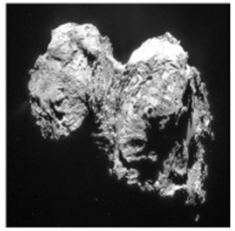
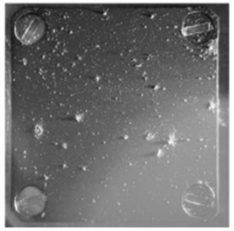


Comet 67P, ~6 km wide

Composition of cometary particles *versus* distance to sun during sample collection

- based on multivariate evaluation of mass spectral data (Rosetta/COSIMA)

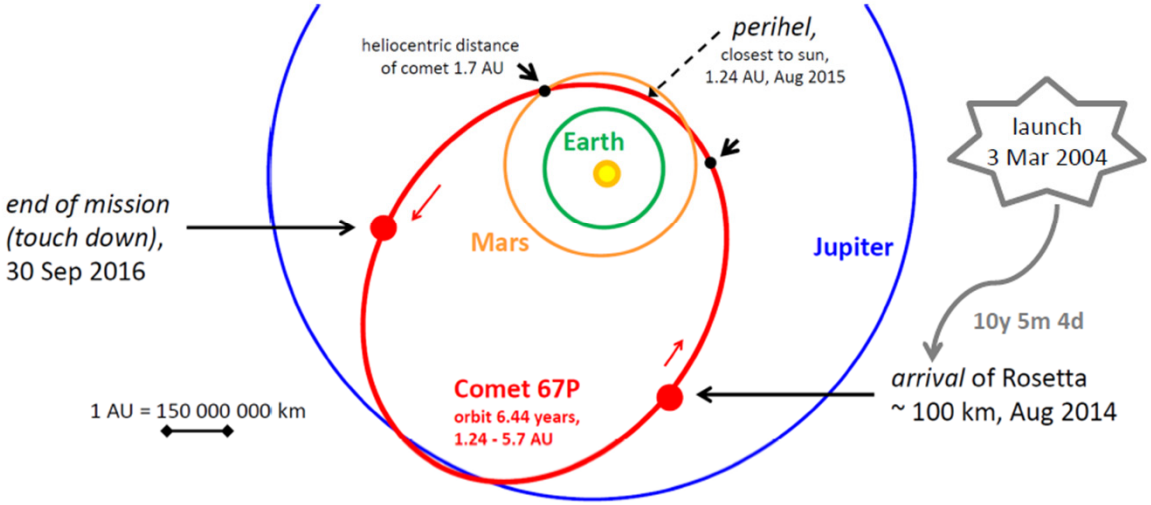


Gold target 1 cm x 1 cm with cometary particles

Varmuza Kurt [1*], Filzmoser Peter [1], Fray Nicolas [2], Cottin Hervé [2], Merouane Sihane [3], Stenzel Oliver [3], Kissel Jochen [3], Briois Christelle [4], Baklouti Donia [5], Bardyn Anais [6], Siljeström Sandra [7], Silén Johan [8], Hilchenbach Martin [3]

[1] **Vienna University of Technology**, Institute of Statistics and Mathematical Methods in Economics, Research Unit Computational Statistics, Vienna, Austria
[2] Laboratoire Interuniversitaire des Systèmes Atmosphériques, Université Paris Est Créteil et Université Paris Diderot, Créteil, France
[3] Max Planck Institute for Solar System Research, Göttingen, Germany
[4] Laboratoire de Physique et Chimie de l'Environnement et de l'Espace, Université d'Orléans et du CNES, Orléans, France
[5] Institut d'Astrophysique Spatiale, Université Paris Sud, Orsay, France
[6] DTM, Carnegie Institution of Washington, Washington, DC, USA
[7] Bioscience and Materials, Research Institute of Sweden, Stockholm, Sweden
[8] Finnish Meteorological Institute, Helsinki, Finland

Rosetta was a space mission (ESA) working more than two years (Aug 2014 to Sep 2016) near **comet** *Churyumov-Gerasimenko* (short *67P*).



COSIMA was an instrument on board of Rosetta.
It collected **cometary particles** with size 10 - 1000 µm on metal targets (distance from the comet 10 - 1500 km). More than 30,000 particles were documented by images, and more than 35,000 secondary ion mass spectra (**TOF-SIMS**) were measured on about 250 particles [1 – 7].

Particle (sample) collection occurred at heliocentric distances (from comet to sun) between 1.8 AU and 3.8 AU (AU for *Astronomical Unit*, 150 million km, approximately the mean distance between earth and sun).

Aim of this work was to check relationships between the **heliocentric distance of sampling** and **mass spectral data**.

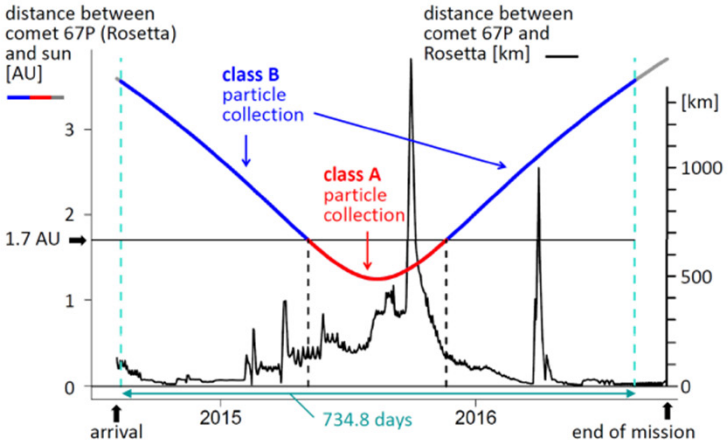
Data

n = 2863 *objects* derived from TOF-SIMS spectra measured at 256 particles, collected in 412 sampling intervals

n_1 = 1459 in **class A**: collection at ≤1.7 AU (*near* sun, warmer)

n_2 = 1404 in **class B**: collection at >1.7 AU (*far* from sun, cooler)

m = 9 variables: ion counts for C^+, CH^+, CH_2^+, CH_3^+, Mg^+, $C_2H_3^+$, $C_3H_3^+$, $C_3H_4^+$, Fe^+, normalized to sum 100 or transformed by *centered log-ratios (clr)* [compositional data].



Strategy for Data Evaluation

❍ Characterization of the **discrimination** of **class A** (particles collected closer to sun) and **class B** (more far from sun).

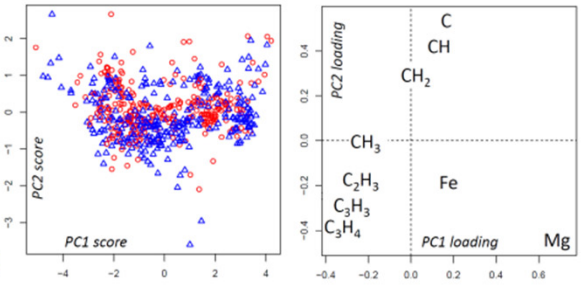
❍ A **good discrimination indicates** a **relationship** between the used mass spectral data (characterizing the chemical composition) and the heliocentric distance of sampling.

❍ Methods applied: PCA; univariate distributions; KNN classification; PLS calibration; the last two with rdCV, repeated double cross validation [8].

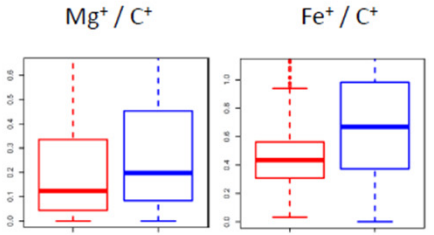
Results

PCA

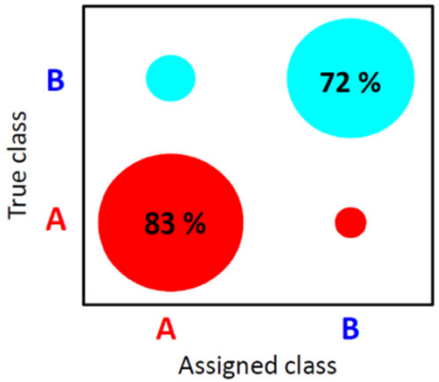
n_1 = 300
n_2 = 300
(randomly)
m = 9
clr ransf.
Variances of PC1, 2: 82.3, 10.9%

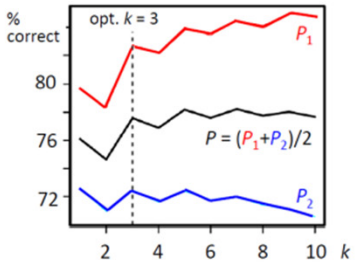


Distribution of ratios

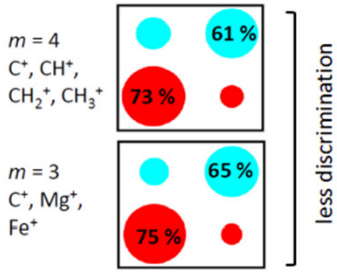


KNN classification





Repeated double cross validation: sum 100; no. segments 3 (out), 4 (in); 50 repetitions; k = 1 ... 10



PLS models (linear) *distance* = f (*spec data*) are not appropriate.

Conclusions

❍ The mass spectral data (characterizing the composition of the particles) are multivariate different for **class A** (collection near aphelion) and **class B** (collection more far from sun).

❍ KNN separates **A** and **B** well; linear approaches (PCA, PLS) do not.

❍ Some ion count ratios show different distributions for **A** and **B**; single ions do not.

❍ Influence of the **distance to the comet** or other parameters cannot be excluded.

[1] Kissel J., et al.: *Space Sci. Rev.*, **128**, 823 (2007)
[2] Hilchenbach M., et al.: *The Astrophysical Journal Letters*, **816**: L32 (2016)
[3] Fray N., et al.: *Nature*, **528**, 72 (2016)
[4] Bardyn A., et al.: *MNRAS*, **469,** Suppl_2, S712 (2017)
[5] Stenzel O. et al.: *MNRAS*, **469,** Suppl_2, S492 (2017)
[6] Isnard R. et al.: Astron. Astroph., aa34797-18 (2019)
[7] Altwegg K. et al.: *Ann. Rev. Astron. Astrophys.* **57**, 113 (2019)
[8] Varmuza K., et al.: *J. Chemom.*, **32**, e3001, 1 (2018)

Supported by the Austrian Science Fund (FWF), project P 26871 - N20.
The authors thank the other members of the **COSIMA team** for their contributions.

190904

* kurt.varmuza@tuwien.ac.at, http://www.lcm.tuwien.ac.at/vk/

Poster at *Conferentia Chemometrica* (CC2019), 8 - 12 Sep 2019, Karcag, Hungary

Abstract

Composition of cometary particles *versus* distance to sun during sample collection - based on multivariate evaluation of mass spectral data (Rosetta/COSIMA)

Varmuza K.*[1], Filzmoser P.[1], Fray N.[2], Cottin H.[2], Merouane S.[3], Stenzel O.[3], Kissel J.[3], Briois C.[4], Baklouti D.[5], Bardyn A.[6], Siljeström S.[7], Silén J.[8], Hilchenbach M.[3]

[1] Vienna University of Technology, Institute of Statistics and Mathematical Methods in Economics, Research Unit Computational Statistics, Vienna, Austria
Email: kurt.varmuza@tuwien.ac.at
[2] Laboratoire Interuniversitaire des Systèmes Atmosphériques, Université Paris Est Créteil et Université Paris Diderot, Créteil, France
[3] Max Planck Institute for Solar System Research, Göttingen, Germany
[4] Laboratoire de Physique et Chimie de l'Environnement et de l'Espace, Université d'Orléans et du CNES, Orléans, France
[5] Institut d'Astrophysique Spatiale, Université Paris Sud, Orsay, France
[6] DTM, Carnegie Institution of Washington, Washington, DC, USA
[7] Bioscience and Materials, Research Institute of Sweden, Stockholm, Sweden
[8] Finnish Meteorological Institute, Helsinki, Finland

./. next page

The instrument COSIMA [1] onboard of the ESA spacecraft Rosetta collected dust particles in the neighborhood of comet Churyumov-Gerasimenko [2]. The distance to the sun during the sample collections varied between 1.2 and 3.8 AU (AU for astronomical unit, 150 000 000 km). The chemical composition of the particle surfaces was characterized by COSIMA using TOF-SIMS (time-of-flight secondary ion mass spectrometry). A set of about 3000 spectra has been selected, and relative abundances for CH-containing positive ions (from organic compounds [3-6]) as well as elemental ions (from minerals [7]) define a set of multivariate data. Evaluation by chemometric techniques indicates different compositions of samples collected at different distances to the sun. The applied methods comprise (1) considering the compositional nature of the used mass spectral data and centered log-ratio transformation [8]; (2) robust PCA [9]; (3) KNN-classification with repeated double cross validation [10].

Supported by the Austrian Science Fund (FWF), project P 26871 - N20.

[1] Kissel J., et al.: *Space Sci. Rev.*, **128**, 823 (2007)
[2] Schulz R., et al.: *Nature*, **518**, 216 (2015)
[3] Fray N., et al.: *Nature*, **528**, 72 (2016)
[4] Bardyn A., et al.: *MNRAS,* **469,** Suppl_2**,** S712 (2017)
[5] Fray N., et al.: *MNRAS,* **469,** S506 (2017)
[6] Varmuza K., et al.: *J. Chemometrics,* **32**, e3001 (2018)
[7] Stenzel O. et al.: *MNRAS*, **469,** Suppl_2, S492 (2017)
[8] Filzmoser P., Hron K., Templ M.: *Applied compositional data analysis*, Springer Nature, Cham, Switzerland (2018)
[9] Hubert M. et al.: *Technometrics*, **47**, 64 (2005)
[10] Varmuza K., Filzmoser P.: In Khanmohammadi M. (ed.), *Current applications of chemometrics*, p. 15-31; Nova Science Publishers, New York, USA (2015)