Data in Brief 54 (2024) 110532

Contents lists available at ScienceDirect

Data in Brief

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

Data Article

Gas chromatography ion mobility spectrometry (GC-IMS) dataset of honey samples with different botanical origins

Joscha Christmann[a,b], Sascha Rohn[b], Philipp Weller[a,*]

[a] *Institute for Instrumental Analytics and Bioanalysis, Mannheim University of Applied Sciences, Paul-Wittsack-Straße 10, 68163 Mannheim, Germany*
[b] *Department of Food Chemistry and Analysis, Institute of Food, Technology and Food Chemistry, Technische Universität Berlin, TIB 4/3-1, Gustav-Meyer-Allee 25, 13355 Berlin, Germany*

ARTICLE INFO

Article history:
Received 2 April 2024
Revised 6 May 2024
Accepted 13 May 2024
Available online 17 May 2024

Dataset link: Honeys biological origin by GC-IMS (Original data)

Keywords:
Non-target screening
Volatile organic compounds
Food fraud detection
Chemometrics
Multivariate data analysis

ABSTRACT

Gas chromatography ion mobility spectrometry (GC-IMS) is a robust and sensitive benchtop technique commonly used for non-target screening of volatile organic compounds. It has been applied to authenticity analysis by generating characteristic "fingerprints" of food samples, well suited for chemometric data analysis. This dataset contains headspace GC-IMS spectra from 50 monofloral honey samples from three different botanical origins, 18 acacia honeys (*Robinia pseudoacacia*), 19 canola honeys (*Brassica napus*) and 18 honeydew honeys (forest flowers). Honeys were sourced from the beekeepers directly or obtained from governmental food inspectors from Baden-Wuerttemberg, Germany. Authenticity was confirmed by pollen analysis in the framework of the official control of foodstuffs. The data was acquired using a setup based on an Agilent 6890N gas chromatograph (Agilent Technologies, Palo Alto, CA) and an OEM Standalone IMS cell from G.A.S Sensorsysteme m. b. H. (Dortmund, Germany). All samples were recorded in duplicates and spectra are presented as raw data in the *.mea* file format. The dataset is available on Mendeley Data: https://data.mendeley.com/datasets/jxj2r45t2x.

* Corresponding author.
E-mail address: p.weller@hs-mannheim.de (P. Weller).

https://doi.org/10.1016/j.dib.2024.110532

Specifications Table

Subject	Analytical Chemistry
Specific subject area	Non-target screening, Food fraud detection, Chemometrics, Multivariate data analysis, Honey, Volatile organic compounds, Gas chromatography
Type of data	Binary files .mea format (G.A.S Sensorsysteme m. b. H., Dortmund, Germany), Raw data
Data collection	The data was acquired by headspace gas chromatography hyphenated with ion mobility spectrometry (HS-GC-IMS). The setup was based on an Agilent 6890N gas chromatograph (Agilent Technologies, Palo Alto, CA) and an OEM Standalone IMS cell from G.A.S Sensorsysteme m. b. H., Dortmund, Germany. Measurements were carried out in duplicates.
Data source location	Mannheim University of Applied Sciences, Institute for Instrumental Analysis and Bioanalytics, Faculty of Biotechnology, 68163 Mannheim, Germany
Data accessibility	Weller, Philipp (2022), "Honeys biological origin by GC-IMS", Mendeley Data, V1, doi: 10.17632/jxj2r45t2x.1 https://data.mendeley.com/datasets/jxj2r45t2x
Related research article	J. Christmann, S. Rohn, P. Weller, Finding features - variable extraction strategies for dimensionality reduction and marker compounds identification in GC-IMS data, Food Research International 161 (2022) 111779. https://doi.org/10.1016/j.foodres.2022.111779.

1. Value of the Data

- The data are of interest for food chemists working on fraud detection of honey.
- The dataset has been analysed extensively in previous publications and can therefore be used to benchmark new statistical methods or implementations against known results.
- The data can be used as additional training data for machine learning models to predict the botanical origin of honey or to pretrain a model with GC-IMS data for a different classification problem.
- The dataset can be beneficial for software developers to develop, test and teach software tools for GC-IMS data analysis.
- The data is also a helpful basis for researchers, who want to dive into data analysis (e.g. with gc-ims-tools) and want to test drive a documented and evaluated dataset

2. Background

The original motivation behind the dataset was the development of an analytical approach to confirm the claimed botanical origins of honey samples [1]. The dataset was further used as example and test data for the development of new Python based software tools and data analysis methods as explained in the related research article. The new data analysis workflows were benchmarked against previously obtained results [2].

3. Data Description

The dataset consists of raw headspace GC-IMS data of 55 monofloral honey samples from various European countries. The samples were 18 acacia honeys (*Robinia pseudoacacia*), 19 canola honeys (*Brassica napus*) and 18 honeydew honeys (forest flowers) listed in Table 1. Honeys were sourced from the beekeepers directly or governmental food inspectors from Baden-

Table 1
Sample list.

Acacia	Canola	Honeydew
Akazien_DE,RO,HU_001	Raps_ EU-N-EU_001	Waldhonig_140616237
Akazien_DE_001b	Raps_BG,RO_001	Waldhonig_CH_H014
Akazien_DE_002	Raps_CL_001	Waldhonig_DE,IT_150383242
Akazien_DE_003	Raps_DE_001a	Waldhonig_DE_001a
Akazien_EU-N-EU_150419327	Raps_DE_001b	Waldhonig_DE_001b
Akazien_EU-N-EU_150567717	Raps_DE_140681199	Waldhonig_DE_140642523
Akazien_EU-N-EU_150567761	Raps_DE_150181273	Waldhonig_DE_150334764
Akazien_FR_001	Raps_DE_150221822	Waldhonig_DE_150380066
Akazien_FR_002	Raps_DE_150298553	Waldhonig_DE_150424564
Akazien_FR_H034	Raps_DE_150410376	Waldhonig_EU-N-EU_001a
Akazien_HR_001	Raps_DE_150515236	Waldhonig_EU-N-EU_001b
Akazien_HU_150558770	Raps_DE_150539325	Waldhonig_EU-N-EU_150177361
Akazien_HU_150583681	Raps_DE_150541561	Waldhonig_EU-N-EU_150356575
Akazien_HU_150595533	Raps_DE_150573236	Waldhonig_EU-N-EU_150380067
Akazien_HU_H046	Raps_EU-N-EU_150138828	Waldhonig_IT,CZ_001
Akazien_IT_150427710	Raps_EU-N-EU_150214363	Waldhonig_IT_001a
Akazien_MD_H038	Raps_PO_001	Waldhonig_IT_140673677
Akazie_CY	Raps_PO_002	Waldhonig_TR_140689478
-	Raps_PO_003	-

Wuerttemberg, Germany. The botanical origin of each sample was confirmed by pollen analysis in the framework of the official control of foodstuffs.

All samples were measured in duplicates for a total of 100 GC-IMS spectra in the dataset, saved as binary *.mea* file from G.A.S Sensorsysteme m. b. H., Dortmund, Germany. A free and open-source file reader is available in the *gc-ims-tools* Python package [3]. An example spectrum is shown in Fig. 1. In each spectrum the retention time ranges to 17 minutes and consists of 6939 data points. Correspondingly, the drift time (21 ms) is represented by 3150 data points [4].

4. Experimental Design, Materials and Methods

Analysis was carried out using an Agilent 6890N gas chromatograph (Agilent Technologies, Palo Alto, CA) hyphenated with an OEM Standalone IMS cell from G.A.S Sensorsysteme m. b. H., Dortmund, Germany. A CTC CombiPal autosampler (CTC Analytics AG, Zwingen, Switzerland) equipped with a 2.5 mL gastight syringe (Gerstel GmbH, Muehlheim, Germany) set to 80°C, to avoid condensation, was used for headspace generation and sample injection. A straight headspace glass liner with 1.2 mm i.d. (Agilent, Waldbronn, Germany) was installed to optimize the transfer onto the column. A 25m × 0.32 mm × 0.25 µm DB-225 (25% phenyl, 25% cyanopropyl methyl siloxane) capillary column (Agilent Technologies, Santa Clara, CA) was used. The carrier gas was nitrogen of 99.99% purity at a constant flow rate of 1.5 mL/min.

The IMS cell was connected to the GC with a heated transfer line set to 120°C. A ^{3}H ion source with 300 MBq activity was used to ionize the analyte molecules and the IMS was operated in positive ion mode. The 10 cm drift tube was set to a constant voltage of 5 kV and a temperature of 90°C. Nitrogen was also used as the drift gas with a constant flow of 150 mL/min controlled by a mass-flow controller (Voegtlin Instruments AG, Aesch, Switzerland). The injection pulse width was set to 150 µs, the sampling frequency to 150 kHz, the repetition rate to 21 µs and the blocking and injection voltages to 70 mV and 2500 mV respectively.

For the analysis of the honey samples 2 g were mixed with 2 mL of a saturated sodium chloride solution (VWR International GmbH, Darmstadt, Germany, and purified water using a Milli-Q system, Millipore, Bedford, MA). The samples were spiked with 18 µL of a 1008 mg/L 2-acetylpyridine stock solution as internal standard. The mix was incubated in a 20 mL headspace vial at 45°C for 15 min and 700 µL headspace volume were injected. After each analysis the

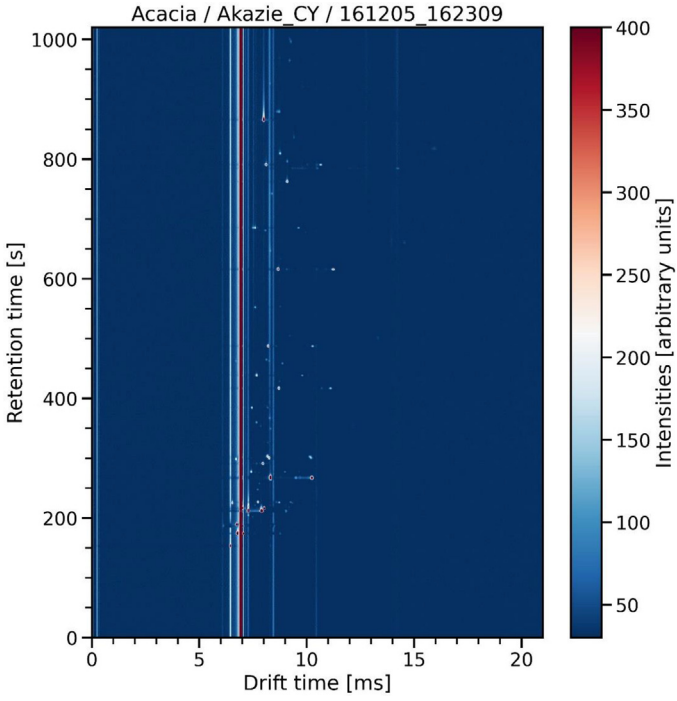

Fig. 1. Example GC-IMS spectrum of an acacia honey sample.

syringe was flushed with nitrogen for 2 min to avoid memory effects. The GC oven was initially held at 40°C for 2 min before ramping up to 120°C at 8°C/min and held at 120°C for further 10 min [1,5,6].

Lmitations

Not applicable.

Ethics Statement

Not applicable.

Data Availability

Honeys biological origin by GC-IMS (Original data) (Mendeley Data).

CRediT Author Statement

Joscha Christmann: Data curation, Writing – original draft, Visualization, Software; **Sascha Rohn:** Writing – review & editing; **Philipp Weller:** Writing – review & editing, Funding acquisition, Supervision, Conceptualization.

Acknowledgements

The authors would like to acknowledge the Chemisches und Veterinäruntersuchungsamt (CVUA) Karlsruhe and Freiburg for supplying honey samples and the supporting data of pollen analysis. We gratefully acknowledge the Center for Applied Research in Biomedical Mass Spectrometry (ABIMAS) for financial support.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of Competing Interests

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.

References

[1] N. Gerhardt, M. Birkenmeier, S. Schwolow, S. Rohn, P. Weller, Volatile-compound fingerprinting by headspace-gas-chromatography ion-mobility spectrometry (HS-GC-IMS) as a benchtop alternative to 1H NMR profiling for assessment of the authenticity of honey, Anal. Chem. 90 (2018) 1777–1785, doi:10.1021/acs.analchem.7b03748.

[2] J. Christmann, S. Rohn, P. Weller, Finding features - variable extraction strategies for dimensionality reduction and marker compounds identification in GC-IMS data, Food Res. Int. 161 (2022) 111779, doi:10.1016/j.foodres.2022.111779.

[3] J. Christmann, S. Rohn, P. Weller, GC-IMS-tools – a new python package for chemometric analysis of GC–IMS data, Food Chem 224 (2022) 133476, doi:10.1016/j.foodchem.2022.133476.

[4] Philipp Weller, Honeys biological origin by GC-IMS, Mendeley (2022).