Chemometrics for Mass Spectral Data Related to Comet Research

Kurt Varmuza

Vienna University of Technology
Institute of Chemical Engineering
Laboratory for ChemoMetrics

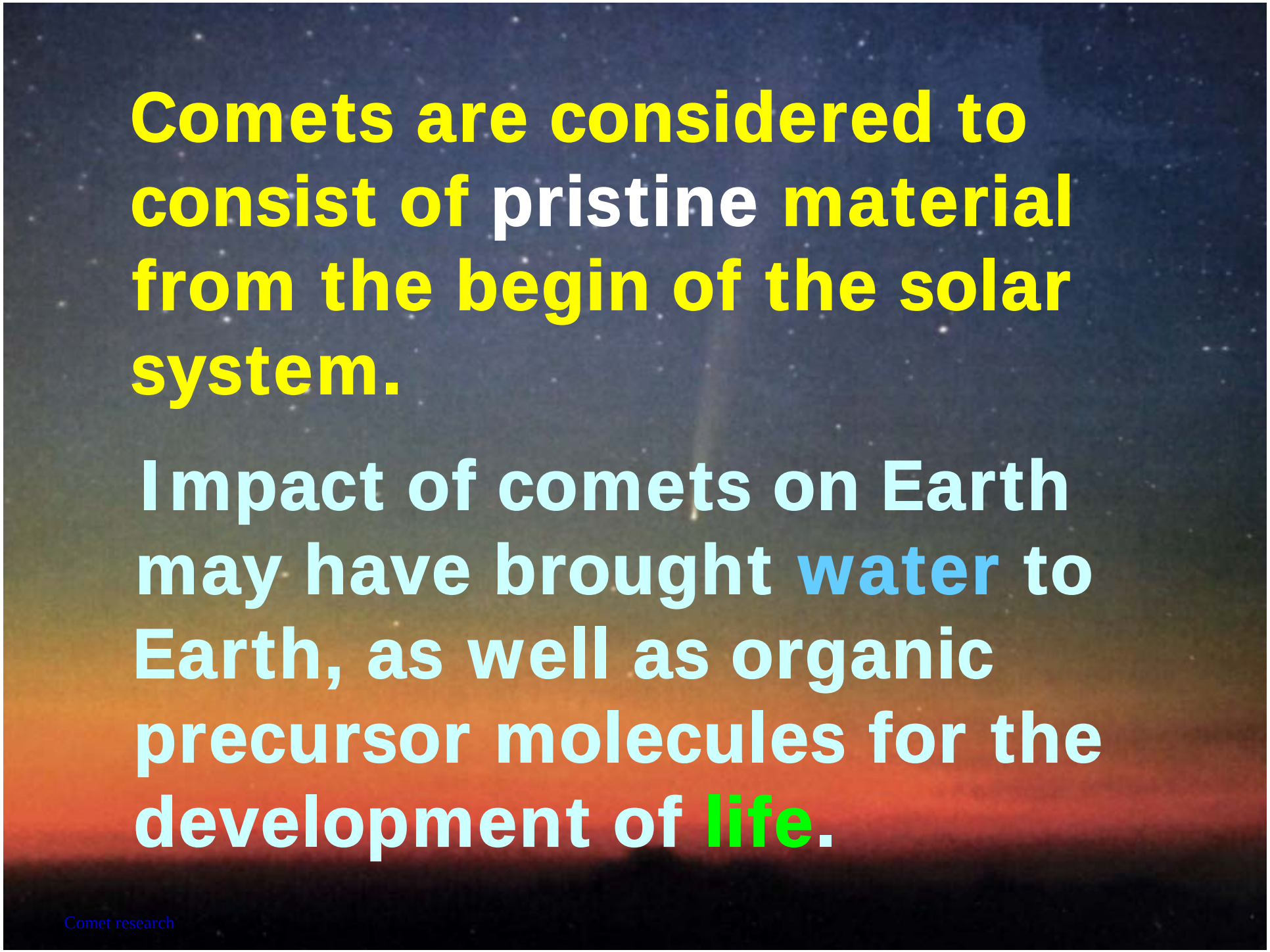
www.lcm.tuwien.ac.at

European User Meeting on Multivariate Analysis, Como, 22 June 2006

Contents

- Introduction
- Method and Instrument
- Inorganics
- Organics

This file contains a subset of the pictures shown in the lecture.


Comets are considered to consist of pristine material from the begin of the solar system.
Impact of comets on Earth may have brought water to Earth, as well as organic precursor molecules for the development of life.

Comet material consist of

- mineral grains,
- volatile organic material,
- complex condensed C,H,N,O-compounds,
- water (ice).

Results from fly-by mission at comets Halley, Hale-Bopp, and Wild-2.

ROSETTA

First spacecraft to orbit a comet,

2.8 x 2.1 x 2.0 m, two 14 m solar panels,
launch mass 3000 kg,

11 instruments in orbiter (165 kg),
9 instruments in lander (100 kg),
propellant (1600 kg)

Cosima

TOF-SIMS mass spectrometer

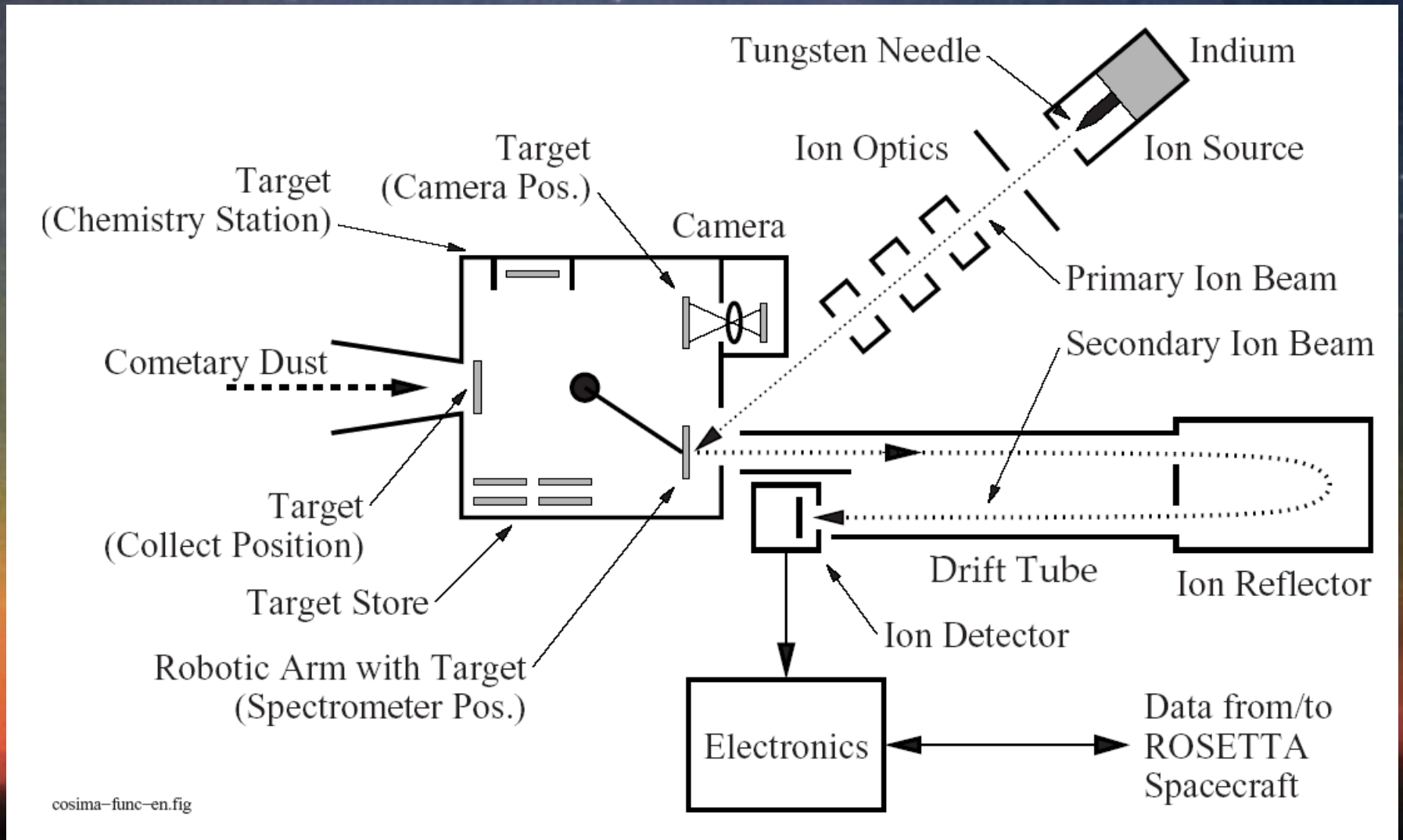
http://rosetta.esa.int

Comet *Chury* !

Name	67P/Churyumov-Gerasimenko
Discovery	1969 K. Churyumov (Univ. Kiev, Ukraine) S. Gerasimenko (Inst. Astrophysics, Dushanabe, Tajikistan)
Diameter (nucleus)	ca 4 km
Orbital period	6.6 years
Distance from Sun	186 - 857. 10^6 km (1.2 - 5.7 AU)

Cometary Secondary Ion Mass Spectrometer

Secondary Ion - Time-of-Flight Mass Spectrometer



Cosima scheme

Cometary Secondary Ion Mass Spectrometer

DATA ?

- Hopefully in year 2014
- First TOF-SIMS data from space
- Data from similar instruments

some minerals	some organic compounds

Chemometric Evaluation of TOF-SIMS Data from Minerals

Engrand C.	Orsay	France
Kissel J.	Katlenburg-Lindau	Germany
Krueger F.R.	Darmstadt	Germany
Martin P.	Orléans	France
Silén J.	Helsinki	Finland
Thirkell L.	Orléans	France
Thomas R.	Orléans	France
Varmuza K.	Vienna	Austria




Rapid Commun. Mass Spectrom. 2006; **20**: 1361–1368

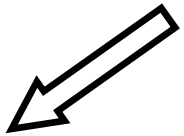
TOF-SIMS of minerals: Samples

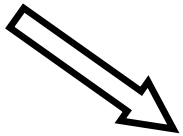
Class i	Mineral	Chemical formula	n_i
1	**Serpentine**	$(Mg, Fe)_3 Si_2 O_5 (OH)_4$	12
2	**Enstatite**	$Mg Si O_3$	9
3	**Olivine**	$(Mg, Fe)_2 Si O_4$	9
4	**Talc**	$Mg_3 Si_4 O_{10} (OH)_2$	9
		sum n	**39**

Mineral grains: 10 - 100 μm, embedded in gold foil (substrate).
These minerals have been identified **in micro meteorites**;
enstatite and olivine also **in comet** Hale-Bopp (Wooden et al., 1999, 2000).

TOF-SIMS of minerals: Peak selection

529 peaks measured (m/z 1.00718 ... 1089.63)





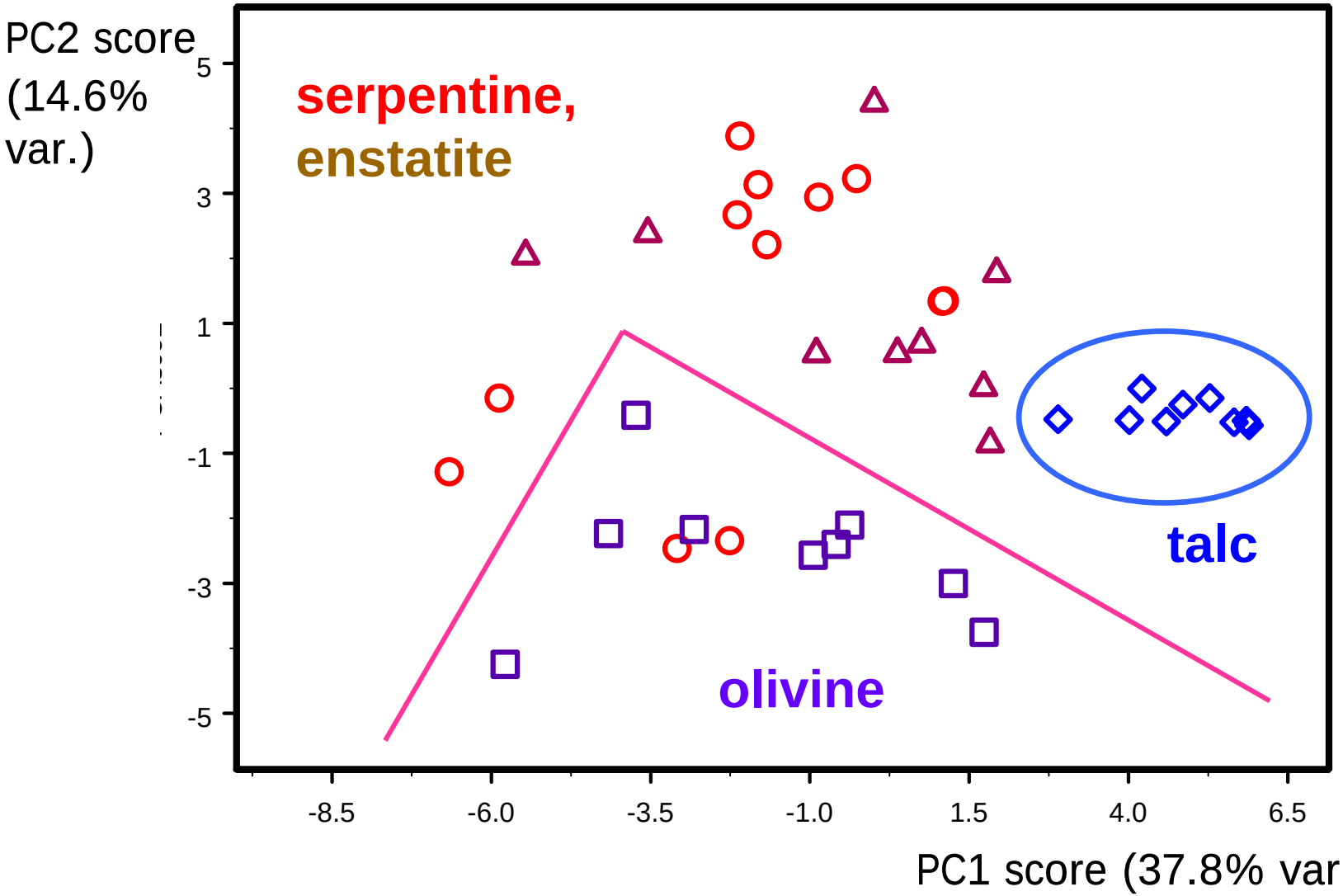
inorganic ions,
organic ions

ions from mineral samples,
ions from contaminations,
ions from substrate (Au)

Peak (feature) selection

- all peaks (overall characterization)
- peaks with inorganic (mineralogical) origin
- peaks from investigated minerals

TOF-SIMS of minerals: PCA (p = 32)



TOF-SIMS of minerals: KNN (p = 32)

KNN classification of mineral type

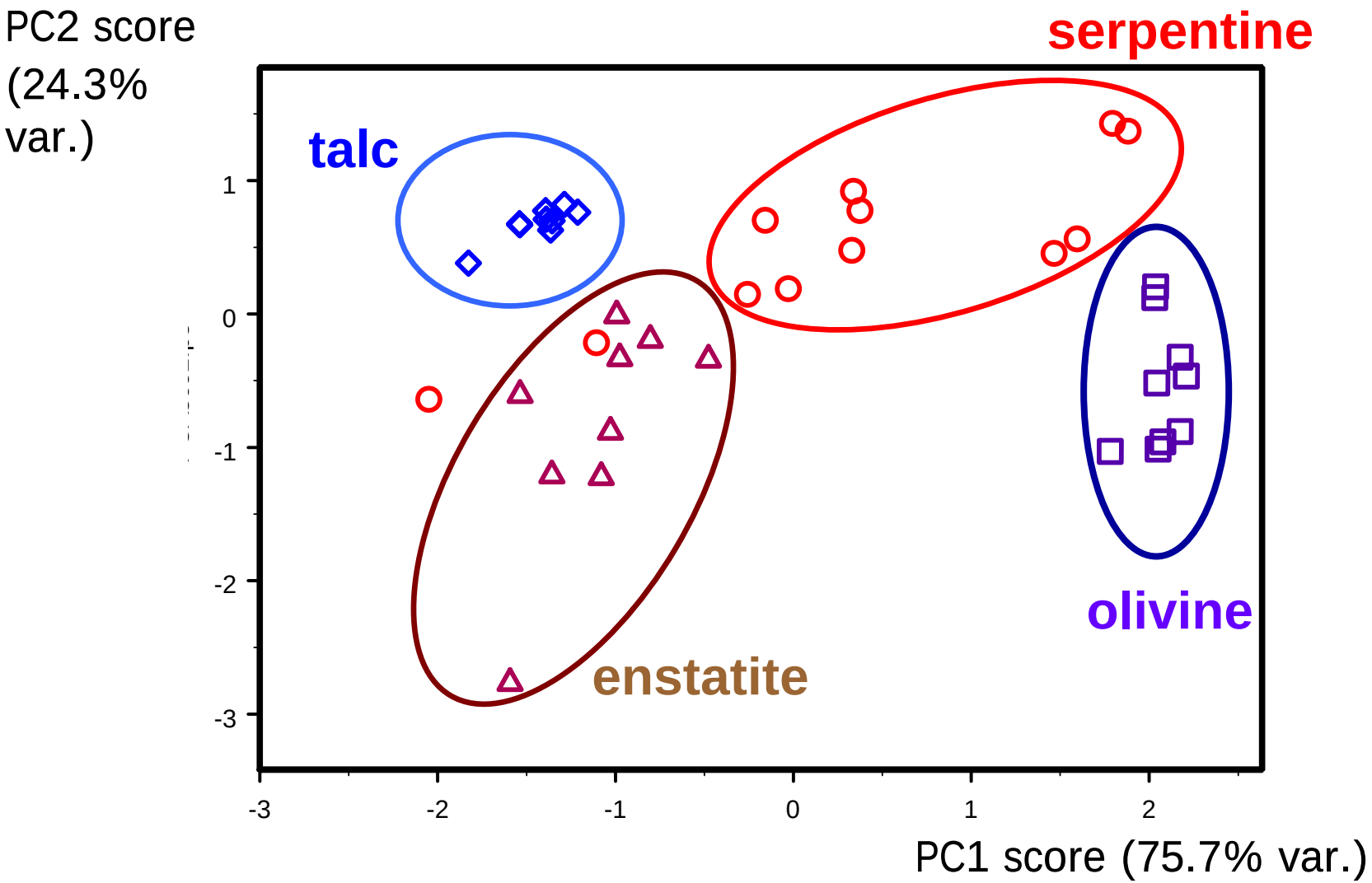
n = 39 (4 classes), p = 32 (normalized to constant sum), Euclidean distance, optimum k is 1 (leave-one-out)

		Number of objects in predicted class			
Correct class		1	2	3	4
1	Serpentine	11	0	1	0
2	Enstatite	0	8	0	1
3	Olivine	1	0	8	0
4	Talc	0	0	0	9

Total predictive ability = 92.3 % (leave-one-out cross validation)

TOF-SIMS of minerals: PCA (p = 3)

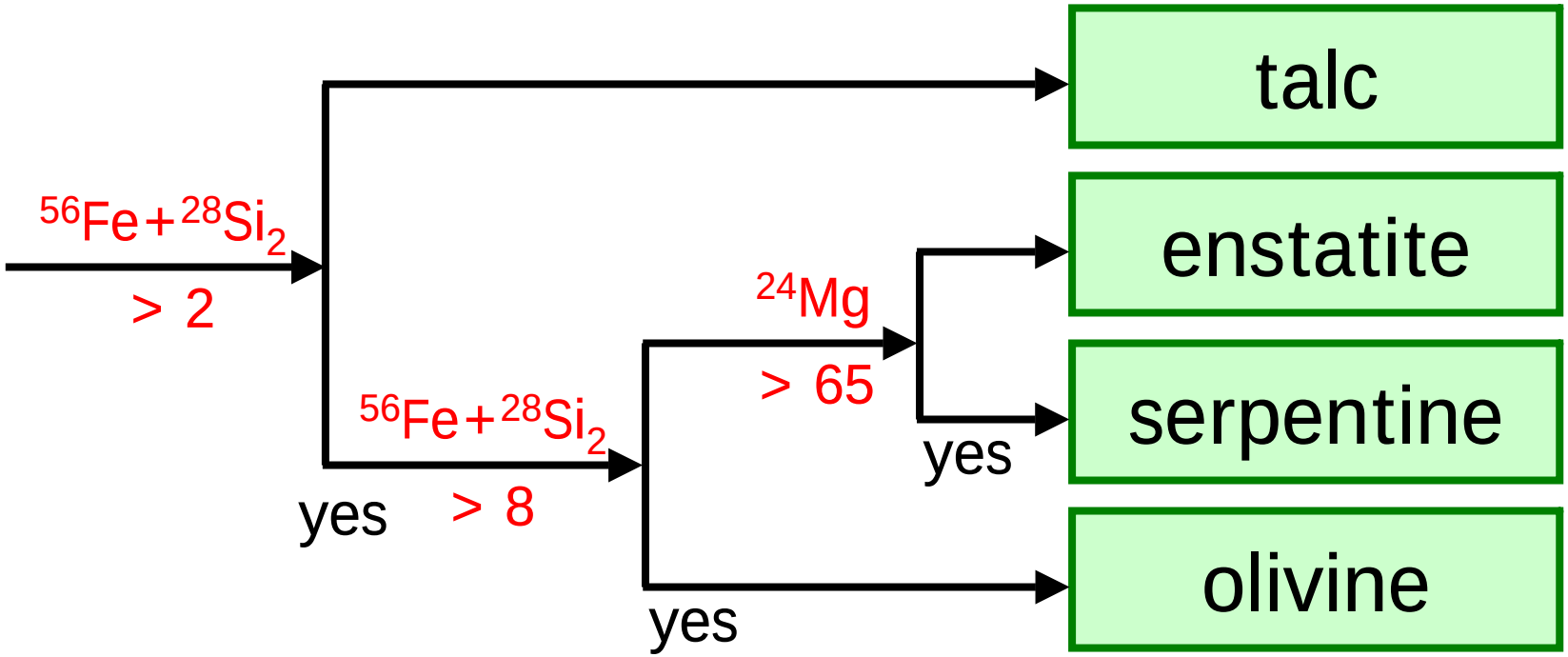
$^{24}Mg^+$, $^{28}Si^+$, $^{56}Fe^+ + {}^{28}Si_2^+$, sum = 100



TOF-SIMS of minerals: CART (p = 3)

Binary decision tree (CART)

n = 39 (4 classes), p = 3 (normalized to sum 100),
Software SCAN (Minitab)



Total predictive ability = 92.3 % (leave-one-out cross validation)

Preliminary Conclusions

- The four investigated minerals can be discriminated by their TOF-SIMS data (PCA, KNN, decision tree).
- Data used: Relative intensities of inorganic, mono-atomic ions up to m/z 57.
- What has main influence on ion yield ?
  - Elemental composition ?
  - Local arrangement of atoms ?

The number of spectra very small ☹

Organic chemical substances in comet dust particles ?!

Demuth W.	Vienna	Austria
Kissel J.	Katlenburg-Lindau	Germany
Krueger F.R.	Darmstadt	Germany
Schmid E.R.	Vienna	Austria
Varmuza K.	Vienna	Austria
Werther W.	Vienna	Austria

Werther, Demuth, Krueger, Kissel, Schmid, Varmuza: J. Chemom. **16**, 99 (2002)

Organic chemical substances in comet dust particles ?!

Experimental data, data interpretation, speculation, ...

Comet *Halley* encounter (Giotto, Vega, 1986), mass spectra, UV emission spectra

Comet *Wild* encounter (Stardust, 2004), mass spectra

Comet material contains:

Highly unsaturated organic compounds, containing N and some O, ev. S;

e.g. polycyclic aromatic compounds, pyrimidines, chinones, and similar ones.

Krueger F.R., Kissel J.

Organic chemical substances in comet dust particles ?!

n = 61 reference compounds (condensed benzene rings, N-aromatic, purines, pyrimidines, ...)

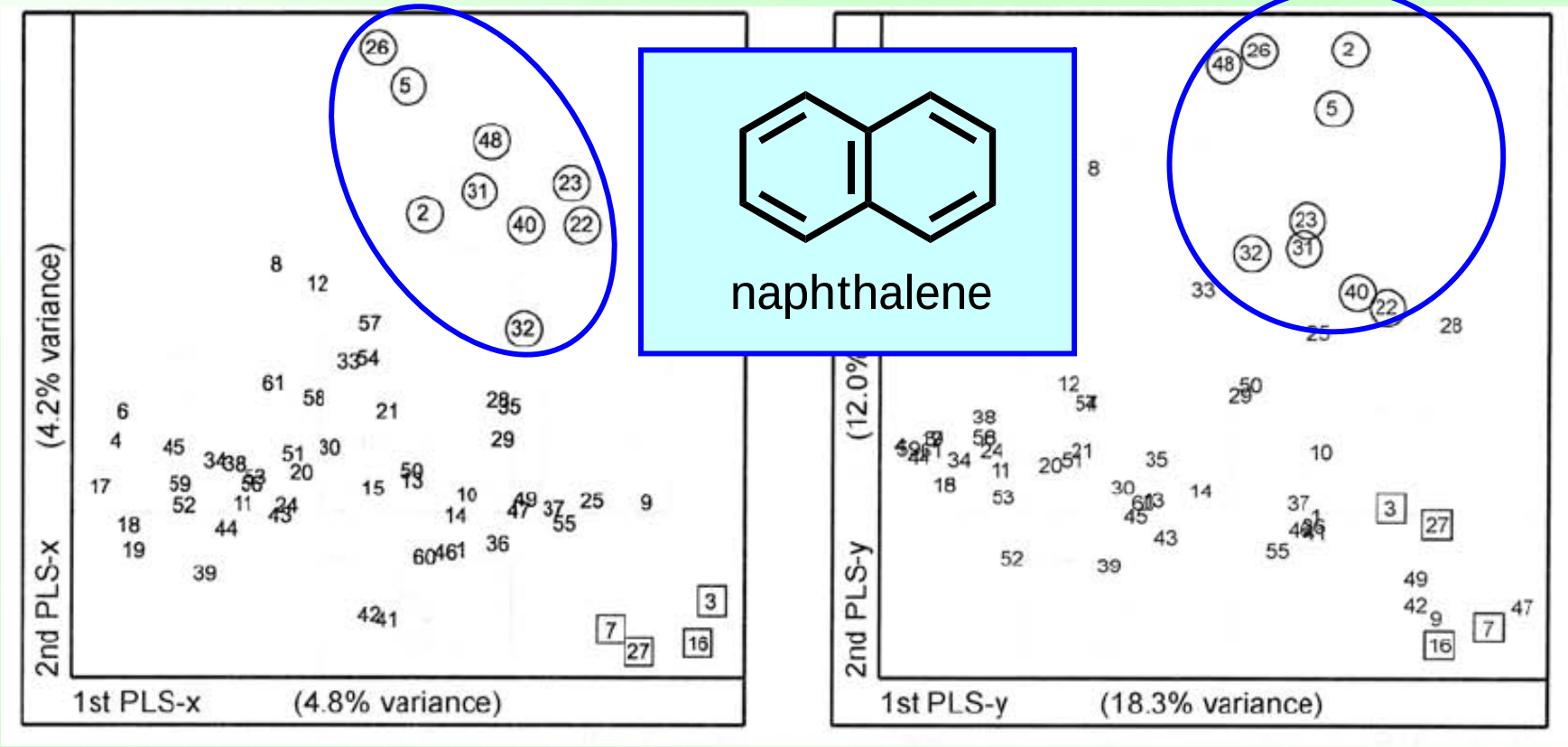
Mass spectrum	300 spectral features [0 ... 100]
Chemical structure	35 binary substructure descriptors for atom-centered fragments [0, 1]

Werther, Demuth, Krueger, Kissel, Schmid, Varmuza: J. Chemom. **16**, 99 (2002)

Data: X,Y

Organic chemical substances in comet dust particles ?!

PLS mapping (positive ion data, molecular descriptors)



Werther, Demuth, Krueger, Kissel, Schmid, Varmuza: J. Chemom. **16**, 99 (2002)

Organic chemical substances in comet dust particles ?!

Preliminary Conclusions

- TOF-SIMS data - if transformed to appropriate spectral features - contain chemical structure information.
- PLS mapping (***X*** = spectra, ***Y*** = structures) can be used for a visualization of spectra-structure relationships.
- Open question: Applicability to data from comet material (expected to be a very complex mixture).

Software **MassFeatGen**

Transformation of low resolution mass spectra (peak lists) into numerical spectral features (modulo-14 sums, autocorrelation, spectra type, ...)

Software **SubMat**

Generation of binary substructure descriptors for organic compounds (molecular structures, substructures in Molfile format)

User Guides and Demo Versions free

www.lcm.tuwien.ac.at [Software]