	100	105	9	0.0588	0.0705
9	100	105	25	0.1729	0.1659
10	40	30	16	0.0553	0.0552
11	40	180	16	0.1383	0.1477
12	100	30	16	0.0835	0.0883
13	100	180	16	0.2566	0.2709

exposure, and cosmetic products [20]. Because primary odor is considered the principal basis for individual scent discrimination in forensic applications [25,32], removing highly variable *m/z* features is expected to reduce contributions from environmentally derived and exogenous chemicals and potentially enrich primary odor. In total, seventeen (17) *m/z* features were excluded, and 133 *m/z* features were retained for subsequent analysis.

3.3. Machine learning-based classification by gender

After the method was optimized, the applicability of the optimized method in objectively discriminating gender of human scent from clothes using different algorithms was tested. The visual inspection of the chromatograms did not reveal any clear differentiation by gender (Supplementary Material, Fig. S1). Therefore, four supervised machine learning techniques were studied including parametric, like LDA and PLS-DA, and non-parametric algorithms like SVM and RF. A total of 62 human scent samples (40 females, 22 males) were collected from volunteers using cotton as an adsorbent and 133 different *m/z* variables were used to generate predictive models, D_{62x133}.

3.3.1. Support vector machine (SVM)

SVM was employed on the data matrix D_{62x133} by utilizing Gaussian Kernel. SVM is a non-parametric algorithm that finds an ideal hyperplane where the distance between different classes is maximized, therefore making the classification effective, especially in high-dimensional spaces because of the use of different kernel functions [42]. A Gaussian SVM, which utilizes a radial basis function kernel (RBF), was used in this study because of its ability to model complex and non-linear relationship boundaries by mapping data to a higher-dimensional space [43] which is better than a linear SVM because of the wider applicability to non-linear datasets.

SVMs are characterized by the Cost (C) and Kernel (γ) hyperparameters which were optimized and tuned through Grid search method described previously [44] to generate the most accurate prediction models. Heatmaps were generated to visualize the calculated accuracies from different combinations of the hyperparameters, γ and C, using 5-fold cross-validation (Fig. 3). The logarithmic function, $\log2\gamma$ and $\log_2 C$, were plotted on the y- and x-axes and the calculated accuracies were represented in a color scale, where the accuracy increases from blue to red. The best accuracies were achieved at high positive $\log_2 C$ ($C = 32$) and high negative $\log2\gamma$ ($\gamma = 0.000977$) with 32 support vectors.

The cost parameter, C, balances the trade-off between low classification errors while maintaining a high margin that separates the different classes [42]. The γ parameter, on the other hand, tells how singular training examples can influence the decision boundary [44]. The combination of high C and low γ observed in the optimized hyperparameters makes the model less tolerant to misclassifications with a small influence of a single training example on the decision boundary.

Using the optimized hyperparameters, the generated predictive

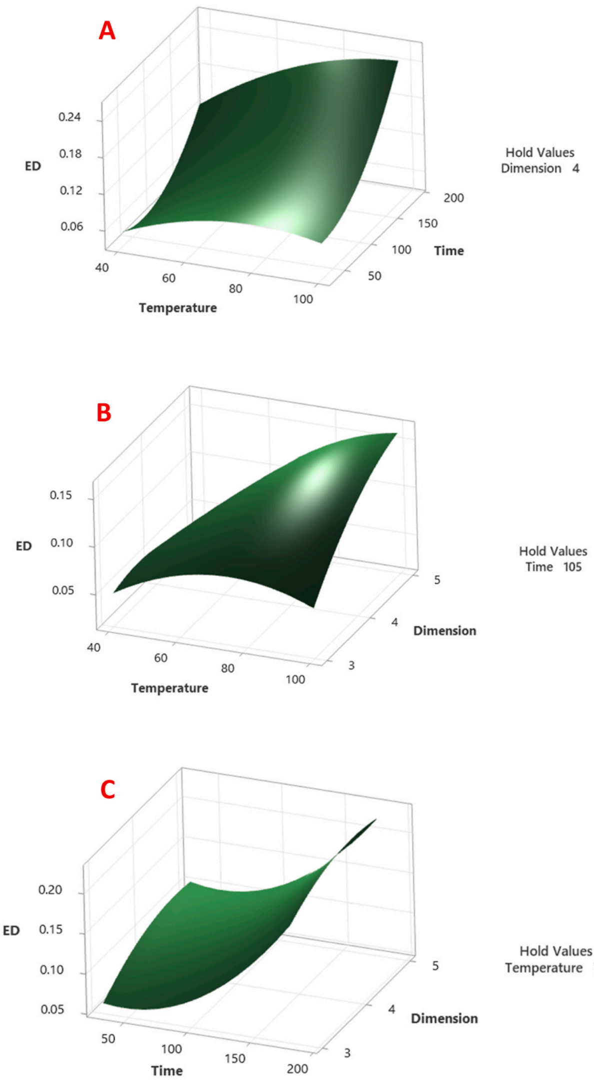


Fig. 2. Response surface plots as a function of (A) extraction time and extraction temperature, (B) extraction temperature and dimension of clothes, and (C) extraction time and dimension of clothes.

Table 5
Intra- and inter-day precision of the analytical method for the extraction and analysis of VOCs in human back samples.

Precision		% RSD	% RSD_{pooled}
Intra-day	($n = 5$)	9.75	N/A
Inter-day	Day 1 ($n = 4$)	7.01	8.76
	Day 2 ($n = 4$)	5.61	
	Day 3 ($n = 4$)	12.23	

model using Gaussian SVM produced 100% training and test accuracy in objectively discriminating genders using human scent from cloth while the cross-validated accuracy falls to 87% (Table 6). This high cross-validated accuracy of gender classification using SVM is similar to other studies employing different methodologies and data processing approaches. In the study of Pojmanova et al. [33], involving chromatographic data to differentiate sex using palm VOCs, linear SVM produced an accuracy of 89.86% with a critical sensitivity of 88.89%. In another study, Sabilla et al. [34] used sweat odor and electronic nose to discriminate gender using linear SVM and achieved an accuracy of 94.12%.

3.3.2. *Random forest (RF)*

Another non-parametric supervised machine learning method, RF, was tested to classify gender using human scent from clothes. Random forest algorithm connects and merges simple decision trees to generate a model with improved accuracy and reliability [45]. Two important parameters, the number of trees (*ntree*) and the number of variables to randomly sample as candidates at each split (*mtry*), were optimized to achieve the best accuracy for the generated predictive model. The maximum and stable accuracy was achieved at 80 trees from a *ntree* vs accuracy scan (0 – 200 trees) while the value of *mtry,* which usually is the square root of the total number of predictors [46], was set to 12 predictors.

The generated RF prediction models registered a test set accuracy of 93% only misclassifying one sample while the cross-validated accuracy is calculated to be 85% (Table 6). The RF generated model produced 100% accuracy in the training set which means that the merged and connected decision trees were able to accurately differentiate the classes given the optimal number of predictors and number of trees. On the other hand, the test set yields 100% sensitivity and 80% specificity (1 misclassification).

3.3.3. *Linear discriminant analysis (LDA)*

LDA, a parametric algorithm, was also evaluated to classify gender using human scent extracted from clothes. LDA is a statistical technique that projects data into a lower-dimensional space to minimize the variance within one class and maximize the variance between classes and the separator between classes are called linear discriminants. These linear discriminants are particularly helpful for pattern recognition and machine learning applications [47,48].

In this particular dataset, only one linear discriminant (LD1) was generated which clearly separated the two classes in the training set (Fig. 4). The generated predictive model from LDA produced a cross-validated accuracy of 83%, while the test set had 90% (9/10) sensitivity and 80% specificity (4/5), respectively (Table 6).

Several studies generated LDA models from human scent to discriminate gender. The study of Colon-Crespo et al. [25] identified 15 VOC marker compounds from hands and used LDA to generate a predictive model for gender. The validated LDA model generated from this study yields 71% highest prediction accuracy. Another study from Frazier et al. [22] using hand scent employed a relatively similar approach in generating LDA models for gender classification by utilizing the total ion chromatogram (TIC) and implementing supervised reduction analysis of the peaks prior to machine learning. The generated LDA model for gender classification had 96.67% accuracy using leave-one-out cross-validation (LOOCV) which is close to the achieved accuracy of the LDA model from this study because of the similarity of the data processing approach employed. On the other hand, a quadratic discriminant analysis was able to achieve an accuracy of 86.96% in gender classification using the peak selection method and validated through LOOCV [33].

3.3.4. *Partial least squares – discriminant analysis (PLS-DA)*

The last supervised machine learning used was the parametric PLS-DA which is identical to the partial least squares regression statistical approach but shifted the focus to pattern recognition rather than a regression model. It is usually utilized when the variables are highly collinear and if the variables are greater than the sample size which is usually the case for chemometric analysis [49].

The optimized PLS-DA model used five components and achieved 100% test set classification accuracy for gender classification while the cross-validated accuracy was calculated to be 98% (Table 6). A similar pattern recognition study was carried out to classify gender using PLS-DA based on VOCs from human sweat. Prior to model development, peak selection approach was used, and the processed data were then subjected to PLS-DA to build a gender classification model. The auto-prediction method obtained an accuracy of 97% [35].

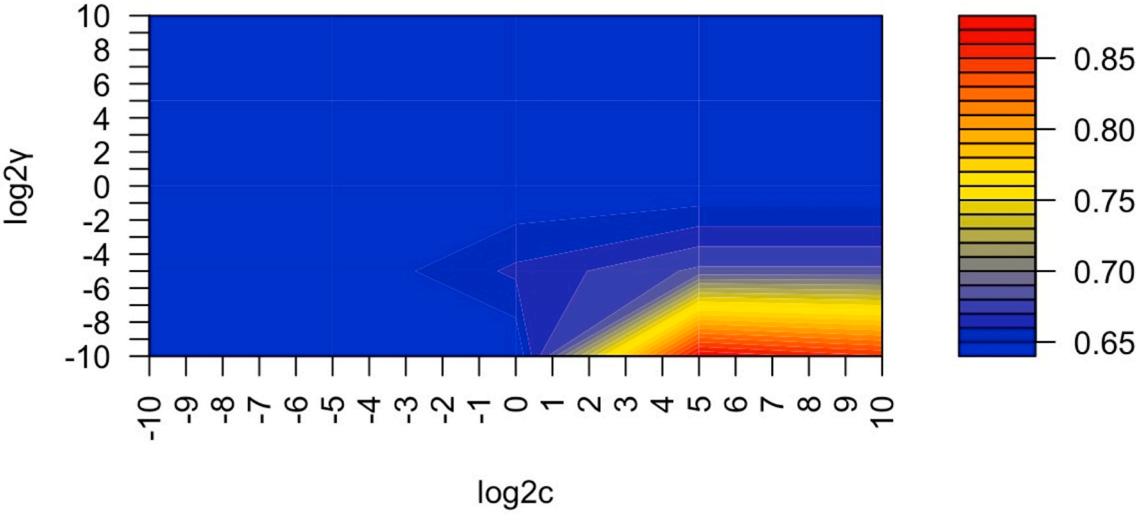


Fig. 3. Heat map generated from the hyperparameters grid search method for the Gaussian SVM model obtained from 5-fold cross-validation.

Table 6
Summary of the performance of all predictive models generated from different supervised learning algorithms.

Model	Training Set, %		Test Set, %		Cross-Validated Accuracy, %
	Sensitivity	Specificity	Sensitivity	Specificity	
SVM	100	100	100	100	87
RF	100	100	100	75	85
LDA	100	100	90	80	83
PLS-DA	100	100	100	100	98

3.4. Chemical basis for TIS-based gender discrimination models and limitations

Previous studies have already established that human scent contains gender-related differences in chemical profile which provides a strong biochemical foundation for discriminative scent modeling. Fatty acids, aldehydes, ketones, alcohols, esters, and hydrocarbons are compound classes that were identified to differentiate gender in human scent [22,25,32,33,50]. In particular, Penn et al. [32] identified 12 gender-characteristic compounds that differ in relative abundance between male and female axillary sweat profiles. Recent work using HS-SPME-GC-MS of hand odor profiles has also demonstrated high-accuracy gender prediction based on VOC signatures [22]. In addition, Pojmanova et al. [33] revealed that gender-related differences in scent profile arise from variation in VOC levels, rather than specific compound signatures, and used this approach to generate a highly accurate gender prediction model.

Building on these established findings, this study explored a simplified, global chemical fingerprint-based analytical strategy. Rather than focusing on identification and quantification of individual compounds, the TIS approach was used, in which the full mass spectral ion intensity distribution is treated as the global chemical fingerprint. In this framework, each *m/z* value functions like a sensor (variable) which captures the scent composition in the form of mass fragments, enabling rapid discrimination without the need for time-consuming compound annotation and quantification [36]. These prior studies support the

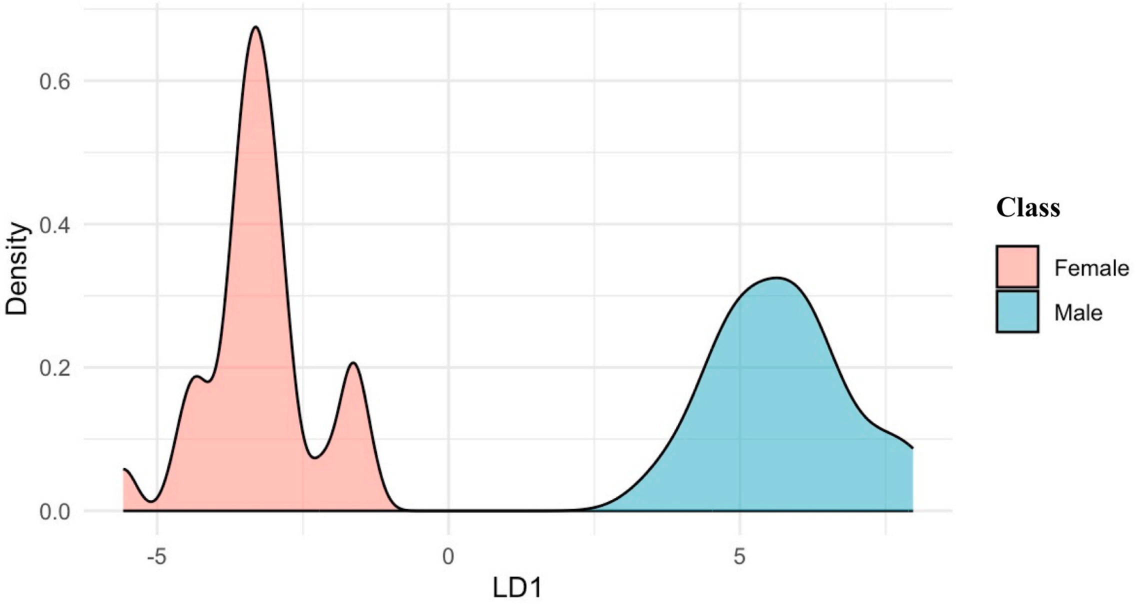


Fig. 4. Separation of classes in the training set using the identified linear discriminant (LD1).

plausibility of gender discrimination using global chemical fingerprint-based TIS models, because the mass fragments in the TIS came from the same compounds present in human scent and therefore reflect their relative abundances. This suggests that the reported high accuracy in gender classification is likely driven by gender-related chemical differences in the scent profile.

The discriminatory information exploited by the TIS-based models therefore likely reflects the cumulative contributions of fragment ions from gender-dependent chemical functional groups or classes, rather than reliance on a small number of specific compounds. As a global chemical fingerprint, TIS also captures fragments from compounds that are not identified or quantified in compound selection models providing a more holistic approach than those used in previous studies. The results obtained from this study complements prior compound-selection studies, supporting that the TIS approach likewise captures gender-related chemical variation but in the context of a global chemical fingerprint.

One of the limitations of this study is the limited sample size and unequal representation of both genders, which may not fully capture the biological variability, particularly within the male population (22/62). Although volunteers represented different regions, age, and dietary habits (Supplementary Material, Table S1), scent profiles are also affected by medication use, hygiene products, detergents, smoking, and environmental exposure. Given a limited number of male sample, the model may partially learn cohort-specific characteristics rather than sex-related biochemical patterns which can increase the risk of overfitting. As a result, the applicability domain of the model should be restricted to individuals with demographic and lifestyle characteristics similar to this cohort, and applying it to broader populations may lead to increased misclassification (false positive) due to unrepresented diets, habits or product use. Future work would expand the population and diversity to develop more robust models. It is important to note that the models learn from new samples and are constantly being updated.

Given the high variable to sample ratio (133/62), the probability of chance correlation was assessed. A permutation test with 1000 iterations was performed in the two models (PLS-DA, SVM) that achieved 100% test and training classification accuracy. Both models had an empirical p-value ≤ 0.001, which means that none of the permutated iterations had an accuracy equal or more than the observed accuracy (Supplementary Material, Fig. S2). The result demonstrates that the high accuracy of these models is highly unlikely to arise from random associations or chance and further support the robustness of the TIS-based modeling approach.

4. Conclusion

This study successfully optimized a stir bar HSSE-GC-MS method for extracting volatile and semi-volatile compounds from clothes specifically used for objective gender discrimination using a global profiling, non-targeted TIS approach and different machine learning algorithms. The optimal extraction parameters, 100 °C for 180 min using a 5×5 cm cloth utilizing the mass spectral range of m/z 101 – 250, were determined using BBD and RSM. This extraction parameters maximized the Euclidean distance between male and female scents and produced a predictive model with a good fit in the experimental data ($R^2 = 0.9874$). Method validation studies results to a low absolute error of 6.60% while the precision values were within the threshold value. The utilization of TIS data in combination with different machine learning algorithms produced highly accurate gender predictive models, resulting to 100% test accuracy using SVM and PLS-DA models while RF (93%) and LDA (86%) also showed high accuracies. Based on these results, future work may build on these findings by expanding the sample size and increasing the population diversity, including varying habits, to further reinforce and generalize the observed patterns in this study.

This validated and optimized HSSE-GC-MS method and utilization of TIS data coupled with different machine learning algorithm gives a fast, automatic and objective method to classify gender using human scent extracted from clothes, which can significantly aid forensic investigations. This method can serve as a rapid screening tool fot gender identification with results obtainable within 3 to 4 h, helping to narrow down suspects or identify victims while more time-consuming techniques are being processed.

CRediT authorship contribution statement

John Marty C. Mateo: Writing – original draft, Validation, Software, Investigation, Formal analysis, Data curation, Conceptualization. **José Luis P. Calle:** Writing – review & editing, Software. **Enrique Durán-Guerrero:** Writing – review & editing, Supervision, Project administration, Methodology, Investigation, Conceptualization. **Miguel Palma:** Resources, Project administration, Methodology, Conceptualization. **Marta Ferreiro-González:** Writing – review & editing, Supervision, Project administration, Methodology, Investigation, Conceptualization.

Funding sources

This study was funded by the European Union through the Erasmus Mundus Master in Quality in Analytical Laboratories (EMQAL).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to thank the Education, Audiovisual, and Culture Executive Agency (EACEA) of the European Union for the financial support of John Marty C. Mateo for his master's study through the Erasmus Mundus Master in Quality in Analytical Laboratories (EMQAL). The authors are equally grateful to the Instituto de Investigación Vitivinícola y Agroalimentario (IVAGRO) and the Department of Analytical Chemistry of the University of Cadiz for providing the necessary facilities to execute this research.

Glossary

BBD	Box-Behnken Design
ED	Euclidean Distance
GC-MS	Gas Chromatography–Mass Spectrometry
HSSE	Headspace Sorptive Extraction
LDA	Linear Discriminant Analysis
MDA	Mean Decrease Accuracy
PLS-DA	Partial Least Squares–Discriminant Analysis
RF	Random Forest
RSM	Response Surface Methodology
SBSE	Stir Bar Sorptive Extraction
SPME	Solid Phase Microextraction
SVOCs	Semi-Volatile Organic Compounds
SVM	Support Vector Machine
TIS	Total Ion Sum Spectrum/Spectra
VIP Scores	Variable Importance Projection (Scores)
VOCs	Volatile Organic Compounds

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.aca.2026.345337.

Data availability

Data will be made available on request.

References

[1] B.A.J. Fisher, Techniques of Crime Scene Investigation, CRC Press, Boca Ratón, 2003, https://doi.org/10.1201/9781420058192.

[2] H.A. Harris, H.C. Lee, Introduction to Forensic Science and Criminalistics, Routledge, New York, 2019, https://doi.org/10.4324/9781315119175.

[3] J.M. Taupin, C. Cwiklik, Scientific Protocols for Forensic Examination of Clothing, CRC Press, Boca Ratón, 2010, https://doi.org/10.1201/b10381.

[4] V. Warrier, T. Kanchan, Clothes and the evidences they carry: a perspective on its forensic examination, J. Indian Acad. Forensic Med. 43 (3) (2021) 296–299, https://doi.org/10.5958/0974-0848.2021.00076.2.

[5] F.A. Procter, G.T. Swindles, N.L.M. Barlow, Examining the transfer of soils to clothing materials: implications for forensic investigations, Forensic Sci. Int. 305 (2019) 110030, https://doi.org/10.1016/j.forsciint.2019.110030.

[6] R.H. McQueen, G.T. Eyres, R.M. Laing, Textile sorption and release of odorous volatile organic compounds from a synthetic sweat solution, Textil. Res. J. (May 2024), https://doi.org/10.1177/00405175241249462.

[7] K.G. Furton, N.I. Caraballo, M.M. Cerreta, H.K. Holness, Advances in the use of odour as forensic evidence through optimizing and standardizing instruments and canines, Philos. Trans. R. Soc. B Biol. Sci. 370 (2015) 20140262, https://doi.org/10.1098/rstb.2014.0262.

[8] R. Peters, R. Veenstra, K. Heutinck, A. Baas, S. Munniks, J. Knotter, Human scent characterization: a review, Forensic Sci. Int. 349 (2023) 111743, https://doi.org/10.1016/j.forsciint.2023.111743.

[9] A. Mitra, S. Choi, P.R. Boshier, et al., The human skin volatolome: a systematic review of untargeted mass spectrometry analysis, Metabolites 12 (9) (2022) 824, https://doi.org/10.3390/metabo12090824.

[10] S.K. Pandey, K.H. Kim, Human body-odor components and their determination, TrAC Trends Anal. Chem. 30 (5) (2011) 784–796, https://doi.org/10.1016/j.trac.2010.12.005.

[11] V. Ruzsanyi, P. Mochalski, A. Schmid, et al., Ion mobility spectrometry for detection of skin volatiles, J. Chromatogr. B. 911 (2012) 84–92, https://doi.org/10.1016/j.jchromb.2012.10.028.

[12] L. Yao, R.M. Laing, P.J. Bremer, P.J. Silcock, M.J. Leus, Measuring textile adsorption of body odor compounds using proton-transfer-reaction mass spectrometry, Textil. Res. J. 85 (17) (2015) 1817–1826, https://doi.org/10.1177/0040517515576325.

[13] Z. Zhao, F. Tian, H. Liao, X. Yin, Y. Liu, B. Yu, A novel spectrum analysis technique for odor sensing in optical electronic nose, Sens. Actuators B Chem. 222 (2016) 769–779, https://doi.org/10.1016/j.snb.2015.08.128.

[14] B. Szulczyński, J. Namieśnik, J. Gębicki, Determination of odour interactions of Three-component gas mixtures using an electronic nose, Sensors 17 (10) (2017) 2380, https://doi.org/10.3390/s17102380.

[15] P. Doležal, P. Kyjaková, I. Valterová, Š. Urban, Qualitative analyses of less-volatile organic molecules from female skin scents by comprehensive two dimensional gas chromatography–time of flight mass spectrometry, J. Chromatogr., A 1505 (2017) 77–86, https://doi.org/10.1016/j.chroma.2017.04.062.

[16] V. Cuzuel, R. Leconte, G. Cognon, et al., Human odor and forensics: towards Bayesian suspect identification using GC × GC–MS characterization of hand odor, J. Chromatogr. B. 1092 (2018) 379–385, https://doi.org/10.1016/j.jchromb.2018.06.018.

[17] P. Pojmanová, N. Ladislavová, Š. Urban, Development of a method for the measurement of human scent samples using comprehensive two-dimensional gas chromatography with mass detection, Separations 8 (12) (2021) 232, https://doi.org/10.3390/separations8120232.

[18] J. Hlavsa, R. Spetlik, J. Čechová, P. Pojmanová, J. Matas, Š. Urban, Sex classification from human scent using image interpretation of 2D gas chromatography-mass spectrometry data, in: Lecture Notes in Computer Science, 2025, pp. 457–470, https://doi.org/10.1007/978-3-031-95911-0_32.

[19] P. Prada, A. Curran, K. Furton, Characteristic human scent compounds trapped on natural and synthetic fabrics as analyzed by SPME-GC/MS, J. Forensic Sci. Criminol. 1 (1) (2014), https://doi.org/10.15744/2348-9804.1.s101.

[20] A.M. Curran, C.F. Ramirez, A.A. Schoon, K.G. Furton, The frequency of occurrence and discriminatory power of compounds found in human scent across a population determined by SPME-GC/MS, J. Chromatogr. B. 846 (1-2) (2007) 86–97, https://doi.org/10.1016/j.jchromb.2006.08.039.

[21] J.S. Brown, P.A. Prada, A.M. Curran, K.G. Furton, Applicability of emanating volatile organic compounds from various forensic specimens for individual differentiation, Forensic Sci. Int. 226 (1-3) (2013) 173–182, https://doi.org/10.1016/j.forsciint.2013.01.008.

[22] C.J.G. Frazier, V.A. Gokool, H.K. Holness, D.K. Mills, K.G. Furton, Multivariate regression modelling for gender prediction using volatile organic compounds from hand odor profiles via HS-SPME-GC-MS, PLoS One 18 (7) (2023) e0286452, https://doi.org/10.1371/journal.pone.0286452.

[23] M. Kusano, E. Mendez, K.G. Furton, Development of headspace SPME method for analysis of volatile organic compounds present in human biological specimens, Anal. Bioanal. Chem. 400 (7) (2011) 1817–1826, https://doi.org/10.1007/s00216-011-4950-2.

[24] V.A. Gokool, H.K. Holness, K.G. Furton, The influence of intra-personal variations in human hand odor on the determination of sample donor, Forensic Sci. Int. 334 (2022) 111235, https://doi.org/10.1016/j.forsciint.2022.111235.

[25] L.J. Colón-Crespo, D. Herrera-Hernández, H. Holness, K.G. Furton, Determination of VOC marker combinations for the classification of individuals by gender and race/ethnicity, Forensic Sci. Int. 270 (2017) 193–199, https://doi.org/10.1016/j.forsciint.2016.09.011.

[26] J. Pawliszyn, Solid Phase Microextraction: Theory and Practice, Elsevier Science B. V., Ámsterdam, 1997. http://ci.nii.ac.jp/ncid/BA32245737.

[27] S. Merkle, K. Kleeberg, J. Fritsche, Recent developments and applications of solid phase microextraction (SPME) in food and environmental analysis—A review, Chromatography 2 (3) (2015) 293–381, https://doi.org/10.3390/chromatography2030293.

[28] A.M. Curran, P.A. Prada, K.G. Furton, Canine human scent identifications with post-blast debris collected from improvised explosive devices, Forensic Sci. Int. 199 (1-3) (2010) 103–108, https://doi.org/10.1016/j.forsciint.2010.03.021.

[29] A.M. Curran, S.I. Rabin, P.A. Prada, K.G. Furton, Comparison of the volatile organic compounds present in human odor using SPME-GC/MS, J. Chem. Ecol. 31 (7) (2005) 1607–1619, https://doi.org/10.1007/s10886-005-5801-4.

[30] A.M. Curran, P.A. Prada, K.G. Furton, The differentiation of the volatile organic signatures of individuals through SPME-GC/MS of characteristic human scent compounds, J. Forensic Sci. 55 (1) (2010) 50–57, https://doi.org/10.1111/j.1556-4029.2009.01236.x.

[31] C. Bicchi, C. Iori, P. Rubiolo, P. Sandra, Headspace sorptive Extraction (HSSE), Stir bar sorptive Extraction (SBSE), and Solid Phase microextraction (SPME) applied to the analysis of roasted arabica coffee and coffee brew, J. Agric. Food Chem. 50 (3) (2002) 449–459, https://doi.org/10.1021/jf010877x.

[32] D.J. Penn, E. Oberzaucher, K. Grammer, et al., Individual and gender fingerprints in human body odour, J. R. Soc. Interface. 4 (13) (2006) 331–340, https://doi.org/10.1098/rsif.2006.0182.

[33] P. Pojmanová, N. Ladislavová, V. Škeříková, J. Kukal, Š. Urban, Sex differentiation from human scent chemical analysis, Separations 10 (5) (2023) 293, https://doi.org/10.3390/separations10050293.

[34] I.A. Sabilla, D.P. Purbawa, R. Sarno, A.A. Fauzi, D.R. Wijaya, R. Gunawan, Men and women classification at night through the armpit sweat Odor using electronic nose, in: IEEE Asia Pacific Conference on Wireless and Mobile, April 2021, https://doi.org/10.1109/apwimob51111.2021.9435273.

[35] S.J. Dixon, Y. Xu, R.G. Brereton, et al., Pattern recognition of gas chromatography mass spectrometry of human volatiles in sweat to distinguish the sex of subjects and determine potential discriminatory marker peaks, Chemometr. Intell. Lab. Syst. 87 (2) (2007) 161–172, https://doi.org/10.1016/j.chemolab.2006.12.004.

[36] M. Barea-Sepúlveda, H. Duarte, M.J. Aliaño-González, A. Romano, B. Medronho, Total ion chromatogram and total ion mass spectrum as alternative tools for detection and discrimination (A review), Chemosensors 10 (11) (2022) 465, https://doi.org/10.3390/chemosensors10110465.

[37] M.E. Sigman, M.R. Williams, J.A. Castelbuono, J.G. Golca, C.D. Clark, Ignitable liquid classification and identification using the summed-ion mass spectrum, Instrum. Sci. Technol. 36 (4) (2008) 375–393, https://doi.org/10.1080/10739140802151440.

[38] M. Gallagher, C.J. Wysocki, J.J. Leyden, A.I. Spielman, X. Sun, G. Preti, Analyses of volatile organic compounds from human skin, Br. J. Dermatol. 159 (4) (2008) 780–791, https://doi.org/10.1111/j.1365-2133.2008.08748.x.

[39] C. Lancioni, C. Castells, R. Candal, M. Tascon, Headspace solid-phase microextraction: fundamentals and recent advances, Adv. Sample Prep. 3 (2022) 100035, https://doi.org/10.1016/j.sampre.2022.100035.

[40] B. Kolb, L.S. Ettre, Static headspace–gas Chromatography: Theory and Practice, Wiley-Interscience, Hoboken, 2006, https://doi.org/10.5860/choice.44-1529.

[41] AOAC, AOAC Peer Verified Methods Program, in: Manual on Policies and Procedures, AOAC International, Arlington, VA, USA, 1998.

[42] C. Cortes, V. Vapnik, Support-vector networks, Mach. Learn. 20 (3) (1995) 273–297, https://doi.org/10.1007/bf00994018.

[43] B. Scholkopf, A.J. Smola, Learning with Kernels, MIT Press, Cambridge, 2018, https://doi.org/10.7551/mitpress/4175.001.0001.

[44] A.A. Misolas, M. Ferreiro-González, M. Palma, Intelligent and automatic characterization of ignitable liquid residues by using total ion spectrum and machine learning, Microchem. J. (2024) 111757, https://doi.org/10.1016/j.microc.2024.111757.

[45] L. Breiman, Random forests, Mach. Learn. 45 (2001) 5–32, https://doi.org/10.1023/a:1010933404324.

[46] J.L.P. Calle, B. Falatová, M.J. Aliaño-González, M. Ferreiro-González, M. Palma, Machine learning approaches over ion mobility spectra for the discrimination of ignitable liquids residues from interfering substrates, Talanta Open 6 (2022) 100125, https://doi.org/10.1016/j.talo.2022.100125.

[47] R.A. Fisher, The use of multiple measurements in taxonomic problems, Ann. Eugen. 7 (2) (1936) 179–188, https://doi.org/10.1111/j.1469-1809.1936.tb02137.x.

[48] G.J. McLachlan, Discriminant Analysis and Statistical Pattern Recognition, Wiley, New York, 1992, https://doi.org/10.1002/0471725293.

[49] S. Wold, M. Sjöström, L. Eriksson, PLS-regression: a basic tool of chemometrics, Chemometr. Intell. Lab. Syst. 58 (2) (2001) 109–130, https://doi.org/10.1016/s0169-7439(01)00155-1.

[50] S. Rankin-Turner, C.J. McMeniman, A headspace collection chamber for whole body volatilomics, The Analyst 147 (22) (2022) 5210–5222, https://doi.org/10.1039/d2an01227h.